

Elucidating the Mechanism of Action of Antimicrobial Peptides by means of Computational Approaches

vorgelegt von

Lic. Chem.

Alejandra de Miguel Catalina

geb. in Madrid, Spanien

von der Fakultät II - Mathematik und Naturwissenschaften

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

-Dr. rer. nat.-

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Roderich Süßmuth

Gutachterin: Prof. Dr. Maria Andrea Mroginski

Gutachter: Prof. Dr. Karl Lohner

Tag der wissenschaftlichen Aussprache: 06. November 2017

Berlin 2017

Abstract

Microorganism such as bacteria, viruses, fungi and some parasites, possess the inherent ability to cease the activity of an antimicrobial -for instance antibiotics, antivirals and antimalarials- against it. This property is known as antimicrobial resistance (AMR) and causes with time the ineffectiveness of standard treatments, helping infections persist and spread to other hosts. AMR is one of the biggest threats to global health, affecting every country and individual citizen. Hospital stays due to afflictions are prolonged, causing higher medical costs and increased mortality. This year (February 2017), the World Health Organization (WHO) published for the first time ever a list of antibiotic-resistant 'priority pathogens', which include selected families of bacteria that present the highest threat to the human health, ordered according to the urgency of required new antibiotics to fight AMR. Considering that just in Europe, 25,000 people die from a 'superbug' per year, the priority of designing new efficient antibiotics is, undoubtedly, an imperious worldwide issue.

Thus, as an alternative to the prevailing antibiotics, the attention has been focused on the antimicrobial peptides (AMPs). Numerous studies have described their efficacy against multidrug-resistant bacteria. AMPs, also named as host defense peptides (HDPs), have been found in all multicellular living organisms. AMPs have apart of antimicrobial properties, important roles in intracellular processes. Unfortunately, there is insufficient knowledge regarding the mechanism of action to unleash their bactericidal effects. With the aim of shedding light at a molecular level of the underlying antimicrobial action mechanism a combined theoretical approach consisting of i) all-atom molecular dynamics (MD) simulations and ii) development of Markov State models (MSMs) has been applied. From the more than 2,000 naturally and synthetic AMPs reported up-to-date, two different families were selected in the current study. First, bacteriocins of Gram-positive bacteria where studied which express the scarcely investigated lanthipeptides comprising two-component lantibiotic, the Lichenicidin lantibiotic. Secondly, from the mammalian cathelicidin family the unique cathelicidin produced by humans was chosen, namely LL-37 and two truncated derivatives known as LL-32 and LL-20.

From both studies, it was possible to determine for the first time with atomistic details the unbiased interaction pathway of Lichenicidin with the lipid II structural subunit of the peptidoglycan layer as well as the interaction pathway of Cathelicidins with membrane surfaces. Understanding the interaction pathways with an atomistic detail will, undoubtedly, contribute to the rational design of improved antimicrobial drugs to fight the ever-increasing AMR of a large plethora of pathogens.

Zusammenfassung

Mikroorganismen wie Bakterien, Viren, Pilze und einige Parasiten besitzen die inhärente Fähigkeit, die Aktivität eines antimikrobiellen Mittel -Antibiotika, Antiviren und Antimalaria- gegen sie einzustellen. Diese Eigenschaft ist bekannt als antimikrobielle Resistenz (AMR) und verursacht mit der Zeit die Ineffektivität von Standard-Behandlungen, hilft Infektionen zu bestehen und das verbreiten auf andere Wirte. AMR ist eine der größten Bedrohungen für die globale Gesundheit, die jedes Land und jeden einzelnen Bürger betrifft. Krankenhausaufenthalte werden verlängert, was höhere medizinische Kosten und erhöhte Mortalität verursacht. In diesem Jahr (Februar 2017) veröffentlichte die Weltgesundheitsorganisation (WHO) erstmals eine Liste von Antibiotika-resistenten "prioritären Pathogenen", die ausgewählte Bakterienfamilien enthält, die die höchste Bedrohung für die menschliche Gesundheit darstellen und die Dringlichkeit der erforderlichen neuen Antibiotika zur Bekämpfung von AMR schildert. Angesichts der Tatsache, dass gerade in Europa 25.000 Menschen an einem "Superbug" pro Jahr sterben, ist die Priorität der Gestaltung neuer, effizienter Antibiotika zweifellos eine prioritäre Aufgabe.

So wurde als Alternative zu den vorherrschenden Antibiotika die Aufmerksamkeit auf die antimikrobiellen Peptide (AMPs) gelegt. Zahlreiche Studien haben ihre Wirksamkeit gegen multidrugresistente Bakterien beschrieben. AMPs, die auch als Wirtsverteidigungspeptide (HDPs) bezeichnet wurden, wurden in allen multizellulären lebenden Organismen gefunden. AMPs haben abgesehen von antimikrobiellen Eigenschaften, wichtige Rollen in intrazellulären Prozessen. Leider gibt es unzureichende Kenntnisse über den Wirkmechanismus, um ihre bakteriziden Wirkungen zu entfalten. Mit dem Ziel, Licht auf molekularer Ebene des zugrundeliegenden antimikrobiellen Wirkmechanismus zu erhalten, wurde ein kombinierter theoretischer Ansatz bestehend aus i) Allatom-Molekulardynamik (MD) Simulationen und ii) Entwicklung von Markov State Modellen (MSMs) angewendet. Von den mehr als 2.000 natürlich und synthetischen berichteten AMPs, wurden zwei verschiedene Familien in der aktuellen Studie ausgewählt. Zuerst wurden Bakteriocine von Gram-positiven Bakterien untersucht, die die kaum untersuchten Lanthipeptide exprimieren, die zwei Komponenten-Lantibiotika, das Lichenicidin-Lantibiotikum. Zweitens wurde von der Säugetier-Kathelicidin-Familie das einzige von Menschen produzierte Kathelicidin gewählt, nämlich LL-37 und zwei verkürzte Derivate, die als LL-32 und LL-20 bekannt sind.

Aus beiden Studien war es möglich, zum ersten Mal mit atomistischen Details den unvoreingenommenen Wechselwirkungsweg von Lichenicidin mit der Lipid-II-Strukturuntereinheit der Peptidoglykanschicht sowie dem Wechselwirkungsweg von Cathelicidinen mit Membranoberflächen zu bestimmen. Das Verständnis der Interaktionswege mit einem atomistischen Detail wird zweifellos zur rationalen Gestaltung verbesserter antimikrobieller Medikamente beitragen, um die ständig wachsende AMR einer großen Fülle von Pathogenen zu bekämpfen.

Publications

A. de Miguel, R. Tapia-Rojo, T. Utesch, M. A. Mroginski. Structure, dynamics and kinetics of two-component Lantibiotic Lichenicidin. PLoS ONE 12(6):e0179962. 2016

Collaborations

During my doctoral thesis, I have promoted the following collaborations and research:

Dr. Rafael Tapia-Rojo from the Columbia University, in order to develop the Markov State Models of Lichenicidin Lantibiotic.

Ms. Enrico Forbrig, scientific research of Prof. Dr. P. Hildebrandt at the TU Berlin together with Prof. Dr. T. Gutsmann from Forschungszentrum Borstel-Leibniz-Zentrum für Medizin and Biowissenschaften.

Additionally, I have also supervised the work of the following students: R. Weimar, C. Suazo, R. Atton and C. Jonas, in the study of other antimicrobial peptides.

Contributions in Conferences

Talk: „Structure and dynamics of lantibiotics by molecular dynamics simulations: a fundamental approach towards antimicrobial activity“. **10th European Biophysical Congress EBSA**, Dresden 2015 (Germany).

Poster: A. de Miguel, N. Malanovic, P. Nibbering, J. W. Drijfhout, K. Lohner, M. A. Mroginski. „Molecular Dynamics simulations aid to mode of action of synthetic peptides derived from human cathelicidin LL-37“. **6th International Meeting on Antimicrobial Peptides**, Leipzig 2016 (Germany).

Poster: A. de Miguel, E. Fobrig, T. Gutmann, P. Hildebrandt, M. A. Mroginski. „Understanding the mechanism of action of human cathelicidin LL-37 through a theoretical and experimental approach“. **AMP2016 International Symposium on Antimicrobial Peptides**, Montpellier 2016 (France).

Poster: A. de Miguel, T. Utesch, M. A. Mroginski. „Structure and Dynamics of Lantibiotics by Molecular Dynamics Simulations: a fundamental approach towards antimicrobial activity“. **Gordon Research Conference** Lucca, 2015 (Italy).

Poster: A. de Miguel, T. Utesch, M. A. Mroginski. „Insights into the antimicrobial mechanism of action of two-component Lantibiotic Lichenicidin by means of Molecular Dynamic Simulations“. **Gordon Research Seminar** Lucca, 2015 (Italy).

Table of Contents

Abstract	iii
Zusammenfassung	v
1. Introduction	1
Chapter I	
<hr/>	
2. Molecular Mechanics (MM)	5
2.1 MM FF.....	5
2.1.1 Chemistry HARvard Molecular Mechanics-CHARMM FF	7
2.1.1.1 Intramolecular (internal/ bonded) Terms	7
2.1.1.1.1 Bond Stretching	7
2.1.1.1.2 Angle Bending (θ).....	8
2.1.1.1.3 Bond Torsions (ϕ).....	8
2.1.1.1.4 Improper Torsions (ω)	8
2.1.1.1.5 Urey-Bradley Term	9
2.1.1.1.6 CMAP Term	9
2.1.1.2 Intermolecular (external/ non-bonded) Terms	9
2.1.1.2.1 Van der Waals Interactions.....	9
2.1.1.2.2 Electrostatic Interactions.....	10
2.2 Boundaries.....	12
2.2.1 Periodic Boundary Conditions (PBC)	12
2.2.2 Non-Periodic Boundary Conditions	13
2.3 Dealing with Non-Bonded Interactions.....	13
2.3.1 Cut-offs and Verlet Neighbour List	14
2.3.2 Particle Mesh Ewald Summation.....	15
3. Molecular Dynamics (MD)	17
3.1 Classical MD Simulations	17

3.2	Stochastic Dynamics.....	18
3.3	Finite Difference Methods for the Integration of the Equations of Motion	19
3.4	Statistical Mechanics	21
3.5	Connecting Macroscopic and Microscopic States	22
3.6	Effective Calculation of Thermodynamic Properties	23
3.6.1	Simulations in the Isothermal/ Isobaric (NPT) Ensemble	23
3.7	MD Simulations Protocol.....	24
3.7.1	Stage 1: Thermal-Equilibration of the System.....	24
3.7.2	Stage 2: Dynamics or Production Run	24
3.7.3	Stage 3: Analysis of the Trajectories.....	25
3.7.3.1	Properties Computed for the Analysis of the Trajectories	27
3.7.3.1.1	Root Mean Square Deviation, RMSD	27
3.7.3.1.2	Root Mean Square Fluctuation, RMSF	27
3.7.3.1.3	Radius of Gyration, RG.....	28
3.7.3.1.4	Dipolar Moment, μ	28
3.7.3.1.5	Monitoring of the Secondary Structure	28
3.7.3.1.6	Interaction Energy, IE	29
3.7.3.1.7	Time-dependent Contact Map, TdCM	30
3.7.3.1.8	Radial Distribution Function, RDF	30
4.	Markov State Models (MSMs)	31
4.1	Definition of Markov State Models	31
4.2	Markov State Model Theory.....	32
4.2.1	Continuous MD.....	32
4.2.2	Discretization of the State Space	35
4.3	Protocol for the Building of MSMs.....	36
4.3.1	Definition of the Microstates	36
4.3.2	Transition Matrix.....	37

4.3.3	Proof of Markovianity.....	37
4.3.4	Detailed Balance	38
4.3.5	Ergodicity.....	38
4.3.6	From the Microstates to the Macrostates	39
4.3.6.1	The Stochastic Steepest Descent (SSD) Algorithm	39
4.3.7	Properties Computed through the Analysis of the Microstate Network	40
4.3.7.1	Population	40
4.3.7.2	Free Energy.....	40
4.3.7.3	Entropy.....	40
4.3.7.4	Rate Constant	41
4.3.7.5	Average ScapeTime.....	41
4.4	MSM Protocol Used in this Thesis	41
5.	Vibrational Spectroscopy	43
5.1	Theory of Vibrational Spectroscopy	43
5.2	Surface-Enhanced Infrared Absorption Spectroscopy	44
5.3	Electromagnetic Mechanism (EM) of Enhancement–Plasmon Resonance.....	45
5.4	Chemical Mechanism of Enhancement	46
5.5	Tethered Bilayer Lipid Membrane (tBLM).....	46
5.6	Experimental Section	47
5.6.1	Liposome Preparation	47
5.6.2	Fourier-Transform Infrared (FT-IR) Transmission	48
5.6.2	Surface-Enhanced Infrared Absorption (SEIRA) Spectroscopy	48

Chapter II

6.	Biomembranes	53
6.1	Fluidity.....	54
6.2	Eukaryotic Cells	55
6.3	Prokaryotic Cells	56

6.3.1	Bacteria	56
6.3.1.1	Gram-Negative Envelope	57
6.3.1.2	Gram-Positive Envelope.....	58
6.3.2	Lipid II - Peptidoglycan Cell-Wall Precursor	59
6.3.2.1	Lipid II - Peptidoglycan Cell-Wall Synthesis	60
7.	Antimicrobial Mechanisms of Action	63
7.1	Approaching to Membrane Targets	64
7.2	Execution of Activity.....	64
7.2.1	Self-Association and Multimerization	65
7.2.2	The Carpet Model.....	65
7.2.3	The Toroidal Pore.....	65
7.2.4	The Barrel-Stave Mechanism.....	66
7.3	Cell Death	66

Chapter III

8.	Bacteriocins.....	69
8.1	Bacteriocins of Gram-positive Bacteria.....	70
8.1.1	Class I: Lanthionine - Containing Bacteriocins.....	71
8.1.1.1	Class Ia: The Lanthipeptides - Lantibiotics.....	71
8.1.1.1-A	Subclass I Lanthipeptides: Dehydratase LanB and Cyclase LanC Enzymes...75	
A-	Nisin Lantibiotic	75
A.1	Nisin Lantibiotic: Experimental Observations	75
A.1.1	Physical and Chemical Properties.....	78
A.1.2	Mode of Action	79
A.2	Nisin Lantibiotic: Molecular Dynamic Simulations of Nisin Natural Variants Produced by Lactococcus lactis	81
A.2.1	Validation of the CHARMM-Compatible FF for Dehydroamino acids.....	81
A.2.2	Conformational Dynamics of Nisin Z, A and Q.....	84

A.2.3 Complex Nisin Z-Lipid II Immersed in a Membrane Mimicking a Gram-Positive Bacteria.....	88
A.2.4 Concluding Remarks of the All-Atom MD Simulation of Nisin Z, Natural Variants A and Q, and Evaluation of the Cage Architecture of the Complex Nisin Z-lipid II	92
8.1.1.1-B Subclass II Lanthipeptides: Bifunctional Lanthionine Synthetase.....	93
B- Two-component Lantibiotic Lichenicidin	99
B.1 Conformational Dynamics of two-component Lantibiotic Lichenicidin in Solution	101
B.1.1 Model Building.....	101
B.1.2 All-Atom MD simulation 2 μ s-long for Lichenicidin Lantibiotic	102
B.1.3 MSMs for Lichenicidin Lantibiotic	105
B.1.3.1 MSMs for Lch α peptide	106
B.1.3.2 MSMs for Bli α peptide	108
B.1.3.3 MSMs for Lch β peptide	111
B.1.4 Concluding Remarks of Conformational Dynamics of Lichenicidin Lantibiotic	115
B.2 Mechanism of Action of Two-Component Lantibiotic Lichenicidin by Means of All-Atom MD Simulation	117
B.2.1 Model Building and Set-up of the All-Atom MD Simulation of Two-Component Lantibiotic Lichenicidin in a Biological Relevant Environment	118
B.2.2 Analysis of the 1.3 μ s-long of All-Atom MD Simulation	121
B.2.2.1 Interaction with Lipid II, Cell-Wall Precursor.....	121
B.2.2.2 Interaction Energies	136
B.2.2.3 Dynamics of the Undecaprenyl Pyrophosphate	187
B.2.2.4 Aggregate-Stabilization	138
B.2.2.4-A Interactions Inter-AG-I Peptides	138
B.2.2.4-B Interactions Inter-AG-II and AG-III Peptides	149
B.2.2.4-C Membrane Effects for the Aggregate-Stabilization	175
B.2.2.5 Role of Lys25 and Glu26.....	182

B.2.2.6 Concluding Remarks	189
----------------------------------	-----

Chapter IV

9. Mammalian Cathelicidins	195
9.1 Diversity, Structure and Classification	195
9.2 Human Cathelicidin LL-37	198
9.2.1 Antimicrobial Activity of Cathelicidin LL-37	1999
9.2.1.1 Mechanisms of Antimicrobial Activity of Cathelicidin LL-37	200
9.2.2 Pleiotropic Effects of LL-37	200
9.3 Cathelicidin LL-37 and its Fragments by Means of MD Simulations.....	201
9.3.1 Conformational Dynamics of LL-37 and its Fragments LL-32 and LL-20 in Solution.....	203
9.3.1.1 Model Building and Set-up of the All-Atom MD Simulation of LL-37 and its Fragments LL-32 and LL-20 in Solution	203
9.3.1.2 Analysis of the Trajectories for LL-37 and its Fragments	204
9.3.2 Theoretical Studies of LL-37 and its Fragments in Biologically Relevant Environments.....	209
9.3.2.1 Model Building and Set-up of the All-Atom MD Simulation of LL-37 and its Fragments LL-32 and LL-20 with Nearby Lipid Membrane	209
9.3.2.2 Analysis of the Trajectories for LL-37 and its Fragments LL-32 and LL-20 with Nearby Lipid Membranes.....	22100
9.3.2.2-A Monomers in a Lipid Environment	2210
9.3.2.2-B Dimerization Studies for LL-37 and Fragment LL-32.....	22166
9.3.3 Agreement with the Experimental Measurements.....	220
9.3.3.1 Transmission IR Experiments	221
9.3.3.2 Surface-Enhanced IR Spectroscopy (SEIRA) Experiments Using Tethered Bilayer Lipid Membranes (tBLMs).....	222
9.3.4 Concluding Remarks Combining Theoretical and Experimental Measurements	226

Appendix.....	231
10. Appendix – Bacteriocins.....	233
10.1 Force Field Parameters for Dehydroalanine and Dehydrobutyrine.....	233
10.1.1 Dha Topology.....	233
10.1.2 Dhb Topology.....	233
10.1.3 Remarks on Atom-Type Assignment.....	234
10.1.4 Force Field Parameters.....	234
10.2 Force Field Parameters for Lanthionine and Methyllanthionine.....	236
10.2.1 Topology for Methyllanthionine.....	236
10.2.2 Topology for Lanthionine.....	237
10.2.3 Force Field Parameters.....	237
10.3 Lipid II.....	238
10.4 Markov State Models.....	238
10.4.1 Time Series of MSM Macrostates for the 2 μ s MD Simulation.....	242
10.4.2 Validation: Chapman-Kolmogorov Test.....	243
10.4.3 TICA Spectrum for the Three Peptides.....	243
10.4.4 Transition Matrix Lch α	244
10.4.5 Transition Matrix Bli α	245
10.4.6 Transition Matrix Lch β	245
11. Appendix – Cathelicidin LL-37 and Fragments.....	246
11.1 All-Atom MD Peptide S9V-LL-20.....	246
11.1.1 Comparative Analysis with Fragment LL-20.....	246
11.2 Overview of the Infrared Absorption Bands (in cm ⁻¹) of the Amino Acid Residues of LL-37 and its Fragments.....	247
12. List of Abbreviations.....	249
13. References.....	252
Danksagung.....	252

1. Introduction

Antimicrobial Resistance (AMR) is the resistance towards drugs used for the treatment of diverse pathogens such as bacteria, fungi, viruses and parasites. Currently, the drug resistance is a worldwide problem caused by the life style and globalization of this century which implies the global trade and travel, promoting the accelerated spreading of microorganisms around the world. If a microorganism is exposed to an antimicrobial, due to natural selection, those able to resist will pass their resistance to their offspring. This fact combined with a *bad-praxis* of consuming antimicrobials promotes, consequently, the development of drug resistance or AMR. Furthermore, the failure of proper handling and quality assurance systems for most of the manufactured drugs, further contributes the conditions for AMR. Depending on the country, patients do not have the same possibilities to access antimicrobials, interrupting the full required treatment against the pathogen and promoting an optimal scenario for the microbes to acquire resistance. Additionally, the use of antibiotics in animal-rearing with the aim of endorsing growth implies a hazardous situation, causing the emergence of resistant microorganism which potentially could infect humans.

With the slogan '*Combat drug resistance- No action today, no cure tomorrow*' the World Health Organization (WHO)^[1] initiated in 2015 a global plan for the successful treatment and prevention of infectious diseases according to quality-guaranteed medicines in a liable-term to all who requires them. The global endorsement has begun in February of 2017 for the first time in the WHO-history the publication of a list involving a selection of 12 families of bacteria that present the highest threats to the human health denoted as antibiotic-resistant 'priority pathogens' due to the elevated number of people who die from a superbug per year. Since the pipeline for new tools to fight drug resistance is practically squeezed, an eye on antimicrobial peptides (AMPs) have been put, to solve this urgent worldwide issue.

Several experimental studies have corroborated the effectiveness of AMPs against multidrug-resistant bacteria, to be consequently proposed as alternatives to the current antibiotics. Different multicellular living organisms are able to generate these host defense peptides (HDPs) not only highlighted because of their antimicrobial properties, but also due to their significant importance in diverse intracellular processes. Scientific research plays a

crucial role in order to amplify the remaining gaps regarding, for instance, the mechanisms used by AMPs to exert their bactericidal effects.

This thesis addresses concisely that issue at an atomistic level. Two selected classes of AMPs were chosen to be studied by means of all-atom molecular dynamics (MD) simulations combined with the development of Markov State Models (MSMs). Several efforts have been made from an experimental point of view in order to understand how the antimicrobial mechanism of action of such complex systems, consisting of two peptide lantibiotics, take place and no conclusive results have been reported so far. Which is why, the study of the two-component lantibiotic lichenicidin endorse an exciting challenge from a theoretical approach. Additionally, the unique human cathelicidin (LL-37) discovered up-to-date and two shorter derived fragments, namely LL-32 and LL-20, were under study using the same technique.

The first chapter of the thesis compromises an overall overview of the methods used section 2-Molecular Mechanics (MM), section 3-Molecular Dynamics (MD), section 4-Markov Models (MSMs) and section 5-SEIRA Spectroscopy consisting in a brief explanation of the experimental technique employed by the multidisciplinary collaboration partners in order to successfully combine theoretical and experimental results.

The second chapter is related to the section 6-Biomembranes and 7-Mechanims of Antimicrobial Peptides with the aim of understanding the biological requisites exposed in chapters 3 and 4.

The third chapter is focused on the bacteriocins produced by Gram-positive bacteria (section 8-Bacteriocins). Extense analysis of the two peptides addressed is here provided. Firstly, biological properties and theoretical results for Nisin lantibiotic, which was used as control peptide and secondly, a summary of all two-component lantibiotics discovered up-to-date combined with biological aspects follow, to conclude with the theoretical results regarding the antimicrobial mechanims of action of the two-component lantibiotic Lichenicidin.

The fourth chapter (section 9) deals with the study of the mammalian cathelicidins. The theoretical results obtained in this research were herein compared with experimental measurement

Chapter I

2- Molecular Mechanics (MM)

3- Molecular Dynamics (MD)

4- Markov Models (MSMs)

5- Vibrational Spectroscopy

2. Molecular Mechanics (MM)

Molecular mechanics (MM), or commonly called by the user, force field (FF) methods, compute the total energy of the system taking only into account the nuclear positions ignoring the electronic motions. Therefore, MM methods cannot provide insight into properties which depend on the electronic distribution of a molecule. MM works properly due to the validity of different assumptions, starting from the Born-Oppenheimer approximation, which enables the calculation of the energy as a function of the nuclear coordinates. MM is based on a simple model of interactions within a system considering all different movements that describe the natural behaviour of it, from the stretching of bonds, to the opening and closing of angles and twisting of torsional angles.

2.1 MM FF

A FF is an assembly of equations and associated constants, mainly designed to reproduce the structural properties of a system describing its energy as a function of its atomic coordinates, hence, they are additionally useful to compute spectral properties such as the vibrational spectra. Other important characteristic of the FFs is the transferability of the functional form and the parameters.

Molecular modelling FFs can be interpreted by means of a combinatorial description of two terms: the intramolecular (internal/ bonded) and intermolecular (external/ non-bonded) forces within the system. Energy penalties are related with the deviation of the bonds and angles from their 'equilibrium' values. Additionally, there are two more functions associated to the torsions and to the interaction between non-bonded parts of the system. Several biomolecular protein FFs have been developed during the last years such as CHARMM [2], Amber [3], GROMOS [4], OPLS-AA [5]. Despite that some differences can be found between them, an invariableness of the four terms (bonds, angles, torsions and non-bonded parts) is found. The total energy can be then described as:

$$E_{total} = E_{bonded} + E_{non-bonded} + E_{other} \quad (2.1.1)$$

$$E_{bonded} = E_{bonds} + E_{angles} + E_{dihedrals} \quad (2.1.2)$$

$$E_{non-bonded} = E_{vdW} + E_{elect} \quad (2.1.3)$$

$$E_{other} = \text{any force field specific term} \quad (2.1.4)$$

The bonded energy is constituted by the contribution of the bond stretching, angle bending together with the dihedral and improper torsions (figure 2.1). In contrast, the non-bonded term is related to the ‘through-space’ interactions. The non-bonded terms involve the electrostatics and the Van der Waals interactions.

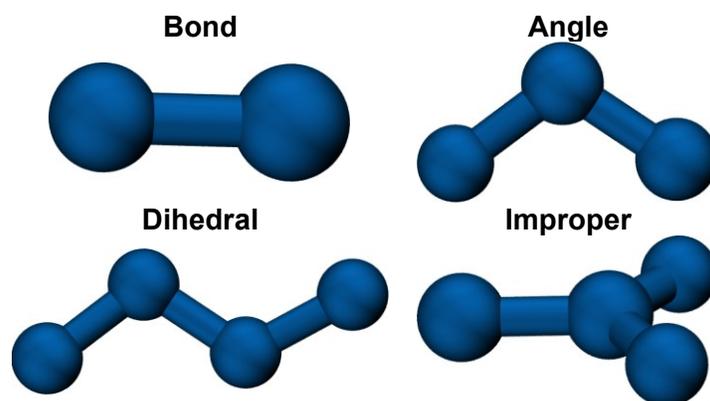


Figure 2.1: Representation of the bonded terms. Bond stretching, angle bending, dihedral and improper torsions.

2.1.1 Chemistry HARvard Molecular Mechanics-CHARMM FF

CHARMM is one of the most commonly used FFs. The CHARMM FF^[2] is represented in the equation (2.1.1.1) and the individual contributions will be explained in detail in this section.

$$\begin{aligned}
 U(\vec{R}) = & \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 \\
 & + \sum_{Urey-Bradley} K_{UB} (S - S_0)^2 \\
 & + \sum_{dihedrals} K_\phi (1 + \cos(n\phi - \delta)) \\
 & + \sum_{impropers} K_\omega (\omega - \omega_0)^2 \\
 & + \sum_{non-bonded} \left\{ \varepsilon_{ij}^{min} \left[\left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{12} \right. \right. \\
 & \left. \left. - 2 \left(\frac{R_{ij}^{min}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi \varepsilon_0 \varepsilon r_{ij}} \right\} \\
 & + \sum_{residues} U_{CMAP}(\phi, \psi)
 \end{aligned} \tag{2.1.1.1}$$

2.1.1.1 Intramolecular (internal/ bonded) Terms

In CHARMM FF, the ‘bonded’ term involve basically the sum of bond stretching, angle bending, bonded and improper torsions together with the Urey-Bradley term and the correction map CMAP.

2.1.1.1.1 Bond Stretching

Commonly, the Hook’s law is used to define the bond stretching for the interaction between pairs of bonded atoms as the variation of the square of the displacement from the reference bond length b_0 . The energy is illustrated by a harmonic potential and the shape of the harmonic potential well is defined by the force constant K_b .

$$U_{bond} = \sum_{bonds} K_b (b - b_0)^2 \quad (2.1.1.1.1)$$

2.1.1.1.2 Angle Bending (θ)

The deviation of the angles from their reference values is also described using an harmonic potential. The function involves the force constant, K_θ and the difference between the angle θ to its equilibrium value θ_0 .

$$U_{angles} = \sum_{angles} K_\theta (\theta - \theta_0)^2 \quad (2.1.1.1.2)$$

2.1.1.1.3 Bond Torsions (ϕ)

The dihedrals evaluate the contribution of each bonded quartet of atoms in the system. The periodic potential illustrating the bond torsions is defined throughout the force constant K_ϕ , the torsion ϕ , the phase shift δ together with the multiplicity n .

$$U_{dihedrals} = \sum_{dihedrals} K_\phi (1 + \cos(n\phi - \delta)) \quad (2.1.1.1.3)$$

2.1.1.1.4 Improper Torsions (ω)

Improper torsions (ω) are commonly denoted as out-of-plane bending, thus are defined by harmonic potentials in analogy to the bond stretching and angle bending functions. The torsional potential is measured through the force constant K_ω and the variation between ω and its equilibrium value ω_0 . Typically, the improper torsional terms are used in the united atom force fields to assess the stereochemistry at the chiral centres.

$$U_{impropers} = \sum_{impropers} K_\omega (\omega - \omega_0)^2 \quad (2.1.1.1.4)$$

2.1.1.1.5 Urey-Bradley Term

It is only used in some special cases to optimize the fit to vibrational spectra and out-of-plane motions. The Urey-Bradley component accounts the cross-term for angle bending using 1,3 nonbonded interactions, where K_{UB} is the respective force constant and S is the distance between the 1,3 atoms in the harmonic potential.

$$U_{Urey-Bradley} = \sum_{Urey-Bradley} K_{UB} (S - S_0)^2 \quad (2.1.1.1.5)$$

2.1.1.1.6 CMAP Term

It is a numerical correction for the protein main chain in order to overcome the inaccuracies of the conformational energetics in the backbone of the peptides. These defects were originally reflected by a high proportion of π -helical conformations for helical peptides.

$$U_{CMAP} = \sum_{residues} U_{CMAP}(\phi, \psi) \quad (2.1.1.1.6)$$

2.1.1.2 Intermolecular (external/ non-bonded) Terms

The non-bonded interactions involve pair of atoms which are separated by three or more bonds in the same molecule or between separated molecules.

2.1.1.2.1 Van der Waals Interactions

The Van der Waals interactions are described by the 12-6 Lennard-Jones (L-J) potential.

The energy term between two atoms i and j involves an attractive $\left(\frac{Rmin_{ij}}{S_{ij}}\right)^6$ part which derives from the London potential, and a repulsive part $\left(\frac{Rmin_{ij}}{S_{ij}}\right)^{12}$. The $Rmin_{ij}$ represents the interatomic equilibrium distance for which no attraction or repulsion can be detected between the atoms i and j . ϵ describes the depth of the potential well and S_{ij} refers to the distance between both atoms i and j (figure 2.1.1.2.1).

$$U_{vdW} = \sum_{vdW} \varepsilon_{ij} \left[\left(\frac{Rmin_{ij}}{S_{ij}} \right)^{12} - 2 \left(\frac{Rmin_{ij}}{S_{ij}} \right)^6 \right] \quad (2.1.1.2.1)$$

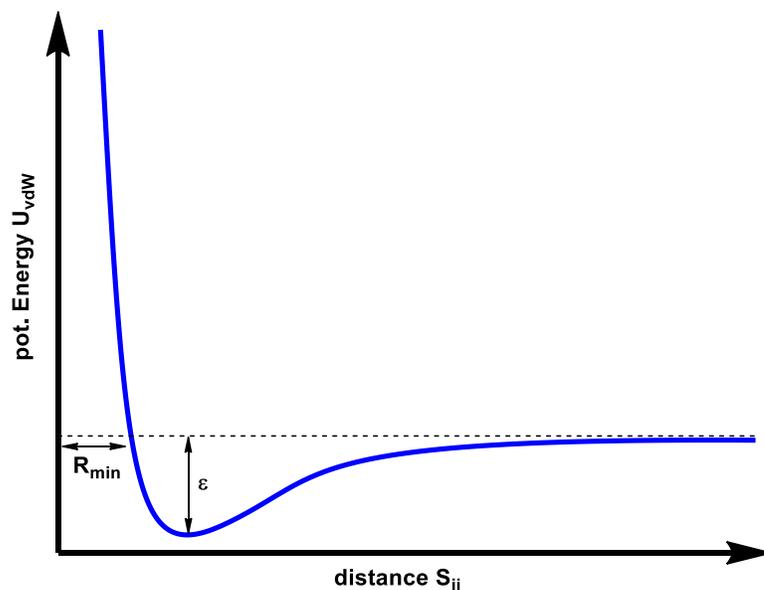


Figure 2.1.1.2.1: Lennard-Jones potential. Lennard-Jones potential describing the vdW energy between two particles. $Rmin_{ij}$ and ε denote for the equilibrium distance, where the force between the two atoms i and j is zero, and the depth of the energy well, respectively.

2.1.1.2.2 Electrostatic Interactions

The electrostatic interactions between two atoms i and j are computed through the Coulomb's law. The electrostatic potential depends on the partial charges q_i and q_j , the distance between both atoms r_{ij} , and the dielectric constant *in vacuo* ε_0 . The electrostatic interaction between two charged particles decay slowly with the interatomic distance (figure 2.1.1.2.2).

$$U_{elec} = \sum_{elec} \left\{ \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon r_{ij}} \right\} \quad (2.1.1.2.2)$$

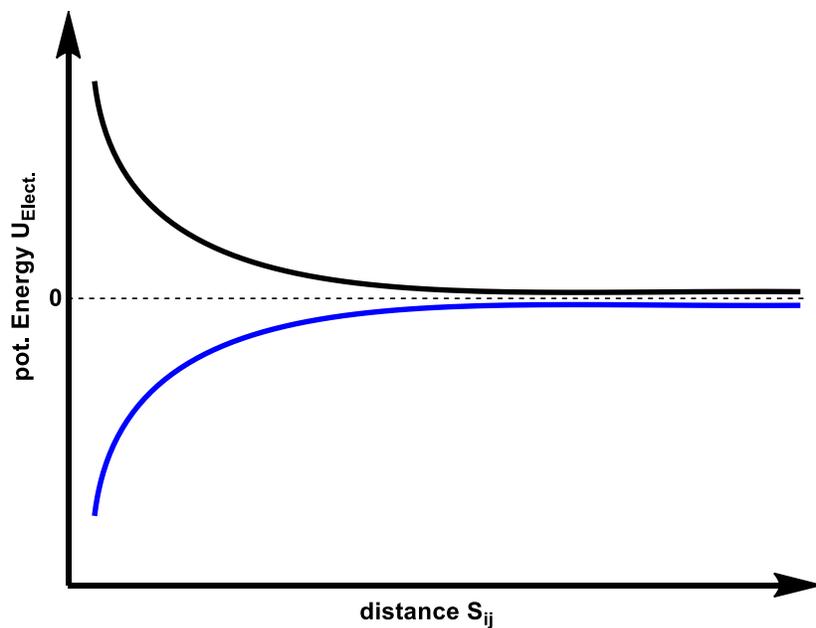


Figure 2.1.1.2.2: Coulombic potential. Coulombic potential between two identical (black traces) and contrary (blue traces) charged particles. Equally charged particles result in positive potentials.

2.2 Boundaries

In order to compute 'macroscopic' properties from MD simulations employing a relatively small number of particles, the correct treatment of the boundaries of the cell volume's simulation box is crucial. Usually, the MD simulations are time-demanding which is also an important reason to keep the system as small as possible in order to find a compromise between realistic data and computation time. The treatment of interacting particles with the artificial borders has to be carefully handled, mainly due to the effects in the potential energy and the dynamics of the system. Thus, in order to deal with the box limits two main approaches have been proposed: periodic boundary conditions and non-periodic methods.

2.2.1 Periodic Boundary Conditions (PBC)

PBC is a widely used technique in classical MD simulations in order to provide realistic environments^[6]. The main goal is to minimize the particle interactions with the box limits in such a way that the PBC mimic mirror cells in all possible directions obtaining a periodic and infinite system keeping the number of particles constant (figure 2.2.1). Consequently, through the use of PBC, the treatment of the long-range electrostatics is significantly improved than by cut-offs forced by the system limits (see section 2.3).

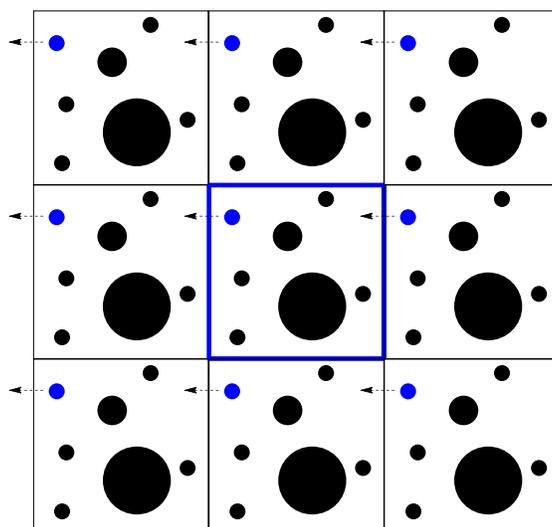


Figure 2.2.1: Periodic Boundary Conditions. The unit cell is placed in the center (blue square) which is surrounded by image cells in all directions. The arrow points out the motion of the blue sphere into the proximal image cell. Representation made in 2D for clarity.

2.2.2 Non-Periodic Boundary Conditions

In early studies of macromolecules, involving proteins and proteins-ligand complexes, all solvent molecules were ignored due to the limited availability of the computational resources. The results of these unrealistic simulations of the proteins *in vacuo* differ from those obtained at the experimental level. The use of the vacuum boundaries minimizes the surface area causing more compact conformations attributable to favourable intramolecular electrostatic and Van der Waals interactions.

Other alternative is the employment of stochastic boundary conditions (SBC)^[6] which is derived from the generalized Langevin equation by Kantorovich and Rompotis. According to this approach, only a small subset of solvent molecules are considered explicitly and the effect of the remaining water are incorporated with the use of a boundary potential. This method consists in the subdivision of the case study in three regions: 1) the fixed atoms, 2) the Newtonian atoms (moving according Newtonian dynamics), and 3) the buffer region of Langevin atoms (moving according Langevin equations)^[7-9]. The Newtonian region compromises the reaction region, while the atoms placed in the Langevin region play the role of dissipating the heat. Furthermore, the fixed atoms are needed for the creation of the correct potential well for the Langevin atoms to move in. The method is consistent with the follow equation

$$m_i \frac{dv_i}{dt} = -m_i \zeta v_i + R_i - \frac{\partial U}{\partial r_i} \quad (2.2.2)$$

Herein, $-m_i \zeta v_i$ denoted the friction while ζ is called friction coefficient. R_i is the random gaussian distributed force while $\frac{\partial U}{\partial r_i}$ denotes the term for balancing temperature of the system exchanged between friction and the random force.

2.3 Dealing with Non-Bonded Interactions

Computation of the non-bonded interactions are the most time-consuming step in MD simulations. Usually, the most common approach is to employ a non-bonded cut-off combined with the minimum image convention. Under periodic boundary conditions, the minimum image convention considers that each atom can see a maximum of one image of every other atom in the system. According to this, the energy/ forces are computed with the

closest atom/ image and when the cut-off is employed the interactions between the pairs of atoms out of the range are set to zero. The cut-off should be selected in such a way that the atom does not see its own image, this means that it should not be longer than the half of the length of the cell.

The use of a defined cut-off allows to deal properly with the non-bonded interactions: i) for Van der Waals interactions this approach is easily adjustable, because as it can be seen in the figure 2.1.1.2.1 the interaction force between the particles tends to zero with the interatomic distance while ii) in opposite, for the long-range interactions the use of the particle mesh Ewald summation is extended.

2.3.1 Cut-offs and Verlet Neighbour List

The determination of the atoms which are within the cut-off range of a specific atom is computationally demanding. In order to efficiently select the surrounded atoms the Verlet neighbour list^[10,11] is commonly used. This method consists in the definition of two spheres around the atom i , defined by the radii r_{cut} and r_{res} . The sphere of r_{cut} include all atoms within the cut-off range of atom i while the sphere defined by r_{res} is associated to the reservoir area and involve all atoms placed around atom i and out of the r_{cut} (figure 2.3.1).

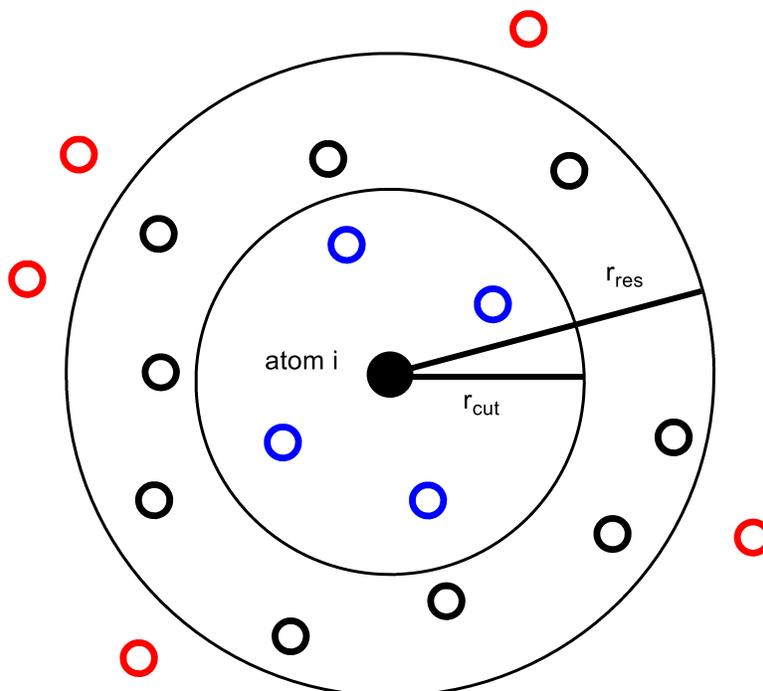


Figure 2.3.1: Verlet neighbour list algorithm. Three groups are distinguished from a central atom i : within the cut-off radius r_{cut} (blue spheres), within the neighbour list r_{res} (black particles) and those out of the r_{res} list (red spheres).

Generally, the radii are defined in such a way that r_{res} is in a range around 2 - 4 Å longer than the r_{cut} for the Van der Waals and short-ranged electrostatics. Another important fact, which has to be accomplished during the MD simulations, is the correct update of the reservoir region according to:

$$r_{res} - r_{cut} > N_{up} \tilde{v} \Delta t \quad (2.3.1)$$

Herein, the N_{up} refers to the update interval for the reservoir list r_{res} , \tilde{v} is the average atom velocity and Δt is the time step. It is crucial to update interval N_{up} of the neighbour list at every step Δt to achieve computational efficiency.

2.3.2 Particle Mesh Ewald Summation

Particle mesh Ewald (PME)^[12] algorithm takes the full electrostatics into account. The algorithm decreases the complexity of the electrostatic force evaluation from $O(N^2)$ to $O(N \log N)$ consequently, decreasing the computational time required for the evaluation of

2. Molecular Mechanics (MM)

the long-range forces. PME is a particular case of the Poisson summation formula, this lattice sum method consists in the replacement of the summation of the interaction energies in real space with an equivalent summation in the Fourier space. In the Ewald method, the initial set of charges, calculated in the real space, are surrounded by a Gaussian distribution to which a neutralising charge distribution of the equal magnitude and of opposite sign, calculated in the reciprocal space, must be added. In doing so, a fast convergence of both terms is achieved.

3. Molecular Dynamics (MD)

Biomolecular studies from an experimental approach are, in several circumstances, impossible, too dangerous, expensive and in some cases, due to time scales, specific properties are unattainable. Because of that, MD simulations can bridge the gaps or compliment the experimental studies. Consequently, an experiment can be replaced, explained, provoked and aided by the use of MD simulations.^[13]

Thus, MD simulations are an efficient approach for the understanding of the molecular mechanism of action of AMPs.^[14] According to the atomistic details achieved through the analysis of the MD trajectories, its contribution to a rational design of new molecular antibiotics must be highlighted.

3.1 Classical MD Simulations

MD simulations is a sophisticated method based on the classical mechanics, concisely on the Newton's second equation of motion with a classical MM force field. Accordingly, the acceleration of a specific system can be computed if the force acting on each atom is known. Once that the acceleration in the system is computed, the equations of motion can be integrated resulting in a trajectory describing the evolution in time of: positions, velocities and accelerations for every single particle in the system.

If the potential energy of the system U is known, such as CHARMM^[2] described in the previous chapter (eq. 2.1.1.1), the force F_i on each atom of mass m_i at position r_i can be computed. Consequently, the new positions of the particles can be computed as a function of the time according to:

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} = m_i \frac{d^2 r_i}{dt^2} \quad (3.1.1)$$

where F_i is the product of the mass m and acceleration a of the particle i , which is defined as the first and second derivative of the velocity v_i and coordinates r_i over time t , respectively. Moreover, F_i is denoted as the negative gradient of the potential energy U calculated by the force field according to

$$F_i = -\nabla_i V \quad (3.1.2)$$

Combination of equations (3.1.1) and (3.1.2) provide the description of the potential energy as a function of the coordinate changes over the time:

$$-\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \quad (3.1.3)$$

where the trajectory of the particle i is directly defined from its initial coordinates r_i , initial velocities v_i and initial acceleration a_i . In case of biomolecular systems, the crystallographic structure obtained from the Protein Data Bank (PDB)^[15] is used for defining of the initial coordinates while the acceleration can be directly computed as the derivation of the potential energy function. In contrast, the initial distribution of velocities has to be estimated, employing usually the Maxwell-Boltzmann distribution at a defined temperature T . Herein, k_B is related to the Boltzmann constant.

$$p(v_i) = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left(-\frac{1}{2} \frac{m_i v_i^2}{k_B T}\right) \quad (3.1.4)$$

As it was mentioned in the previous chapter (2. Molecular Mechanics (MM)), several empirical force fields exist. The work described in this thesis was developed employing CHARMM force field, mainly because it is continuously developed and improved to fit accurately to the experimental studies. Furthermore, there are a plenty of programs available to execute the simulations. Herein, a parallel MD code designed for high-performance simulations of large biomolecular systems, Nanoscale Molecular Dynamics (NAMD^[16]) was used to run all simulations.

3.2 Stochastic Dynamics

The commonly known Langevin stochastic dynamics^[8,9] are an extension of the classical dynamics consisting in the addition of two terms to the classical equations of motion: the friction ζ_i and a stochastic random term R_i :

$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} + \frac{R_i}{m_i} - \zeta_i \frac{dr_i}{dt} \quad (3.2)$$

Implementation of this contribution does not affect the conservation of the total energy of the system due to the opposing effect provided by the two last terms: the

stochastic random term R_i supplies energy E_{kin} (collisions and interactions with solvent molecules) while the friction term ζ_i removes E_{kin} representing the frictional drag on the system moving through a solvent. The energy conservation is, therefore, enabled by the coupling of the stochastic random term R_i to an external bath.

3.3 Finite Difference Methods for the Integration of the Equations of Motion

The *finite difference methods*^[6] are used in order to carry out MD simulations with continuous potentials. The main idea behind splits the integration into several small broken stages separated in time by a fixed time dt .

The total force at the specific time t is computed as the vector sum of the interactions of each particle with the surrounding particles. Directly from the force, the acceleration is determined which combined with the velocities together with the positions at time t allows the computation of the new positions and velocities at time $t + dt$. As the force is constant during the time step, it allows the evaluation of the forces for the particles in the new positions, resulting in a concatenated determination of the new positions and velocities at the time $t + 2dt$ and subsequent.

All different algorithms employed for the integration of the equations of motion in MD simulations, such as Verlet algorithm^[11], Leap-frog algorithm^[6], velocity Verlet algorithm^[11] or Beeman's algorithm^[17] establish that all dynamic properties can be approximated as Taylor series expansions.

The calculations made in this work were performed with NAMD program^[16]. In this program, the available method is velocity Verlet integration, for the acquisition of the new positions and velocities. Accordingly, the velocity Verlet algorithm is implemented as a three-stage routine providing at the same time the new positions, velocities and accelerations.

$$r(t + \delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 \quad (3.3.1)$$

$$v(t + \delta t) = v(t) + \frac{1}{2}[a(t) + a(t + \delta t)]\delta t \quad (3.3.2)$$

3. Molecular Dynamics (MD)

Firstly, the new positions are computed through the (3.3.1) employing the velocities v and accelerations a at time t . Subsequently, the velocities at time $t + \frac{1}{2}dt$ are computed according to:

$$v\left(t + \frac{1}{2} \delta t\right) = v(t) + \frac{1}{2} a(t)\delta t \quad (3.3.3)$$

followed by the determination of the new forces from the new positions obtaining $a(t + dt)$. Concluding the procedure with the calculation of the velocities at time $(t + dt)$ according to:

$$v(t + \delta t) = v\left(t + \frac{1}{2} \delta t\right) + \frac{1}{2} a(t + \delta t)\delta t \quad (3.3.4)$$

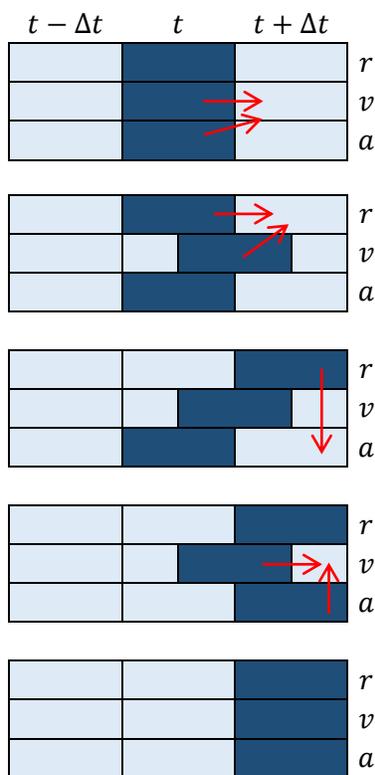


Figure 3.3: Velocity Verlet algorithm. Three steps procedure to solve the equations of motion in NAMD.

3.4 Statistical Mechanics

Through the use of MD simulations, microscopic information associated to the atomic positions and velocities is derived. The extrapolation of these details to macroscopic observables, including pressure, heat capacities or energies, which constitute a thermodynamic state, demands statistical mechanics. Statistical mechanics^[18] provide the required mathematical expressions to establish a bridge between macroscopic properties and microscopic simulations of a system. While the thermodynamic state is defined by a combination of variables such as the number of particles N , the temperature T , the pressure P derived from thermodynamic equations. The mechanical or microscopic state is related to the atomic positions q and the momenta p which are directly considered as coordinates in a multidimensional space, namely phase space. The phase space, for a system of N particles, is characterized by $6N$ dimensions and any single point, denoted as G , can describe the state of the system. An ensemble is the collection of points in the phase space satisfying the conditions of a specific thermodynamic state.

MD simulations create a succession of points in the phase space as a function of the time, corresponding to the same ensemble which additionally are associated to different conformations of the system and their respective momenta.

Different ensembles have been defined:

- *Microcanonical* ensemble NVE where the thermodynamic state is characterized by a constant number of atoms N , volumen V and energy E . This state represents isolated systems.
- *Canonical* ensemble NVT where the thermodynamic state is defined through a constant number of particles N , volume V and temperature T .
- *Isobaric-Isothermal* ensemble NPT which is characterized by fixed number of atoms N , pressure P , and temperature T .
- *Grand canonical* ensemble μPT where the chemical potential, μ together with the volume, V and the temperature, T are fixed.

3.5 Connecting Macroscopic and Microscopic States

An experiment at the macroscopic level is usually associated to a large number of particles covering several number of conformations. The interconnecting link established by statistical mechanics with the microscopic level involve the macroscopic properties to be understood in terms of ensemble averages. Consequently, the ensemble averages can be then described as an average of several replicas of the system simultaneously studied. Thus, the average values are described as ensemble averages in statistical mechanics.

Because of that, the observable of interest is described as a function of the momenta p and the positions r

$$A(p^N, r^N) \quad (3.5.1)$$

while the ensemble average is described by

$$\langle A \rangle_{ensemble} = \iint dp^N dr^N A(p^N, r^N) \rho(p^N, r^N) \quad (3.5.2)$$

MD dynamics simulations compute a time average of A according to the following expression:

$$\begin{aligned} \langle A \rangle_{time} &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(p^N(t), r^N(t)) \delta t \\ &\approx \frac{1}{M} \sum_{t=1}^M A(p^N, r^N) \end{aligned} \quad (3.5.3)$$

where t is the simulation time, M is the number of time steps and $A(p^N, r^N)$ is instantaneous value for A .

The relation between ensemble averages (Statistical Mechanics) with time averages (MD) is given by the ergodic hypothesis, which assumes an equivalence between the ensemble average and the time average.

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} \quad (3.5.4)$$

In order to satisfy the previous assumption, the simulation have to acquire enough sampling time with the aim of covering all possible representative conformations to allow,

afterwards, the calculation of the experimental observable associated to structural variations, dynamics and related thermodynamic properties.

3.6 Effective Calculation of Thermodynamic Properties

With the aim of reproducing experimental data, more commonly, the MD simulations are run under *NPT* ensemble^[19,20], where the temperature and the pressure are kept constant. Consequently, all simulations carried out in this work were run under a *NPT* ensemble, which is implemented into the NAMD program^[16].

3.6.1 Simulations in the Isothermal/ Isobaric (*NPT*) Ensemble

In NAMD, for the *NPT* ensemble simulations, a new set of equations of motion inspired by the Langevin-piston method^[21] and the Hoover's method for constant pressure simulations^[22] are used.

The pressure is controlled by adjusting the size of the unit cell together with the rescaling of the atomic coordinates over the course of the simulation. In this modified Nosé-Hoover method, the fluctuations in the barostat together with the temperature control are controlled by Langevin dynamics (eq. 2.2.2), in order to properly simulate a *NPT* ensemble. The combination of both provided the equation:

$$\frac{dp}{dt} = \sum_{i=1} \sum_{j=i+1} r_{ij} F_{ij} - \frac{\zeta}{Q} p(t) \quad (3.6.1.1)$$

where the heat-bath mass is approached to Nk_bT and the friction coefficient is given by the following expression:

$$\frac{\Delta\zeta}{dt} = \sum_i \frac{1}{2} m_i v_i^2 - \frac{3}{2} Nk_B T \quad (3.6.1.2)$$

3.7 MD Simulations Protocol

Generally, the standard protocol employed for the performance of molecular dynamic simulations can be divided in three well differentiated stages: preparation of the system, dynamics and analysis of the trajectories (figure 3.7).

3.7.1 Stage 1: Thermal-Equilibration of the System

Step 1

Generally, the starting configuration of the system of interest is directly downloaded from the protein structure data bank^[15]. Subsequently, if there is a lack of information, such as missing atoms or gap regions, the structure must be manually completed. Additionally, it is crucial to control the protonation states of the amino acids depending on the pH of the media. Afterwards, the system must be solvated and ionized according to the experimental conditions of interest.

Step 2

Once the system has been prepared, with the aim of avoiding any possible steric instabilities an energy minimization of the system is required. Through the minimization local energy minima is found. Different algorithms can be used for this purpose, for the simulations carried out in this work, an extension of the steepest descend algorithm^[23], the conjugate gradient method^[24] was used.

After the energy minimization, the equilibration phase allows the system to reach a thermal equilibrium. The length of this stage depends on the relaxation time required by the studied system to achieve stability.

3.7.2 Stage 2: Dynamics or Production Run

Once the first stage is fulfilled, it is possible to perform the production run in order to determined and analyse the dynamic properties of the system. The production run is extended until the desired simulation time t which is reached by steps of Δt .

3.7.3 Stage 3: Analysis of the Trajectories

The output of a classical MD simulation is a trajectory describing the time dependent evolution of all atoms of the system in a defined period of time t . Additionally, according to the statistical mechanics, macroscopic observables, such as energy, temperature, pressure or volume can be also determined. The post-processing of the data enables the evaluation of interaction energies, distances and angles and further properties of interest which can be computed employing different plug-ins or in-house developed programs. In this thesis, properties such as: root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (RG), dipolar moment (μ), monitoring of the secondary structure, the interaction energy (IE) and time-dependent contact map (TdCM), radial distribution function (RDF), have been used for an in-depth analysis of the trajectories of the AMPs. All of them are going to be explained in the next section 3.7.3.1.

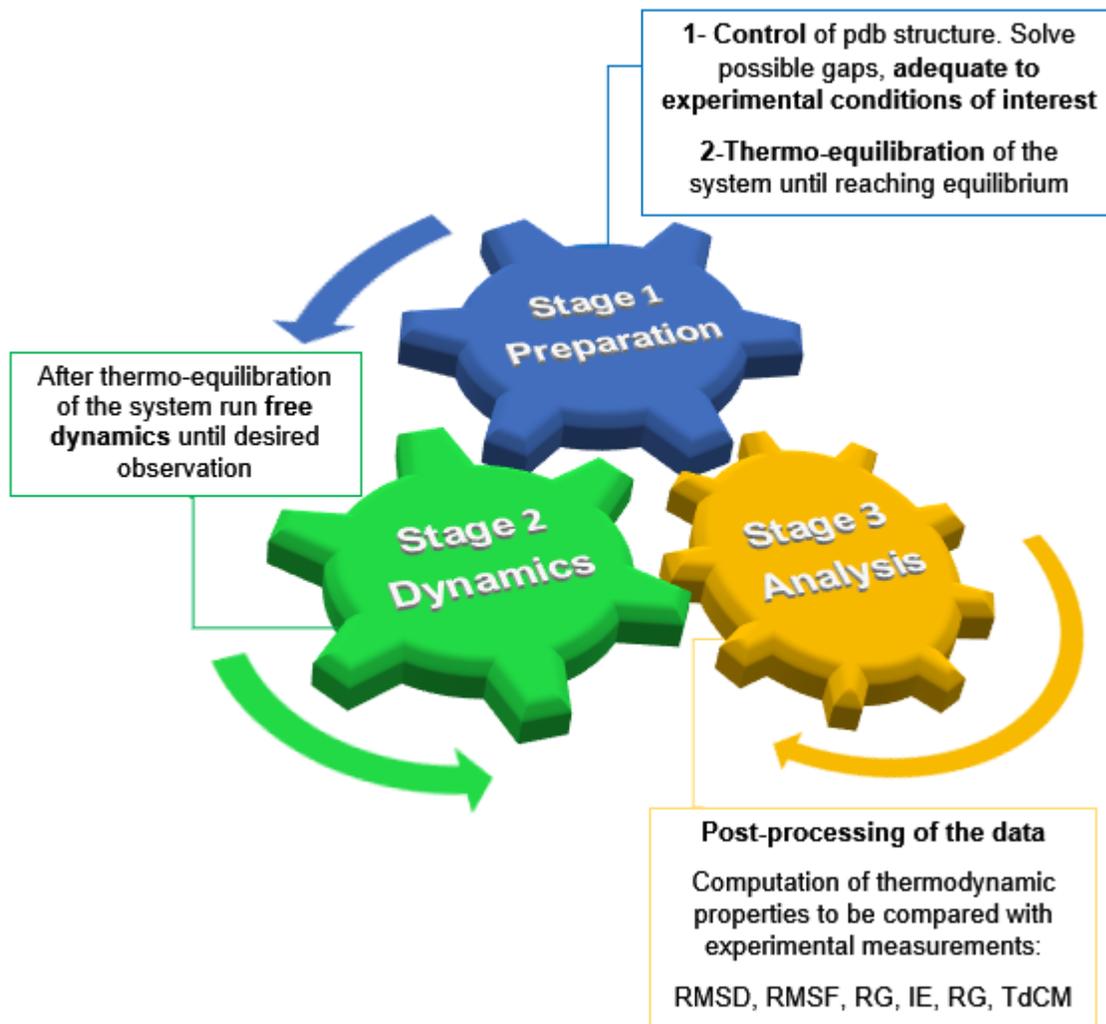


Figure 3.7: MD simulations protocol. Three-step procedure to solve the equations of motion in NAMD consisting of preparation, dynamics and anylisis of the data.

3.7.3.1 Properties Computed for the Analysis of the Trajectories

3.7.3.1.1 Root Mean Square Deviation, RMSD

Essentially, the RMSD provide information regarding the stability of the studied system through comparison of the native structure (reference structure) with all conformations acquired at different simulation times. Beforehand, the conformations obtained during the simulation must be aligned to the reference one in order to measure correctly the deviation.

The RMSD is computed according to:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_i (r_{k,i} - r_{l,i})^2} \quad (3.7.3.1.1)$$

which is defined as the difference between two frames, namely k and l , which is associated to the square root of the sum of the square displacements of all atoms i , and it is normalized by total number of atoms N . High values of RMSD involve notorious deviations from the initial structure. Usually, in order to analyze variations in the secondary structure, it results useful to compute exclusively the RMSD of the backbone of the peptide of interest. It is usually measured in Å.

3.7.3.1.2 Root Mean Square Fluctuation, RMSF

In contrast with the measure of the RMSD, the analysis of the RMSF provides information regarding the variation of a single atom of the system over the time, t_T . Usually, the chosen atom is the C_α and through the screening of its behavior in the course of the simulation, flexible regions of the system can be accurately detected. Commonly, the starting configuration is used as reference, r_{ref} . Another alternative is the use of an average position as reference providing improved results by discarding protein translations and rotational movements. The RMSF is computed as follows, and it is usually measured in Å.

$$\text{RMSF} = \frac{1}{t_T} \sqrt{\sum_i (r(i) - r_{ref})^2} \quad (3.7.3.1.2)$$

3.7.3.1.3 Radius of Gyration, RG

The radius of gyration is the measure of the compactness/ shape of the studied-system. It is expressed according to:

$$RG = \sqrt{\frac{1}{N} \sum_i (r(i) - r_{ref})^2} \quad (3.7.3.1.3)$$

where the initial structure is used as reference, N refers to the total number of particles and r is associated to the configurations obtained during the simulation time. It is usually measured in Å.

3.7.3.1.4 Dipolar Moment, μ

The dipolar moment provides information regarding the separation between the positive and negative charges in the system. It is measured using as reference distance the center of mass (COM) or center of geometry (COG) of the protein taking into consideration the mass m_i per atom i , the total mass of the protein M together with the coordinates, r_i per particle i .

$$r_{COM} = \frac{1}{M} \sum_{atoms} m_i r_i \quad (3.7.3.1.4)$$

Once the r_{COM} is defined, together with the partial charges q_i defined from the force field, and the distance per atom, r_i to the reference point, it is feasible to compute the dipolar moment:

$$\mu = \sum_i q_i (r_i - r_{COM}) \quad (3.7.3.1.5)$$

3.7.3.1.5 Monitoring of the Secondary Structure

In this work, for monitoring of the secondary structure, STRIDE program was used. STRIDE^[25] is based on a knowledge-algorithm taking into account the combination of the hydrogen bond energy and statistically derived backbone torsional angle information, which is afterwards optimized and the resulting assignments are in optimal agreement with

crystallographer's labels. The program is implemented in Visual Molecular Dynamics^[26] (VMD) program as Timeline plug-in.

Figure 3.7.3.1.5 shows the typical output, where the following color code is used for the designation of the possible variety of secondary structures: teal (turns), yellow (extended configurations), green (isolated bridges), pink (α -helix), blue (3_{10} helix), red (π -helix), white (coil).

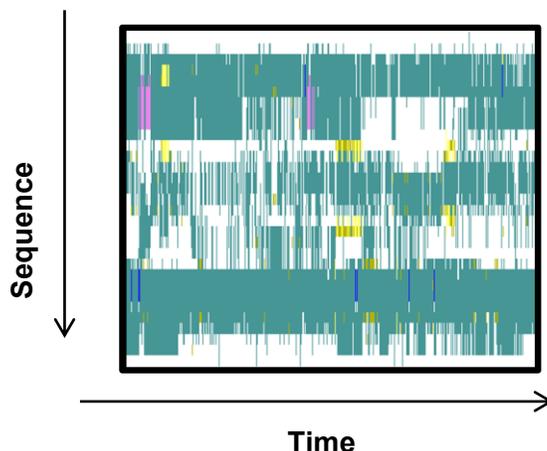


Figure 3.7.3.1.5: Output of the secondary structure evolution. Color code denote the variety of secondary structures found: teal (turns), yellow (extended configurations), green (isolated bridges), pink (α -helix), blue (3_{10} helix), red (π -helix), white (coil).

3.7.3.1.6 Interaction Energy, IE

The evaluation of the interaction energies, E_{int} is useful to determine the attracting or repelling character between two reaction partners A and B . The calculation takes only into consideration the non-bonded terms, since the formation or breaking of bonds are not usually described by classical MD simulations. The E_{int} between A and B is determined from overall self-energy between both partners, E_{AB} and removal of the individual contributions or self-energies of the respective associated reactants, namely E_A and E_B (eq. 3.7.3.1.6). This calculation was carried out employing the VMD^[26] plug-in NAMDEnergy.

$$E_{\text{int}} = E_{AB} - E_A - E_B \quad (3.7.3.1.6)$$

3.7.3.1.7 Time-dependent Contact Map, TdCM

This *in-house* made program monitors the contacts between two reaction partners placed at a distance lower than 3 Å in the simulation time t . The program is most useful when combined with the calculation of the energies of interaction described in the preceding section. The results obtained provide the graphical monitoring of the number of interaction per residue $\{i, i + 1, i + 2 \dots\}$ of a specific peptide m_i to the reaction partner m_j composed by the residues $\{j, j + 1, j + 2 \dots\}$ during the simulation.

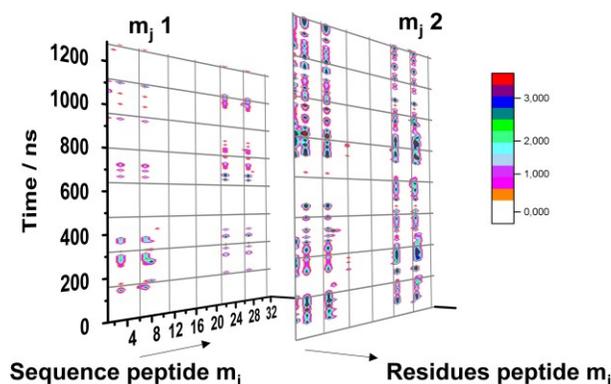


Figure 3.7.3.1.7: TdCM of peptide i and two residues of peptide j . Overall role control per residue for identification of the interaction type.

3.7.3.1.8 Radial Distribution Function, RDF

It is also commonly known as pair correlation function, $\rho(r)$. It describes how the density is modified as a function of the distance employing a specific particle as reference over the trajectory, or a subset of it, of the thermodynamic ensemble of interest. In biomolecular systems, it is of extreme interest because it allows a direct comparison with the experimental measured values. In MD simulations, the $\rho(r)$ is defined as:

$$\rho(r) = \frac{1}{N_{frame}} \sum_i^{N_{frame}} \sum_{j \in sel_1} \sum_{k \in sel_2; k \neq j} \delta(r - r_{ijk}) \quad (3.7.3.1.8)$$

where N_{frame} is related to the number of frames, r_{ijk} is the distance between the j and atom k for frame i , while δ is the Dirac delta function.

4. Markov State Models (MSMs)

4.1 Definition of Markov State Models

Markov States Models (MSMs)^[27] have become an useful method to understand and improve the analysis of MD simulations. MSMs represent the free energy landscape as a kinetic network that determine the molecule's conformations and dynamics at a given time interval. The states are related to the free energy basins through which a transition probability matrix can be computed, providing information concerning the probability of jumping from one state to another, and therefore, about the free energy barriers. It is mandatory that the network is Markovian in the defined time interval (namely lag time τ), which implies that the probability of the system to evolve to a new state depends exclusively on the current state and not on the former one. Additionally, a precise balance must be held with the purpose of maintaining the microscopic reversibility.

The key factor for building the MSMs is the definition of the states themselves. The states must be differentiated in such a way that they correspond to the free energy basins where the boundaries between them are correctly defined. Because of that, the exclusive use of a geometric guideline to define them is insufficient and the combination of kinetic and geometric criteria must be considered. Usually, the MSMs are executed in a two-step protocol: firstly, under a geometric and secondly under a kinetic criteria. In the first step, an amount of elements are defined according to a chosen geometric evidence, which will represent the microstate of the system. This first network is characterized by a large number of elements thwarting its study. Subsequently, according to this first selection, two different scenarios can be obtained: fast kinetic transitions associated to fast relaxing processes and slow kinetic relaxations.

Consistent with the observables obtained during the first step, the lumping of the microstate network into kinetic relevant clusters, associated to the free energy basin, is carried out. This second step defines the macrostates of the system where the boundaries between them are related to the free energy barriers.

4.2 Markov State Model Theory

4.2.1 Continuous MD

The state space Ω holds all the dynamical properties needed to define the current state of the system. In MD simulations, the dynamical variables contained in the Ω are understood in terms of positions and velocities. The state of the system at time t is defined as $x(t) \in \Omega$ and characterized by the following properties:

- $x(t)$ is a Markov process in Ω . This means that the change of the system depends exclusively on the current $x(t)$. In all-atom MD simulations, the classical equations of the motion are solved and the trajectories are markovian by definition.
- $x(t)$ is ergodic. The dynamical process is aperiodic and Ω does not have disconnected subsets that cannot be reached with one trajectory. In other words, in $t \rightarrow \infty$, each point can be infinitely visited. This assumption involves that any average of a function $f: \Omega \rightarrow \mathbb{R}^d$ is given by a single stationary distribution $\pi(x)$ which means that for every initial state x is accomplished:

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T dt f(x(t)) = \int_{\Omega} dx f(x) \pi(x) \quad (4.2.1.1)$$

and the fraction of time in which the system remains in any of its states is defined by the stationary density $\pi(x) : \Omega \rightarrow \mathbb{R}_{0+}$ being $\int_{\Omega} dx \pi(x) = 1$. This stationary density involves that if $P(x, y; \tau)$ is the transition probability between two points, namely x and $y \in \Omega$, within the time interval τ , then

$$\pi(y) = \int_{\Omega} dx P(x, y; \tau) \pi(x) \quad (4.2.1.2)$$

Confirming that the stationary density is unique and in most of the studies is directly associated to the thermodynamic ensemble.

- $x(t)$ is reversible. In equilibrium, the fraction of systems going from x to y per unit of time is identical to the fraction of systems from y to x . Assuming that the transition probability satisfies the condition:

$$\pi(x)P(x, y; \tau) = \pi(y)P(y, x; \tau) \quad (4.2.1.3)$$

Focusing now on the evolution of an ensemble density, it can be said that the probability density $p_t(x)$ differs from $\pi(x)$ at time t distributed in Ω . This density will change in an interval τ with the action of the transition probability density $P(x, y; \tau)$. The change itself can be then described by the action of a continuous operator, the propagator $\mathcal{P}(\tau)$ defined as:

$$p_{t+\tau}(y) = P(\tau) \circ p_t(y) = \int_{\Omega} dx P(y, x; \tau) p_t(x) \quad (4.2.1.4)$$

In resemblance with $\pi(x)$, the modified probability density is achieved through the application of $P(\tau)$ to $p_t(x)$. Taking also into account that in an infinite time, the probability density becomes $\pi(x)$. Alternatively, the transfer operator $\mathcal{T}(\tau)$ can be used. The main difference is located in that $\mathcal{T}(\tau)$ propagates functions $u_t(x)$, differing from the probability densities by a factor of the stationary density equal to: $p_t(x) = \pi(x)u_t(x)$.

Both operators show the following characteristics:

- Fulfil the Chapman-Kolmogorov equation allowing to propagate long times

$$p_{t+n\tau}(x) = [P(\tau)]^n \circ p_t(x) \quad (4.2.1.5)$$

$$u_{t+n\tau}(x) = [\mathcal{T}(\tau)]^n \circ u_t(x) \quad (4.2.1.6)$$

- $P(\tau)$ has eigenfunctions $\phi_i(x)$ and eigenvalues λ_i

$$P(\tau) \circ \phi_i(x) = \lambda_i \phi_i(x) \quad (4.2.1.7)$$

in contrast with $\mathcal{T}(\tau)$, which has eigenfunctions ψ_i with the same eigenvalues λ_i

$$\mathcal{T}(\tau) \circ \psi_i(x) = \lambda_i \psi_i(x) \quad (4.2.1.8)$$

4. Markov State Models (MSMs)

All eigenvalues λ_i are real and are placed in the range of $-1 \leq \lambda_i \leq 1$ if the dynamics are reversible. Additionally, both eigenfunctions are connected by the factor of the stationary density $\pi(x)$ according to:

$$\phi_i(x) = \pi(x) \psi_i(x) \quad (4.2.1.9)$$

- The eigenfunction with the largest eigenvalue $\lambda_1 = 1$ is related to the stationary distribution $\pi(x)$, so it involves:

$$P(\tau) \circ \pi(x) = \pi(x) = \phi_1(x) \quad (4.2.1.10)$$

Therefore, the eigenfunction $\psi_i(x)$ is constant on the continuous space Ω like:

$$\phi_1(x) = \pi(x) \psi_i(x) = \pi(x) \quad (4.2.1.11)$$

The eigenspectrum of the operators P and \mathcal{T} allow to decompose the dynamics into m slow and fast molecular processes. \mathcal{T}_{slow} contains all relevant dominant slow processes while \mathcal{T}_{fast} include the fast processes of scarce interest.

$$\begin{aligned} u_{t+n\tau}(x) &= \mathcal{T}_{slow}(\tau) \circ u_t(x) + \mathcal{T}_{fast}(\tau) \circ u_t(x) \\ &= \sum_{i=1}^m \lambda_i^n \langle u_t, \psi_i \rangle_{\pi} \psi_i(x) + \mathcal{T}_{fast}(\tau) \circ u_t(x) \end{aligned} \quad (4.2.1.12)$$

The slow dynamics are a superposition of the dynamical processes with an associated eigenfunction ψ_i or ϕ_i and an eigenvalue λ_i . These processes decay with the time, in long-time limit, only the first term with $\lambda_1 = 1$ remains, recovering to the stationary distribution $\phi_1(x) = \pi(x)$. The other eigenfunctions are associated to processes which decay with time, being dynamical rearrangement features, while the ensemble relaxes to the equilibrium distributions. The physical measurable timescale for each process is described:

$$t_i = -\frac{\tau}{\log \lambda_i} \quad (4.2.1.13)$$

Which allow the rewriting of the eq. (4.2.1.12) as follows where all fast processes have been neglected.

$$u_{t+n\tau}(x) \approx 1 + \sum_{i=2}^m e^{-\frac{n\tau}{\tau_i}} \langle u_t, \psi_i \rangle_{\pi} \psi_i(x) \quad (4.2.1.14)$$

4.2.2 Discretization of the State Space

MD simulations are Markovian by construction in a full continuous state space Ω . In practice, MSMs precise discretization of the state space in order to acquire a manageable description of the dynamics. MSMs are based on the partition of the state space into discrete states which combined with the transition matrix is used to model the transitions observed in the dynamics.

The development of the Markov model is obtained by discretizing the state space followed by the estimation of the corresponding transfer operator from the simulation data and not through the discretization of the propagator. A practical example is described in the following:

Let's consider a discretization of the state space Ω into N sets S_i . The process can be merely a partition with sharp boundaries of the considered degrees of freedom or, another option, a reduced amount after the application of a dimension reduction technique. The stationary probability π_i found in set i is given by the full density

$$\pi_i = \int_{x \in S_i} dx \pi(x) \quad (4.2.2.1)$$

herein, S_i is the i^{th} partition of the state space Ω involving that $S = \{S_1, \dots, S_N\} : \cup_{i=1}^N S_i = \Omega$ and the markov model is then defined by the transition probability matrix $\tilde{T}(\tau) \in \mathbb{R}^{N \times N}$, which is the discrete approximation of the transfer operator \mathcal{T} . Each element T_{ij} is the time-stationary probability to find the system in set j at time $t + \tau$ given that at time t was in set i .

According with the explanations made in the previous section 4.2.1, if the population of a sets S_i at time t is the column vector $p(t) \in \mathbb{R}^N$, it is possible to compute then the probability after time τ

$$p_j(t + \tau) = \sum_{i=1}^N p_i(t) T_{ij}(\tau) \quad (4.2.2.2)$$

Consequently, the stationary probability of the discrete states π_i will be

$$\pi^T = \pi^T \tilde{T}(\tau) \quad (4.2.2.3)$$

It is possible to develop an eigenvalue decomposition of the matrix $\tilde{T}(\tau)$ in order to find the N associated dynamical processes and describe the system with exclusively the m slow ones under a specific threshold benchmark.

4.3 Protocol for the Building of MSMs

In the next section, the steps will be briefly explained, which have to be fulfilled in order to generate a MSM for this case study.

4.3.1 Definition of the Microstates

To effectively discretize the conformational space, a geometric criterium is used for the classifications of the conformations observed during the MD simulation. Accordingly, the microstate network is developed. This geometric partition should be kinetically relevant which involves that between conformations, where the system can rapidly jump, should be grouped together. This geometric partition, or commonly known distance metric, allows the discretization of the conformational space into individual bins. It is useful to seek for a meaningful order parameter, because a collective variable as a direct partitioning of the coordinate space $\{r\}$ would yield into massive partitions. An example is the use of the RMSD, which is a commonly employed technique, such as in PCA^[28] or TICA^[29] for reducing the initial number of coordinates and perform a fine partitioning directly on them. The partitioning of the conformational space has to be done directly (i.e into bins of equal volume) or by means of a clustering algorithm such as k -centers clustering or k -medoids clustering^[27].

At this stage, the trajectory is transferred into a sequence of discrete bins. The microstate network is described by the transition count matrix C_{ij} , which counts the jumps between the bins i and j while the occupation vector π_i provides the weight of the node i . All these findings allow to understand the dynamical trajectory $\{r(t)\}_{i=1}^N$ as a weighted and

directly connected network, where each node represent the conformational microstates, linked by possible transitions between the defined microstates.

4.3.2 Transition Matrix

The microstate network is defined by the transition probability matrix \tilde{T} , which is determined as:

$$T_{ij}(\tau) = \frac{C_{ij}}{\sum_k C_{ik}} \quad (4.3.2.1)$$

4.3.3 Proof of Markovianity

MSMs have to be Markovian in the studied lag time interval. Despite that the MD simulations are directly Markovian (any configuration can be obtained from the previous one), when discretizing the Ω , a coarse-graining is actually performed, so it could be only markovian at longer timescales. In order to validate the markovianity of the model the Chapman-Kolmogorov test has to be done which involves:

$$\tilde{T}(n\tau) = \tilde{T}(\tau)^n \quad (4.3.3.1)$$

where n is the number of steps of length τ . The given equivalence satisfies that in a model, taking n steps with lag time τ , it must be equivalent as taking exclusively a single step with lag time equal to $n\tau$.

The most common procedure to do this validation is through the analysis of the relaxation timescales of the system, which is directly associated to the eigenvalues of the transfer matrix (section 4.2.1)

$$t_i = -\frac{\tau}{\log \lambda_i} \quad (4.3.3.2)$$

Herein, t_i is the relaxation time and λ_i the i^{th} eigenvalue. In order to proof markovianity, it has to be fulfilled that the relaxation times of the Markov Model with a lag time of $n\tau$ should be identical to those with a lag time τ

$$t_i = -\frac{n\tau}{\log \lambda_{i,T(n\tau)}} = -\frac{n\tau}{\log \lambda_{i,T(\tau)}^n} = -\frac{-n\tau}{n \log \lambda_{i,T(\tau)}} = -\frac{\tau}{\log \lambda_{i,T(\tau)}} \quad (4.3.3.3)$$

Where $\lambda_{i,T(\tau)}$ is an eigenvalue of $T(\tau)$. In the figure 4.3.3 is shown the evolution of the relaxation timescale as a function of the lag time, their stable tendency is directly related to the satisfaction of the markovianity.

4.3.4 Detailed Balance

Detailed balance is also understood as microscopic reversibility, which involves that from every time there is a transition from state i to state j , which has to be compensated with a transition from state j to state i . Different reasons exist by which a detailed balance is not achieved:

- Limited sampling
- Non-attainment of a true equilibrium sampling
- Poor definition of the microstates

If the states are visited only once along during the trajectory, they become source or sink states, that should be trimmed off the model. The detailed balance can be enforced by symmetrizing the count matrix.

$$\hat{C}_{ij}(\tau) = \frac{C_{ij} + C_{ji}}{2} \quad (4.3.4.1)$$

herein, \hat{C}_{ij} is the estimation of the reversible counts and C_{ij} the 'real' transitions.

4.3.5 Ergodicity

The network has to be interconnected. This means that any state can be reached from any other arbitrary state. Lack of ergodicity involves an insufficient sampling during the simulation, being necessary to increase the simulation time in order to get more data or directly discarding the disconnected components by choosing exclusively the largest ones.

4.3.6 From the Microstates to the Macrostates

Lump the microstates to gain physical insights of the system, it is necessary to take into account:

- what microstates should be merged
- number of macrostates to be built

Several methods have been developed to address these two issues, the most popular one is the Perron Cluster Cluster Analysis (PCCA) which employs the eigenspectrum of the transition probability matrix to build the coarse-grained model. This method is based on the identification of the slow timescales which are understood as the dominant eigenvalues in the spectrum of the transition probability matrix. PCCA requires a clear gap between the fast and the slow scales in the eigenvalues. Other methods based on the same principles are commonly used as PCCA+, SHC, BACE and SSD. For the generation of the lichenicidin MSMs the latter one was used.

4.3.6.1 The Stochastic Steepest Descent (SSD) Algorithm

This algorithm was designed^[30] for detecting basins of attraction over a discretized free energy surface, in the case study, the microstate network. The algorithm clusters the nodes depending on the probability flux relaxation, and according to that, to the kinetic criteria associated to the system. Employing as starting point a random state i , the initial probability ($\pi_i(0) = \delta_{a,i}$ for $i = 1, \dots, N$) is allowed to evolve in time, promoting the relaxation of the Markov chain as $\pi(\tau) = \tilde{T}\pi(0)$. If the system starts placed in the node a , it is relaxed through the maximum probability flux, moving to some node b , where the probability is again concentrated ($P_i(1) = \delta_{b,i}$). The process is repeated until a node is reached defined by the probability flux. This last node constitutes a *minimum* in the free energy landscape. Through the reiteration of the process over the whole network, a set of pathways is achieved, which move the system to different minima. Consequently, the nodes leading to the same free energy minimum are defined to be in the same basin of attraction and to be clustered in the same macrostate. At this point, the network has been divided onto a set of macrostates or basins of attraction.

4.3.7 Properties Computed Through the Analysis of the Microstate Network

According to the coarse-graining metastable macrostates found by the SSD algorithm or any other, different thermodynamic and kinetic properties can be computed such as:

4.3.7.1 Population

Herein, the population of the macrostate α , π_α can be determined by

$$\pi_\alpha = \sum_{i \in \alpha} \pi_i \quad (4.3.7.1.1)$$

where π_i is the population of the microstate i . The transition probability from the microstate α to β .

$$T_{\alpha\beta} = \frac{\sum_{j \in \beta} \sum_{i \in \alpha} T_{ij} \pi_j}{\sum_{j \in \alpha} \pi_j} \quad (4.3.7.1.2)$$

The previous equation assess a correct normalization and accomplishment of the balance condition. This variable can be also understood as the ‘weight’ of each state which is associated to the depth of the free energy well.

4.3.7.2 Free Energy

This property is usually calculated by comparison with the most occupied state. The free energy from the state α to β is computed according to:

$$\Delta G_\alpha = -kT \log \frac{\pi_\alpha}{\pi_\beta} \quad (4.3.7.2.1)$$

4.3.7.3 Entropy

The entropy according to Shannon’s information theory which is based on probability theory and statistics:

$$S_\alpha = -k_B \sum_{i \in \alpha} \pi_i \log \pi_i \quad (4.3.7.3.1)$$

4.3.7.4 Rate Constant

The rate constant for the transition from basin α to β

$$k_{\alpha\beta} = \frac{T_{\alpha\beta}}{\tau} \quad (4.3.7.4.1)$$

4.3.7.5 Average Scape Time

The average scape time of any basin (α) to another (β), which is directly associated to the energy barriers between the states

$$t_{\alpha} = \frac{\tau}{(1 - T_{\alpha\beta})} \quad (4.3.7.5.1)$$

4.4 MSM Protocol Used in this Thesis

In this section, the protocol used for the development of the MSMS of Lichenicidin will be explained.

- Being Ω the state space of the system, comprising the biomolecule and surrounding bath particles, $x(t)$ denotes the state of the system at time t . First of all, it is required to choose the subset of the Ω interesting for the case study. For all-atom MD simulation, the subset of the Ω of interest is the whole biomolecule, protein/ peptide, or most commonly, parts of it like the backbone or the C_{α} -coordinates. For lichenicidin lantibiotic the latter option was used.
- Secondly, it is useful to reduce the number of coordinates through the application of a specific method, such as PCA^[28] or TICA^[29]. In this thesis, TICA was employed. The resulting small set of coordinates become the configuration space over which the Markov state model will be built.
- Subsequently, the microstate network through the discretization of the configurational space into discrete bins is generated.
- Once the microstate network is defined, the existence of a detailed balance together with the markovianity at the selected lag time must be controlled.

4. Markov State Models (MSMs)

- The application of the SSD algorithm to the microstate network is then applied. With the aim of defining the basins of attraction of the system, the free energy minima, which promote the definition of the macrostate network.

Through a simple glance of the network several information can be obtained. From the direct identification of the most relevant states, which combined with its occupation reflect its stability. Additionally, the evaluation of the transition matrix is used to determine the rates between the states. Furthermore, the calculation of different properties (see section 4.3.7) were computed.

5. Vibrational Spectroscopy

The discovery of the infrared (IR) radiation in the spectrum of the sun by William Herschel in 1800 set the foundation for the invention of vibrational spectroscopic techniques. In the following century researchers transferred this knowledge to investigations of matter and developed the method of IR spectroscopy. After focusing on inorganic molecules in the beginning, innovations allowed its application to biomolecule characterization studies in life science^[31].

Currently, there is special interest in the highly complex mechanisms of cellular membranes and their interaction with signaling molecules, proteins and peptides. For that purpose, the immobilization of a model membrane system on metal surfaces is a suitable approach to assure the necessary stability of the system and allow applying electrochemical techniques. Particularly, the increasing importance of surface-sensitive techniques, e.i. surface-enhanced infrared absorption (SEIRA) spectroscopy, offers new insights to peptide-membrane interactions. In the presented work, a tethered bilayer lipid membrane (tBLM) system was constructed to study the peptide-membrane interaction of the antimicrobial peptide LL-37 as well as its fragments LL-32 and LL-20 by means of SEIRA spectroscopy.

5.1 Theory of Vibrational Spectroscopy

The interaction of electromagnetic radiation with matter can induce the transition between vibrational states of molecules that is utilized by vibrational spectroscopy, i.e. infrared (IR) spectroscopy. These transitions can be attributed to a resonant absorption of IR radiation ($10\text{-}12500\text{ cm}^{-1}$)^[31]. If the energy of a photon $h\nu_{\text{IR}}$ of polychromatic IR radiation matches the energy difference between an initial and final vibrational state ($h\nu_i$ and $h\nu_f$, respectively) the transitions occur:

$$h\nu_{\text{IR}} = h\nu_i - h\nu_f \quad (5.1.1)$$

Molecular vibrations are based on the model that delineates a harmonic oscillator assuming that atoms are point masses connected by mass-less springs^[32]. A molecule

composed of N atoms possesses $3N-6$ vibrational degrees of freedom ($3N$ reduced by translation and rotation, each in respect to the three axis in space), denoted as normal modes. Since linear molecules can rotate in respect to two axes, it contains $3N-5$ normal modes. As the simplest case, the vibration of a diatomic molecule A-B with the masses m_A and m_B as well as the reduced mass μ is described. Based on Hooke's law for a simple oscillator, the force constant f , the reduced mass μ and the velocity of light c is related to the vibrational frequency ω and can be expressed in wavenumbers $\tilde{\nu}$ (in cm^{-1}):

$$\tilde{\nu} = \frac{1}{2\pi c} \sqrt{\frac{f}{\mu}} \quad (5.1.2)$$

The infrared absorption intensity is described using the absorbance A (in OD=optical density), i.e. the negative decadic logarithm of the ratio between the light intensity after passing through the sample I and the initial light intensity I_0 . According to the Lambert-Beer law this value is proportional to the concentration c , the path length of the sample d and the molar absorption coefficient ε , the latter being related to the transition dipole moment of the considered vibration:

$$A = -\lg\left(\frac{I}{I_0}\right) = \varepsilon \cdot c \cdot d \quad (5.1.3)$$

5.2 Surface-Enhanced Infrared Absorption Spectroscopy

SEIRA effect was first described by Hartstein *et al.* using Ag and Au-films to adsorb monolayers of organic compounds on surfaces^[33]. This effect is achieved by electromagnetic and chemical enhancement increasing the IR-signal by a factor of 10-100 within a distance of ca. 8 nm. This distance dependence excludes signal impairments by the bulk solution^[34,35]. In this work, the SEIRA technique was employed in the attenuated total reflexion (ATR) mode^[36].

5.3 Electromagnetic Mechanism (EM) of Enhancement–Plasmon Resonance

The electric field of the incident IR radiation polarizes metal islands by the excitation of collective electron resonances (localized plasmon modes). The dipole p , which is induced in the metal, generates a local EM field surrounding the particles (see figure 5.3). This induced local EM field is polarized perpendicularly to the surface at every point of the metal particle (surface selection rule) and its magnitude depends on the sixth power of the distance to the metal surface d

$$|E_{\text{local}}|^2 = \frac{4p^2}{d^6} \quad (5.3.1)$$

so that the enhancement factor F at a distance d to a nanoparticle with a radius a_0 can be calculated by

$$F(d) = F(0) \cdot \left(\frac{a_0}{a_0 + d}\right)^6 \quad (5.3.2)$$

This enhanced EM field interacts with adsorbed molecules and can excite their vibrational states.

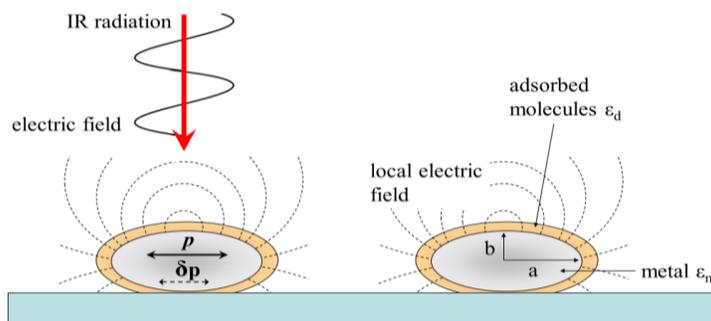


Figure 5.3: Electromagnetic mechanism contributing to surface enhanced infrared absorption (SEIRA). The polarization of ellipsoidal formed metal particles by incident IR radiation induces a dipole p parallel to the surface. The resulting local electric field enhances vibrations of adsorbed molecules on the particles, which lead to a dipole change δp and interference with the optical properties of the metal. Adapted with kind permission from J. Kozuch^[37].

5.4 Chemical Mechanism of Enhancement

Additionally to physisorption, chemisorption enhances the signal of adsorbed molecules, since the donor-acceptor interaction between the adsorbate and the metal surface changes the vibrational polarizability of the molecules. Furthermore, a charge transfer between the molecular orbitals and the metal might increase the absorption coefficient. Chemisorption and physisorption often lead to an alignment of the molecular dipoles with a preferential orientation in respect to the metal surface. Vibrations perpendicularly oriented to the surface lead to IR absorption that is three times higher than randomly oriented ones, described by the surface selection rule^[35].

5.5 Tethered Bilayer Lipid Membrane (tBLM)

In this approach, a lipid-tethered lipid bilayer membrane was constructed (figure 5.5). A thiol group connects the hydrophilic spacer 6-mercaptohexanol (6MH) and a linker molecule containing a hydrophobic cholestanyl headgroup (WK3SH) to a nanostructured metal surface^[38,39]. The phase separation of both SAM molecules allows alternating island formation of lipid monolayers and bilayers after spreading of POPC:POPG liposomes on top of the SAM surface. In between, the hydroxyl-headgroup of 6MH and the lipid headgroup there is a water reservoir offering an aqueous environment on both sides of the bilayer patches. The presented system was already successfully used for peptide^[37] and protein^[40] studies and fulfills the prerequisite for investigations of the membrane interplay of LL-37 the derivatives LL-32 and LL-20.

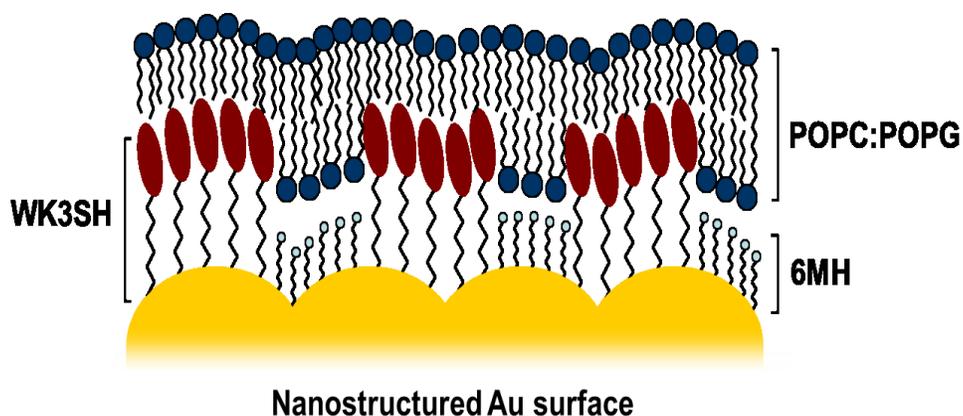


Figure 5.5: Schematic representation of a tethered bilayer lipid membrane (tBLM) adsorbed on a nanostructured Au-film. The thiol-bound self-assembled monolayer consists of hydrophobic WK3SH and hydrophilic 6MH-linker molecules. POPC:POPG (90:10) liposomes are adsorbed to the SAM-surface leading to alternating lipid bilayer and lipid monolayer islands. In between the lipid bilayer and the 6MH there is an aqueous reservoir. Adapted with kind permission from J. Kozuch.^[37]

5.6 Experimental Section

5.6.1 Liposome Preparation

First 50 μ l methanol and 50 μ l chloroform were mixed in a test tube. In a second step 9 μ l of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) in chloroform (25 mg/mL, Avanti Polar lipids) and 1 μ l of 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (sodium salt) (POPG) in chloroform (25 mg/mL, Avanti Polar lipids) were added to the prepared solution. After drying of the solution under a nitrogen stream, the application of vacuum overnight removed residual solvent traces and a lipid film was formed at the bottom of the tube. After addition of 0.5 ml 20mM Tris/100mM NaCl buffer at pH 7.4 a liposome solution of a concentration of 0.5 mM was obtained.

5.6.2 Fourier-Transform Infrared (FT-IR) Transmission

FT-IR transmission experiments were conducted using a CaF_2 window with a $2\ \mu\text{m}$ deepening. After deposition of $3\ \mu\text{L}$ of a peptide sample at a concentration of 5-20 mM and its homogeneous distribution over the whole surface within the deepening, the hydrated sample film was sealed with a plain CaF_2 cover window. An additional isolation of the sample was provided by a thin layer of silicone grease between both windows to avoid sample dehydration and the attached windows were placed in a suitable window holder. Infrared spectra were recorded at room temperature with a resolution of $4\ \text{cm}^{-1}$ at a spectral window from $4000\ \text{cm}^{-1}$ to $1000\ \text{cm}^{-1}$ using a Bruker IFS66v/s or a Tensor 27 spectrometer equipped with a liquid-nitrogen-cooled MCT detector (J15D series, EG&G Judson).

5.6.2 Surface-Enhanced Infrared Absorption (SEIRA) Spectroscopy

Based on the protocol of Miyake et al. a SEIRA active nanostructured Au-film was formed by electroless deposition on a trapezium shaped silicon prism ($W \times L \times H$: $20 \times 25 \times 10\ \text{mm}^3$)^[41]. The following electrochemical cleaning by cyclic-voltammetry (in $0.1\ \text{M}\ \text{HClO}_4$) led to an Au surface with a real surface area of about $1.65\ \text{cm}^2$, increasing the geometric surface of $0.79\ \text{cm}^2$ by a roughness factor of 2.1.

The Au-electrode was rinsed with water and 1-propanol and dried under a nitrogen stream before its functionalization by a tethered mixed self-assembled monolayer (SAM). The linker-molecule dihydrocholesteryl (2-(2-(2-ethoxy)ethoxy)ethanethiol (WK3SH)-linker containing a hydrophobic cholestanol-headgroup was mixed with hydrophilic 6-mercaptohexanol (6MH). Both molecules were incubated overnight at $4\ ^\circ\text{C}$ in a mixture of a molar ratio of 60:40 in solution leading to a ratio of 75:25 on the surface^[38].

Additionally, the prepared POPC:POPG (90:10) liposomes were adsorbed onto the SAM surface for 2-3 hours forming a negatively charged tethered lipid layer membrane (tBLM) system.

The successful assembly of the SAM and tBLM system was validated using electrochemical impedance spectroscopy (EIS) by determination of the capacitance of the system, which facilitates calculation of its resistance. The EIS measurements were

performed using a μ AutolabIII/FRA2 instrument and FRA software. A three electrode configuration with an Ag/AgCl (3 M KCl) reference electrode, a Pt mesh as counter electrode and the Au film as the working electrode was applied to record EIS spectra in the frequency range of 0.05 Hz to 100 kHz at a DC of potential of 250 mV (vs Ag/AgCl) and amplitude of 25 mV (rms).

SEIRA spectroscopic measurements were performed in the attenuated total reflection infrared (ATR-IR) mode in a Kretschmann configuration under an angle of incidence of 60° . All spectra were recorded in a spectral window of 4000 to 1000 cm^{-1} and with a resolution of 4 cm^{-1} using a Bruker Tensor 27 or a Bruker IFS66v/s spectrometer with a liquid nitrogen cooled photovoltaic or conductive MCT (Mercury Cadmium Telluride) detector, respectively. For each spectrum, 400 scans were accumulated. The calculation of difference spectra allowed the distinction of each incubation step by subtraction of the absorbance spectra of two consecutive steps, i.e. the spectra of peptide incubation from the spectrum the previously formed tBLM.

Chapter II

6- Biomembranes

7- Antimicrobial Mechanisms of Action

6. Biomembranes

Eukaryotic (animals and plants) and prokaryotic (bacteria and archaea) organisms^[42] differ in many aspects, such as the cell size (10-100 μm for eukaryotes and 1-10 μm for prokaryotes systems), although the most relevant difference is the absence of a wide type of components in the less evolved prokaryotes organisms, for instance lack of: nucleus, lysosomes and peroxisomes, microtubules, endoplasmatic reticulum, mitochondria, Golgi apparatus or chloroplasts. Despite of this fact, they also share a key factor in common: the presence of a biomembrane to enclose all the integral cell elements.

Biomembranes are constituted by the reordering of lipids. Lipids are molecules characterized by a hydrophobic (apolar) and a hydrophilic (polar) region, showing an amphipathic character. There are mainly three types of lipids: the phospholipids (PLs) (most abundant in the lipid bilayers), cholesterol, and glycolipids. In A-figure 6.1 is shown the chemical structure of the PLs used in this thesis: the zwitterionic lipids POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), and negatively charged lipid POPG. For all lipids in general, the hydrocarbon tails can vary in length, being the normal range between 14 to 24 carbon atoms. Additionally, one of the tails commonly shows one or more cis-unsaturations while the other tail remains totally saturated. The presence of double bonds in the tail promote the formation of small kinks. This fact combined with the differences in the length affects the grade of compactness in their packaging and consequently, the fluidity of the membrane.

Due to its amphipathic character, the lipids in aqueous environment are able to aggregate in liposomes, micelles (spherical shape) or sheets (bilayers) as it is illustrated in B-figure 6.1. Despite the fact that they enclosed themselves in global well-ordered structures, they are able to move allowing rotation, lateral diffusion or flexion. Additionally, studies carried out with electron spin spectroscopy showed that lipid molecules in synthetic bilayers could rarely migrate from a monolayer on one side to the other. This typical movement was named as “flip-flop”.

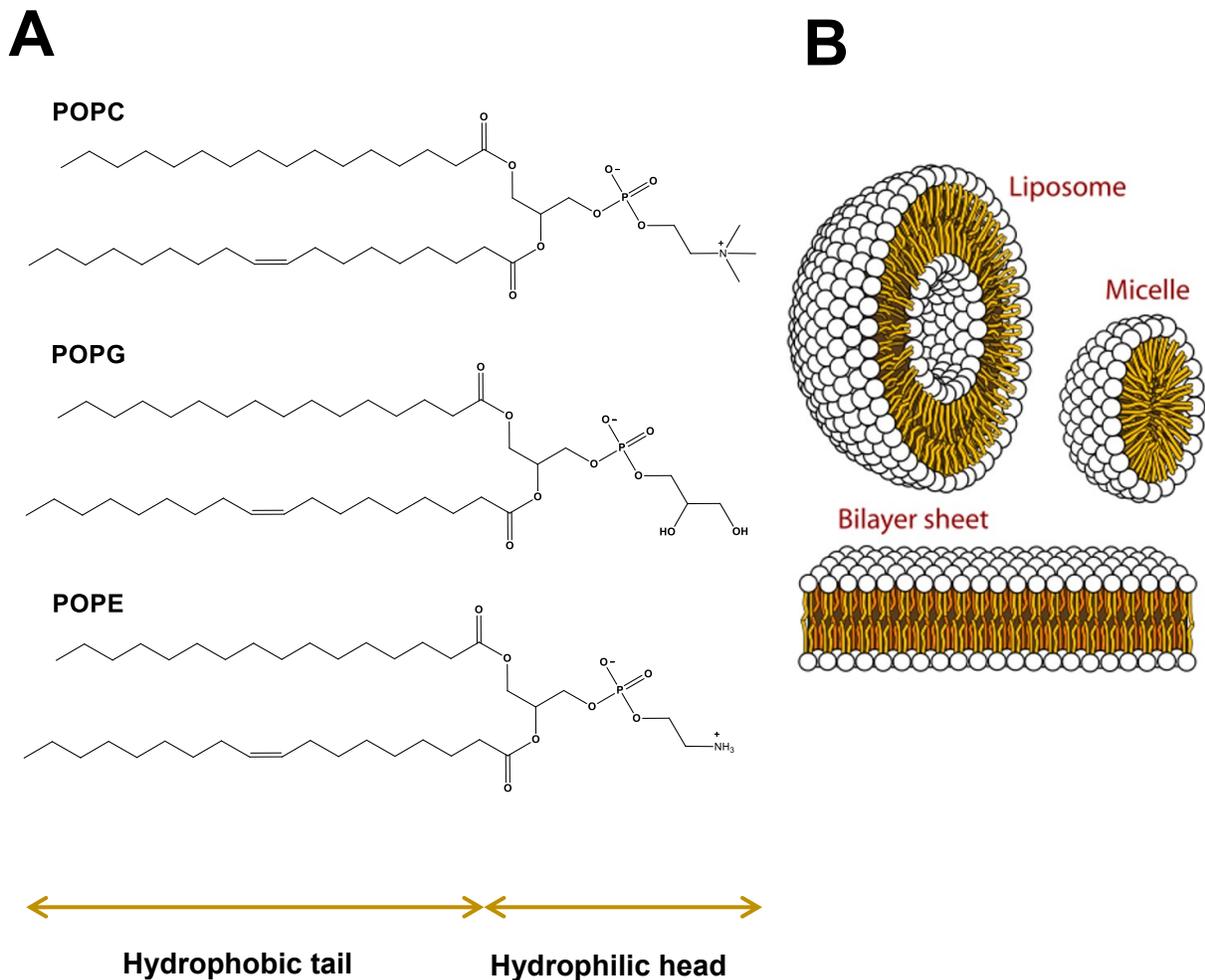


Figure 6.1: A- Chemical structure of the PLs used in this thesis. POPC, POPG and POPE. B- Ordered structures adopted by the lipids. The amphipathic character consisting of a hydrophilic head combined with hydrophobic tail promote them to aggregate in well-confined structures.

6.1 Fluidity

According to the temperature phase transition, lipid bilayer changes from a liquid to a rigid crystalline state. Lower temperatures indicate that the membranes are more difficult to freeze, most probably due to two factors: i) the presence of short chain hydrocarbon tails, reducing the tendency of interacting with each other, and ii) the presence of kinks caused double bonds hindering their packing.

The cholesterol plays also an important role in the fluidity of the membranes in animal cells. It is suitable to find one cholesterol molecule (figure 6.1.1) per phospholipid molecule. The cholesterol provides mechanical stability in such a way that places its hydroxyl group close to the polar head groups of the PLs while the rigid and steroid rings interact with the hydrocarbon tails closest to the polar head groups leaving the rest of the chain still flexible. By this way, the membrane fluidity turns out to be limited. Even high concentrations of cholesterol found in eukaryotic plasma membranes avoid possible phase transitions. Conversely, presence of cholesterol decreases the membrane permeability.

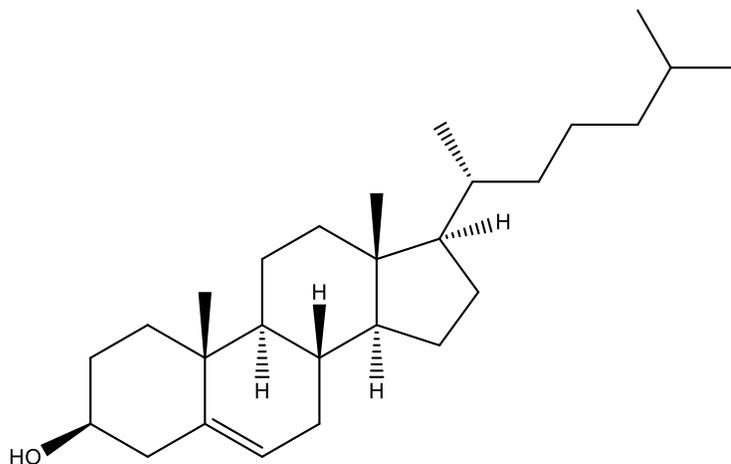


Figure 6.1.1: Cholesterol chemical structure. Characterized by a hydroxyl group and steroid rings providing mechanical stability to lipid membranes.

6.2 Eukaryotic Cells

The eukaryotic-mammalian cells are principally composed by a variety of lipids, (mainly POPC, POPE phosphatidylethanolamine, sphingomyelin, phosphatidylserine) at physiological pH all of them are zwitterionic lipids with the exception of phosphatidylserine, which is negatively charged. This thesis was focused on the bactericidal activity of the AMPs which is why the eukaryotic cell membranes will not be explained.

6.3 Prokaryotic Cells

Bacterial membranes are predominantly composed of one type of phospholipid and there is absence of cholesterol, obtaining a determined mechanical stability by the presence of a overlaying cell wall.

6.3.1 Bacteria

Bacteria are found in water, soil (including in the deep Earth's crust), or even under inhospitable conditions such as acidic thermal springs, below zero degrees or in the nuclear waste. The bacteria can be distinguished by their shape (figure 6.3.1.1), finding mainly three classes: coccus (round shape), bacillus (stick-like) and spiral (twisted) although star- and rectangular- shaped have been also found are less common.

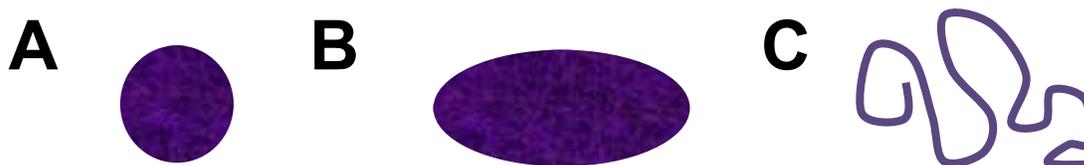


Figure 6.3.1.1: Main bacterial shapes. A- Coccus (round shape). **B-** Bacillus (stick-like). **C-** Spiral (twisted).

Due to their pleomorphic properties, the bacteria cannot be classified exclusively according to that criteria, which is why the Gram-staining is commonly used. The Hans Christian Gram (1853-1398) technique (Gram staining method^[43]) discovered in the city hospital in Berlin in 1884 allowed to classify the bacteria in two groups: 'Gram-positive' and 'Gram-negative'. Gram-positive bacteria are those which are able to take up a violet coloured under the microscope after the Gram-stain test while the Gram-negative provide a negative result because no crystal violet stain used in the test is retained (figure 6.3.1.2). The different staining relates to their cell wall composition.

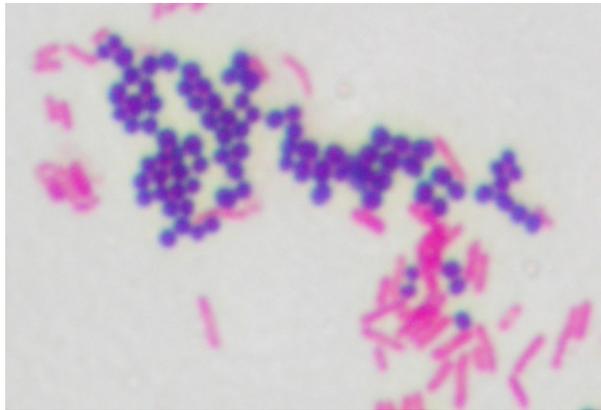


Figure 6.3.1.2: Gram staining method. Bacterial classification in two groups: 'Gram-positive' and 'Gram-negative'. Gram-positive bacteria take up a violet colour under the microscope after the Gram stain test while the Gram-negative provide a negative result.^[44]

6.3.1.1 Gram-Negative Envelope

The Gram-negative bacteria are characterized by the presence of two membranes, the inner or cytoplasmic (IM) and the outer membrane (OM), enclosing between them a protein-rich periplasmic space containing a peptidoglycan (PGN) layer usually of 7-8 nm of thickness.^[45]

The OM varies in composition, whereas the external leaflet is mainly composed of glycolipids, principally lipopolysaccharide (LPS), lipoproteins (LPP) and β -barrel proteins, the internal leaflet contains PLs.^[46]

The LPS are exclusive components of Gram-negative bacterium and involve three parts: 1) the lipid A, 2) the oligosaccharide core and 3) the O antigen. The presence of lipid A is essential for the bacteria survival in opposite to the O-antigen part which is not crucial for the correct bacterial function.

LPP and the β -barrel proteins represent the proteins of the OM. In contrast to the LPP, the transmembrane proteins adopt a β -barrel conformation through of the association of the β -sheets in cylindrical manner. Due to the adoption of this specific arrangement, they are usually denoted as outer membrane proteins (OMPs). A large variety of OMPs exist, the most abundant is the OmpA but also other porins such as OmpF and OmpC are present,

usually existing as trimers. Their main function is to allow the passive diffusion of small molecules.

The PGN layer consist of a long polymer of repeating disaccharide N-acetylglucosamine (GlcNAc)-N-acetylmuramic (MurNAc) linked through pentapeptide side chains.

The IM is a PL bilayer. The most common PLs are POPE and POPG finding some small amounts of phosphatidylserine (POPS), cardiolipin and polyisoprenoid carriers (C55) as well.

In the figure 6.3.1.1 is displayed a graphical representation of the Gram-negative cell envelope.^[45]

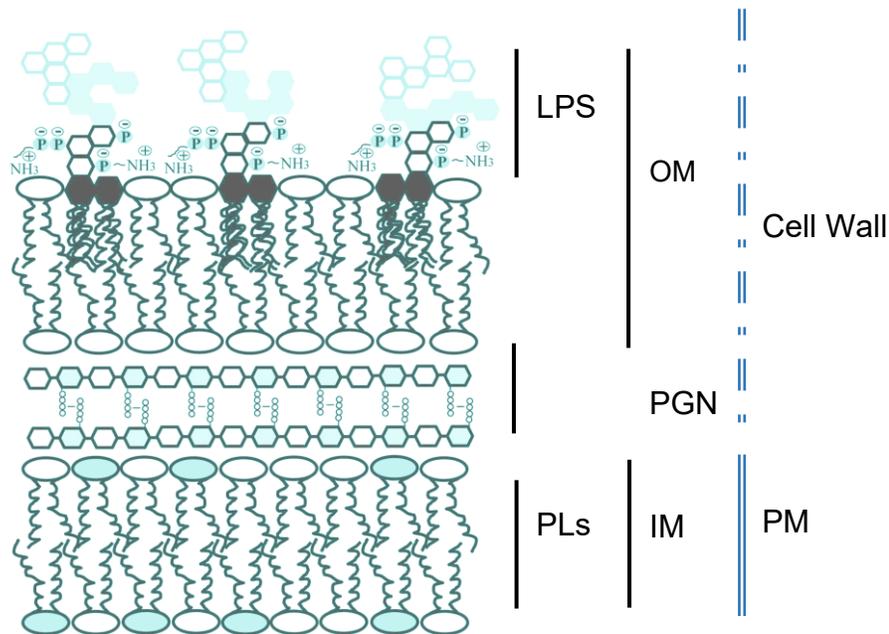


Figure 6.3.1.1: Gram-negative bacteria envelope. Cell wall is composed by the OM and PGN layers while the PM consists of PLs. Figure displayed with kind permission from N. Malanovic.^[45]

6.3.1.2 Gram-Positive Envelope

The main difference of the Gram-positive bacteria in comparison with the negative ones is the lack of OM. Instead of the OM, a 40-80 nm thick PGN layer is present.^[45] Interweave between these PGN layers a variety of components can be found, such as anionic polymers,

commonly called teichoic acids (TAs), and proteins.^[43] There are two classes of TAs depending on the component which anchor to, if there are covalently attached to the PGN, are namely as wall teichoic acids (WTAs) while if they are attached to the head group of the membrane lipids are denoted as lipoteichoic acids (LTAs).^[47]

The proteins found in the Gram-positive bacteria envelope are analogous to those found in the Gram-negative counterpart. In this case, due to the lack of OM, the LP can be attached or associated to the PGN, LTA or to WTA. In the figure 6.3.1.2 is displayed a graphical representation of the Gram-positive cell envelope.

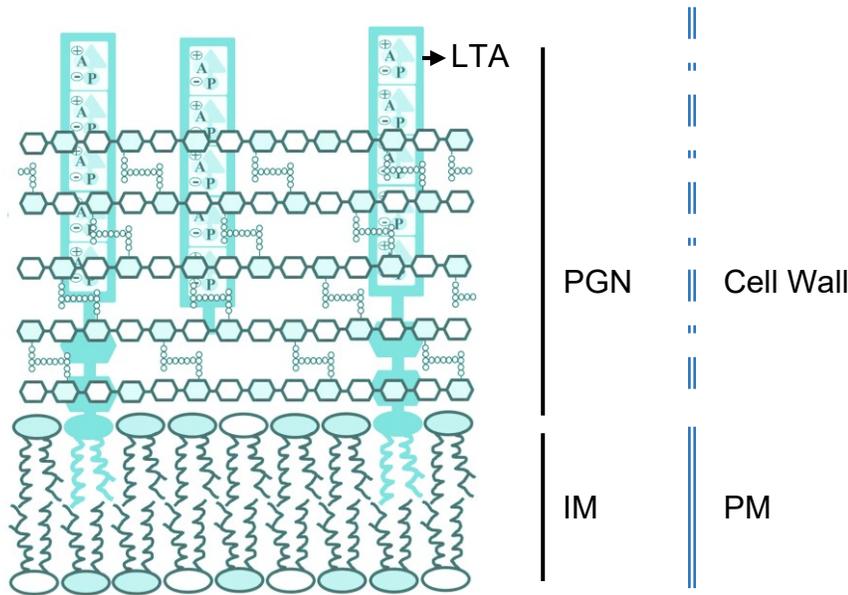


Figure 6.3.1.2: Gram-positive bacteria envelope. Cell wall is composed of a dense layer of PGN and lack of OM in comparison with the Gram-negative bacteria. PM is characterized by high amounts of POPG. Figure adapted with kind permission from N. Malanovic.^[45]

6.3.2 Lipid II - Peptidoglycan Cell-Wall Precursor

The lipid II is the structural subunit of the PGN layer. PGN is a complex polymer of alternating glycan, MurNAc and GlcNAc, residues cross-linked by a short peptide appended to MurNAc. The glycan composition is well-conserved between the species and the peptide stem composition typically comprises a sequence of L-alanine, γ -D-glutamyl-diaminopimelyl (or L-lysyl)D-alanyl-D-alanine. The peptapeptides are usually 4-3 and less common 3-3 crosslinked providing mechanical stability to the cell. The pentapeptides, in Gram-positive

bacteria are connected through inter-bridge structures consisting of five amino acids. The AA composition varies among the species. Figure 6.3.2.1 shows a graphical representation of the PGN.^[45]

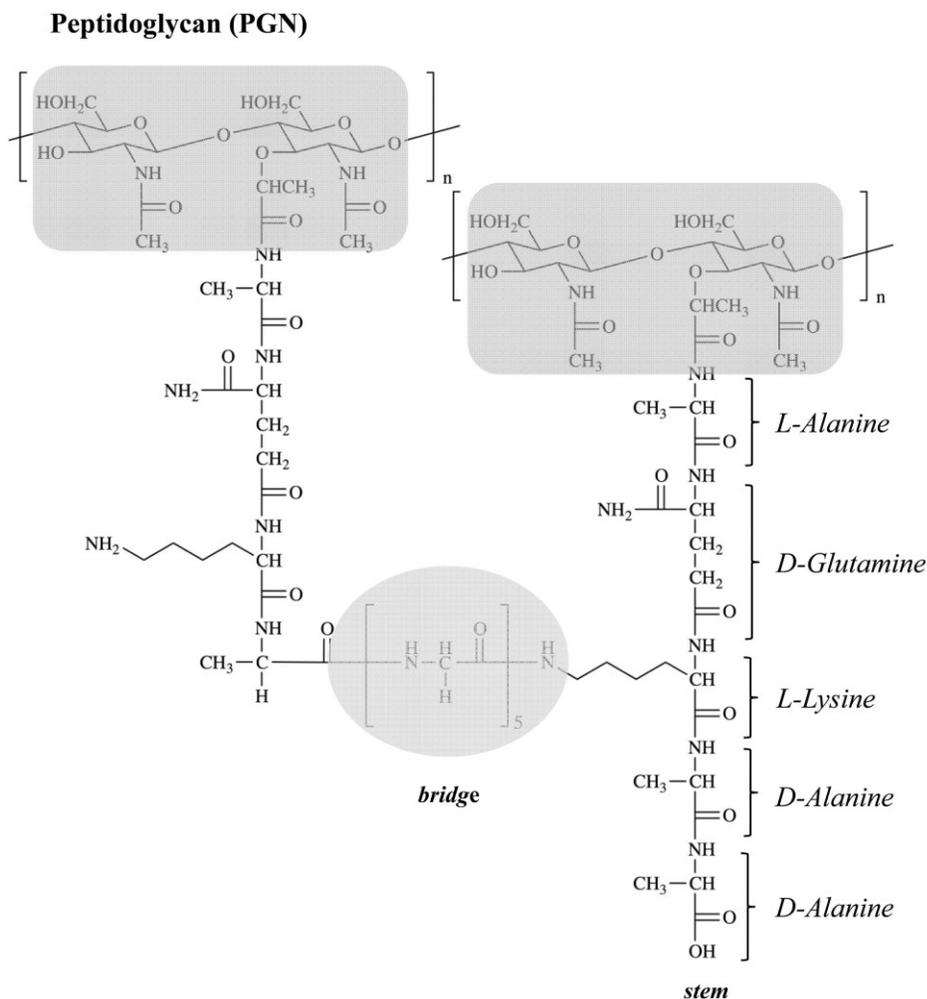


Figure 6.3.2.1: Chemical structure of PGN. Figure displayed with kind permission from N. Malanovic.^[45]

6.3.2.1 Lipid II - Peptidoglycan Cell-Wall Synthesis

The synthesis of the cell wall takes place stepwise. Firstly, on the cytoplasmatic side of the PM the coupling of the UDP-MurNAc to the bactoprenyl-phosphate to yield Lipid I is produced. Afterwards, the Lipid II is obtained by the association of the GlcNAc sugar due to the peripherally membrane-associated protein MurG. In this way, the complete peptidoglycan subunit linked via a pyrophosphate to the membrane-embedded

polyisoprenoid anchor is achieved. At the end, through an unknown mechanism the lipid II is translocated to the periplasmatic side of the plasma membrane.

The synthesis of lipid II is limited to the amount of bactoprenyl phosphates molecules which is approximately 2×10^5 molecules per cell. Interestingly, in Gram-positive bacteria the PGN is 20 layers thick while in the Gram-negative bacteria the average layer thickness is about 1.5 layers.

7. Antimicrobial Mechanisms of Action

The AMPs constitute the first defence line in the innate immune system. Like lipid membranes which enclose the cytosolic organelles, AMPs are also amphipathic compounds and in most of the cases are positively charged at physiological pH. There are specific features associated to the antimicrobial activity and selective toxicity of AMPs highlighting i) conformation, ii) charge, iii) hydrophobicity, and subsequently, iv) amphipathicity. These molecular fingerprints are interconnected such that an alteration in one feature will promote a compensatory modification in the others.

For example, an increase of the overall positive charge in the AMP is usually associated with an improvement of the antimicrobial activity^{[48][49]} due to the fast approaching of the peptides to the negatively charged phospholipid membranes of bacteria. It should, however, be noted that always a boundary is present: if the increase of the positive charge is too high, the peptides interact excessively strong with the phospholipid head groups avoiding the translocation of the peptide into the cell interior and consequently, decreasing or losing the antimicrobial activity.^[50] Additionally, an adequate and delicate balance between a specific hydrophobicity and consequently, amphipathic conformation promote an effective membrane permeabilization.^[51]

The optimal balance between the aforementioned features is essential for the AMPs to promote the cell death:

- First, an initial peptide-approach to the membrane target is required. Moreover, factors like threshold concentration and conformational phase transition have to be considered, in order to understand the different hypothesis which have been elucidated to explain their mode of action.
- Second, after a self-association promoting multimerization^[52], execution of their activity through one of the proposed mechanisms i) barrel-stave mechanism, ii) toroid-pore mechanism or iii) carpet mechanism take place.

- Third, the peptide-mediated cell death is produced by a ‘multi-hit process’^[53] involving from a membrane dysfunction, to complementary events such as the inhibition of extracellular biopolymer synthesis and/ or the inhibition of the intracellular functions.

7.1 Approaching to Membrane Targets

It is generally accepted that the electrostatic interactions drive the initial attraction of the positively charged peptides to the negatively charged lipid membrane in Gram-positive bacteria. Alternatively, for Gram-negative bacteria the self-promoted uptake mechanism^[54], focused on a competitive displacement of LPS-associated divalent cations by the peptides, was proposed.

Additionally, some AMPs employ a denoted receptor-mediated membrane interaction, due to their targeting to unique epitopes on the microbial surface, such as the case of the lantibiotics to the bacterial lipid II component.

7.2 Execution of Activity

A threshold concentration, accumulation of the peptides on the surface, is required to allow the subsequent step to enter and cross the lipid bilayer. This requirement, together with the parallel-to-transmembrane surface orientation of the peptide, is undoubtedly influenced by the membrane potential $\Delta\psi$. Additionally, for α -helical AMPs a conformational phase transition must be produced. It has been well-documented that several AMPs are disordered in solution, finding from random coil to extended conformations, which upon interactions with the lipid bilayers are able to adopt well-structured α -helices conformations.^[54] In contrast to the α -helical AMPs, the β -sheets AMPs are ordered in solution as well as in lipid environments due to the presence of stabilizing disulphide bridges. Interestingly, the quaternary structures of such peptides in solution can be affected upon interaction with the membranes resulting in a structural disassociation and possibly promoting an effective monomeric antimicrobial mechanism.

7.2.1 Self-Association and Multimerization

The AMPs tend to build complex quaternary structures due to their intrinsic physiognomies. These amphipathic peptides reorient their hydrophobic and hydrophilic domains accordingly to the membrane environment or adjacent peptides contributing to the creation of transmembrane pores or channels.

Attributable to the variability of microbial membrane structure, different permeabilizing mechanisms are expected. Unfortunately, up-to-date there is not a general agreement in this regard, and several mechanisms have been proposed:

7.2.2 The Carpet Model

The electrostatic interactions allow an initial binding of the peptides to the target membrane. Subsequently, a high accumulation of the peptides over the surface induce a displacement of the phospholipid head groups, which combined with a decrease of the barrier properties of the membrane promote its disruption. In this model, the membrane integrity is lost due to the unfavourable energetics caused by the achievement of a specific threshold peptide concentration on the surface. In this scenario, neither channel formation is detected nor the peptides are inserted into the membrane core (A-figure 7.2.1).

7.2.3 The Toroidal Pore

The main difference with the Barrel-Stave model (B-figure 7.2.1) is that the peptides approach the membrane in a perpendicular manner and cooperatively interact with the lipids to generate a toroid or “wormhole” channel. Studies focused on the stability-disintegration and half-life^[55,56] of such pores, resulted in the translocation of the peptides to the cytoplasmic leaflet suggesting that this mechanism could most probably be employed by the peptides to access to most potential intracellular targets.^[57]

7.2.4 The Barrel-Stave Mechanism

The peptides bind to the membrane and a conformational phase transition promotes a localized membrane thinning by the displacement of the polar-head phospholipid groups. In this scenario^[54], the hydrophobic side of the peptide is inserted into the membrane. A threshold concentration must be achieved to promote the self-aggregation of the peptides (protein-protein interactions “staves”) inserting themselves deeply into a water-filled channel (“barrel”). Increase of the aggregate-size is understood as the enlargement of the membrane pore. Consequently, due to the lipid translocation, the peptides are able to get the inner leaflet membrane as a result of the concentration gradient of the surface-bound peptide and the trans-negative $\Delta\psi$ (C-figure 7.2.1)

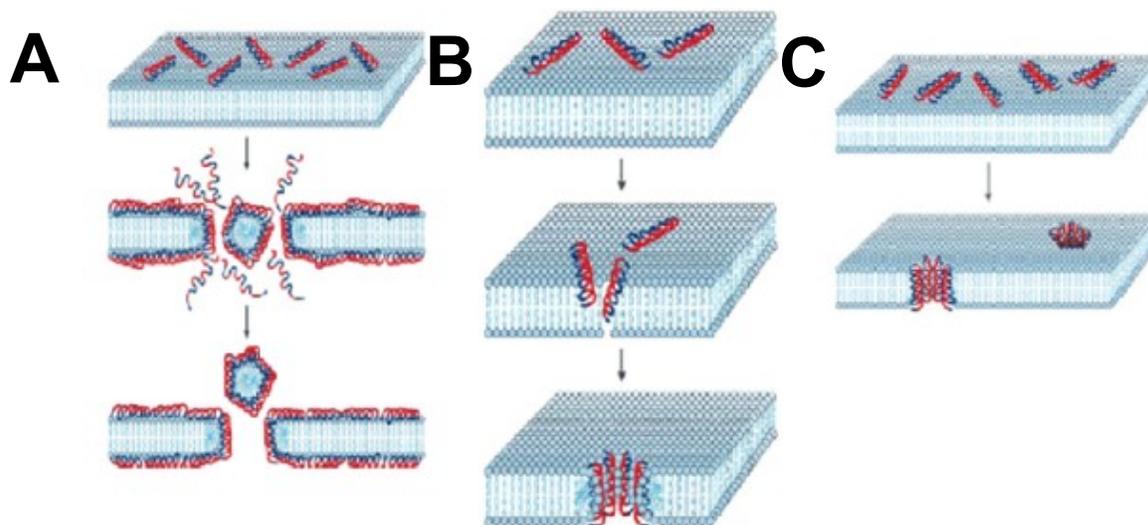


Figure 7.2.1: Proposed mechanism of action. A- Carpet model. **B-** Toroidal pore. **C-** Barrel Stave model^[58].

7.3 Cell Death

Cell death is caused by a consequence of the AMPs activity. The membrane perturbation leads directly or indirectly to the cell death because several complementary events are involved in the membrane dysfunction such as: leakage of ions and metabolites, depolarization and blocking of the biopolymer synthesis. Additionally, the peptide translocation across the membrane allows them to get intracellular targets.

Chapter III

8- Bacteriocins

A- Nisin Lantibiotic

B- Two component Lantibiotic Lichenicidin

8. Bacteriocins

Microbes show a broad-spectrum of microbial defense systems, among others, the bacteriocins. The designation 'bacteriocine' was devised for first time to describe an archetype Gram-negative bacteriocin.^[59] Nevertheless, many years beforehand Pasteur I and Joubert (1877) pointed out for first time that *bacteriocin-mediated inhibition*. This fact was attributable to their results regarding the inhibition of *Bacillus anthracis* by bacteria isolated from urine analysis.

Up-to-date, bacteriocins have been found in almost all leading bacterial families (Eubacteria and Archaeobacteria) examined^[60], which have been defined in an early stage described as proteinaceous compounds with bactericidal mode of action. Concisely, they were characterized by a narrow killing range directly associated to the producing strain via selected attachment to specific epitopes. Initially, it was proposed that their production boost the death of the precursor organism. Due to the complexity of such components and constant advances in the bacteriocin research, the initial definition has been currently modified as *modified or unmodified peptide antimicrobials produced by bacteria which are protected by a dedicated immunity system*.^[59]

Two classes of bacteriocins can be found in Gram-negative bacteria: the colicins and the microcins^[59], while in Gram-positive bacteria the classification scheme becomes more tedious. Bacteriocins produced by Gram-positive bacteria, concisely those from lactic acid bacteria (LAB) have been intensely studied due to their potential properties as food preservatives and as 'tools' for medical treatments in humans and animals.^[61,62]

Additionally, Archaea are also bacteriocins-like peptides producers, commonly referred as archaeocins which can be classified as halocins or sulfobiocins.

Since the focus of this thesis lies on bacteriocins produced by Gram-positive bacteria^[63], in the following, those will be explained in detail.

8.1 Bacteriocins of Gram-Positive Bacteria

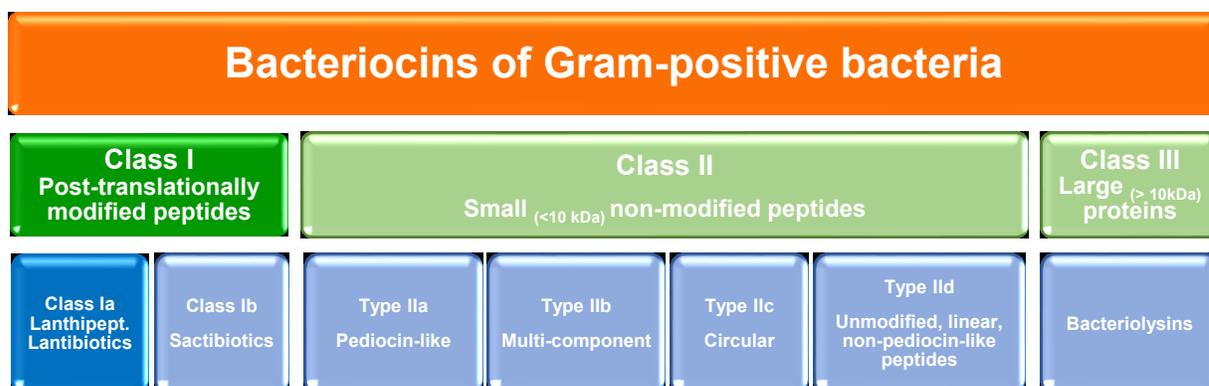
The role of the bacteriocins in nature has been always questioned, it seems that the determining factor is to provide the producing organism an environmental benefit over possible competitors.^[59,63,64]

Due to their heterogeneity, classification of the bacteriocins from Gram-positive bacteria results very complicated. The initial LAB bacteriocins classification scheme proposed by Geis *et al.* (1983)^[65] was out-dated by Klaenhammer^[66] in 1993. Klaenhammer-scheme has served as basis for the posterior classifications.^[59]

Despite the discrepancies found in the proposed classifications, the scientific community agreed with the general subdivision for the Class I - Lanthionine and Class II - non-lanthionine containing peptides. Alternative classifications based on the presence of disulphide and monosulphide bonds^[63], or based on the structure^[67] have been also proposed.

The existence of several and continuously-modified schemes reflect the complexity and sophistication of these small (non-)bactericidal-weapons. The discovery of new bacteriocins combined with development of physico-chemical advances allow to classify them in an improved refined manner, therefore, further changes in the classification tree are expected in the following years.

Rea *et al.*^[59] adapted the latest classification scheme for bacteriocins of Gram-positive bacteria proposed. The proposed subclass Ib-Labyrinthopeptins by Rea *et al.* is included in the subclass Ia, related to Meindl *et al.* (2010)^[68] and Knerr *et al.* (2012)^[69] observations (scheme 8.1.1).



Scheme 8.1.1: Adapted the latest classification scheme for bacteriocins of Gram-positive proposed by Rea *et al.*^[59]. The proposed subclass Ib-Labyrinthopeptins by Rea *et al.*^[59] is included in the subclass Ia, according to Meindl *et al.* (2010)^[68] and Knerr *et al.* (2012)^[69] observations.

8.1.1 Class I: Lanthionine - Containing Bacteriocins

Initially, this group was designated only for the post-translationally modified peptides, the lantibiotics. Currently, recently discovered post-translationally bacteriocins promoted the following subclassification: Ia) originally lantibiotics, currently named as lanthi peptides and Ib) the sactibiotics. The work reported in this thesis is specifically focused on the Nisin and Lichenicidin lantibiotics belonging to the Ia), which are going to be explained in detail after an introductory general description of this subclass.

8.1.1.1 Class Ia: The Lanthi peptides - Lantibiotics

The family name lantibiotic has evolved to 'lanthi peptides'^[69,70] in order to encompass those posttranslational modified bacteriocins with non-antimicrobial activity, such as the case of SapB, AmfS and SapT.^[59]

The lantibiotics are ribosomally synthesized as propeptides and posttranslationally modified peptides (RiPPs) to their biologically active forms.^[71] During the posttranslational modifications (PTMs), the thioether cross-links of *Meso*-lanthionine (Lan) and (2*S*,3*S*,6*R*)-3-methylanthionine (Mlan) are introduced, which is the reason for the conferred name

8. Bacteriocins

containing lanthionine antibiotics. The presence of lanthionine does not make an AMP directly lantibiotic^[59] due to the aforementioned lack of antimicrobial activity.

The thioether bridges (figures 8.1.1.1.1 and 8.1.1.1.2) are yield by means of the dehydration of serine and threonine residues to the unsaturated amino acids 2,3-dehydroalanine (Dha) and (Z)-2,3-dehydrobutyrine (Dhb) followed by an intramolecular Michael addition.

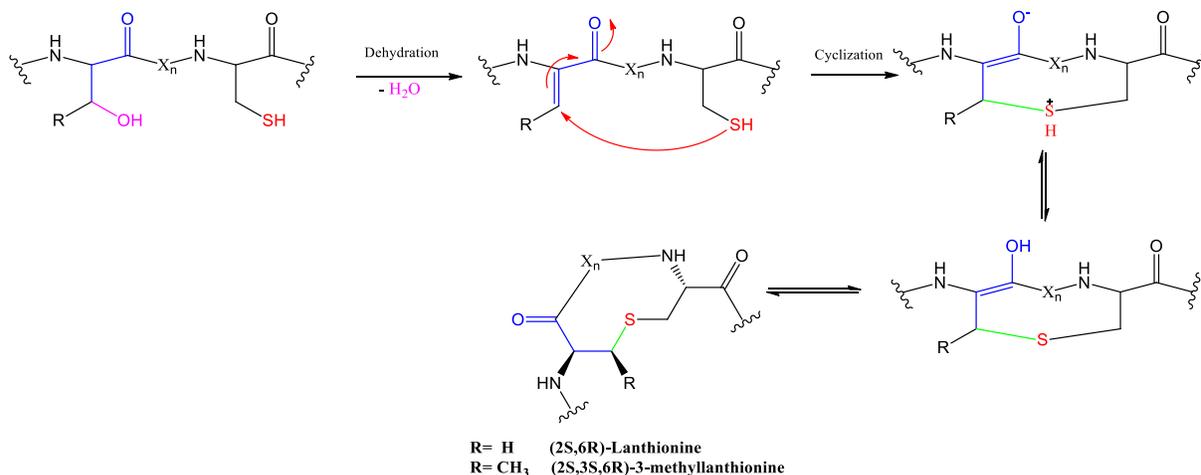


Figure 8.1.1.1.1: Mechanism of thioether cross-link formation in lantibiotics. Dehydration of Ser and Thr followed by an intramolecular Michael addition to yield Lan and Mlan rings.

The PTMs are the source of dehydrated amino acids like 2,3-dehydroalanine (Dha) and (Z)-2,3-dehydrobutyrine (Dhb) (figure 8.1.1.1.2), which are frequent elements of the peptidic backbone. Alternatively, residues as S-aminovinyl-D-methycysteine, S-aminovinyl-D-cysteine, lysinoalanine, 2-oxybutyrate, 2-oxopropionate, 2-hydroxypropionate, β -hydroxy-aspartate are also common but less usual.

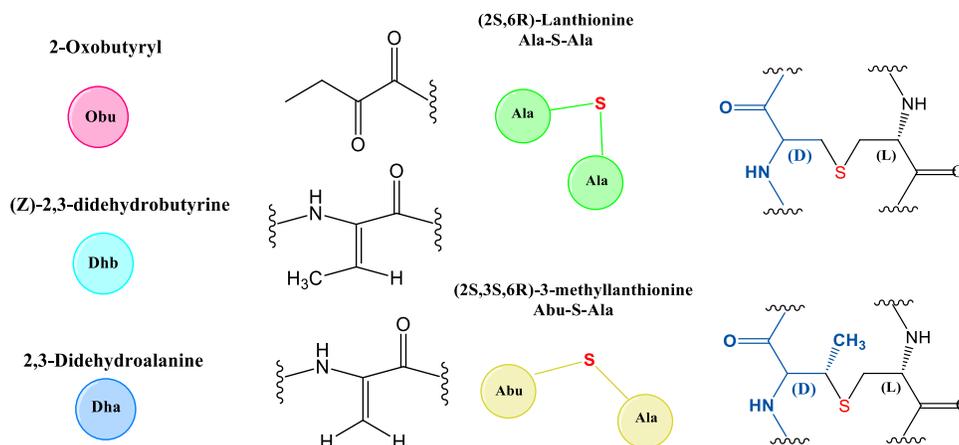


Figure 8.1.1.1.2: Residues obtained after the PTMs. Dehydroamino acids Dha and Dhb, together with Lan and Mlan rings.

Lantibiotics show a relevant medical impact due to their high activity against Gram-positive bacteria such as *Staphylococcus*, *Streptococcus*, *Enterococcus* and *Clostridium* and as well as powerful activity to certain Gram-negative pathogens like *Neisseria* and *Helicobacter*.^[69] A broad spectrum of variables involving: the shape, the number of peptides for optimal activity, the enzyme employed for the posttranslational changes and secretion and primary structure in the propeptide have been used to develop extensive classification schemes comparable to those developed for bacteriocins. The updated summary from the initial scheme designed by Rea *et al.*^[59] is provided in the table 8.1.1.1.1.

8. Bacteriocins

Classification proposed by	Classification-scheme		
Jung 1991 ^[72]	Type A Elongated, membrane acting cationic peptides	Type B Globular, active through cell wall inhibition	
Klaenhammer 1993 ^[66]	No subclassification		
Piper <i>et al.</i> 2009 ^[73] Cotter <i>et al.</i> 2005 ^[74]	Based on the similarities in the primary structure of the propeptide: 12 Subclasses Classification protocol initially used for LAB lantibiotics by Twomey <i>et al.</i> (2002) ^[75]		
Pag & Sahl 2002 ^[76] Willey & van der Donk 2007 ^[77]	Based on the enzyme employed for the posttranslational changes and secretion:		
	Subclass I Modification-LanB/C Export-LanT Leader cleavage–LanP	Subclass II Modification-LanM Export/ Cleavage LanT	Subclass III Modification-RamC No antimicrobial activity.
Heng <i>et al.</i> 2007 ^[78]	Type A Linear	Type B Globular	Type C Two-Component
Goto <i>et al.</i> 2010 ^[70]	Agreement to Pag & Sahl ^[76] , and Willey & van der Donk ^[77] subclassification, and further extended to the Subclass IV Modification-LanL/ Export-LanT-like Employed for first time the term of 'lantipeptide' to denote: " <i>those compounds that by structure and biosynthetic strategy are clearly related to lantibiotics but that are not known to possess antimicrobial activity</i> "		
Meindl <i>et al.</i> 2010 ^[68]	Agreement to Pag & Sahl ^[76] , and Willey & van der Donk ^[77] subclassification. Discovery of the Labyrinthopeptins which are included in the Subclass III. Characterized by the presence of carbacyclic labionin. Modifications-LabKC/ Export-LanT-like		

Table 8.1.1.1.1: Summary of the proposed classification schemes.

According to the latest classification accepted^[77], an extensive explanation will be made for the classes whereas the research has been developed, for instance, Subclasses I (Nisin) and II (Lichenicidin). More information about the diverse lanthipeptide family can be found in the following reviews.^[59,79]

8.1.1.1-A Subclass I Lanthipeptides: Dehydratase LanB and Cyclase LanC Enzymes

Two enzymes catalyse the reactions to yield the Lanthionine and Methyllanthione cross-links, LanB and LanC. LanB dehydrates the threonine and serine residues while LanC mediates the cyclation. A transmembrane ATP-binding cassette (ABC) transporter, LanT, is employed to export the modified peptide. LanP, subtilisin-like serine protease, is used to cleavage the leader peptide. The peptide, which exemplifies this subclass, is the Nisin lantibiotic.

A- Nisin Lantibiotic

A.1 Nisin Lantibiotic: Experimental Observations

Despite that its discovery date was made in 1928 in England, it was not until 40 years later when the elucidation of its structure and properties was initiated^[80]. In the meanwhile, Whitehead and co-workers in New Zeland followed by Meanwell (1943) and Mattick & Hirsch (1944) also reported the existence of lactococcal strains, which showed microbial antagonism inhibiting the growth of other LABs.^[64]

Nisin, initially called 'group N Inhibitory Substance'^[64] (the termination 'in' is related to antibiotic), has been the first antibiotic used as preservative in the food industry in several countries^[81,82] over the last decades. Furthermore, its large field of application has been exemplified successfully in the shelf-life of dairy products (hard and semi-hard cheeses, processed cheeses and cheese spreads, heat-processed milk and milk drinks). Additional advantages of nisin in food preservation have been demonstrated by the decrease of i) the spread exposure of pathogen food borne ii) the employment of other chemical preservatives. Its use permits soft food treatments allowing an improved conservation of vitamins and organoleptic properties.

8. Bacteriocins

So far, eight different nisin natural variants have been discovered. Nisin A produced by *Lactococcus lactis* was the first one to be studied and characterized. Afterwards, from the same lactococci were isolated the nisins Z, F, and Q variants^[83] (structural composition differences in comparison with Nisin A displayed in table A.1.1, graphical representation in figure A.1.1). Interestingly, from *Streptococcus uberis* was isolated the nisin U and U2^[84] which show more exchanged amino acids and a lack of the last three residues in the C_t region in comparison with nisin A^[85] (table A.1.1 and figure A.1.1). Nisin P is segregated from *Streptococcus gallolyticus* subsp. *Pasteurianus*^[86] and *Streptococcus suis*^[87] which possess similar exchanged residues to Nisin U and U2 (see figure A.1.1) and the same length. Moreover, nisin H^[88] has been sequenced and analysed very recently by a strain of *Streptococcus Hyointestinalis* and it differs from nisin A in five positions. According to the structural characteristics found in Nisin H, it seems to be an intermediate between nisins of lactococcal and streptococcal origins.

According to the definition of 'natural variant', it is expected that i) the associated producers should exhibit cross immunity to other variants^[89] ii) despite that nisins are produced by different bacteria, the residues sensitive to change are placed in the same position (with the exception of residues 6 and 18 only exchanged in nisin H) (see table A.1.1), and iii) all peptide must show the same ring pattern: the lanthionine bridge (ring A involving the residues 3 to 7), as well as the methyllanthionine bridges (ring B: residues 8 to 11, ring C: residues 13 to 19, ring D: residues 23 to 26 and ring E: residues 25 to 28) are unalterable amino acids in any of them. The rings are separated by the hinge region involving the amino acids 20 to 22 – residues 20 and 21 extremely sensitive to be exchanged.

With the exception of variants U, U2 and P, the rest are composed by 34 amino acids. All share the presence of dehydrated amino acids, such as dehydrobutyrine Dhb2 and dehydroalanine Dha5 and Dha33. Interestingly, only in natural variants U, U2 and P is feasible to detect an additional dehydroamino acid at the position 17th. It has been also suggested to denote the nisin U and U2 as nisin-like peptides, as it is the case of salivaricin D, due to the impossibility of nisin A machinery to generate the aforementioned variants.^[83]

The presence of thioether cross-linkages provide them specific functionalities such as thermostability, and attributable to the presence of very reactive double bonds unique mode of action may also be associated. All variants show a specific amphipathicity which

will affect to their mode of action. Furthermore, with the exception of nisin H, the lack of aromatic amino acids hinders their absorbance at 260 or 280 nm. So far, it has been reported that bioengineered nisins containing aromatic residues displayed reduced activity.^[90,91] This fact is also reflected in nisin H (unique natural variant with aromatic residues) which in comparison to nisin A at equal concentrations and purity, showed less activity against the plenty target microorganisms studied.^[92]

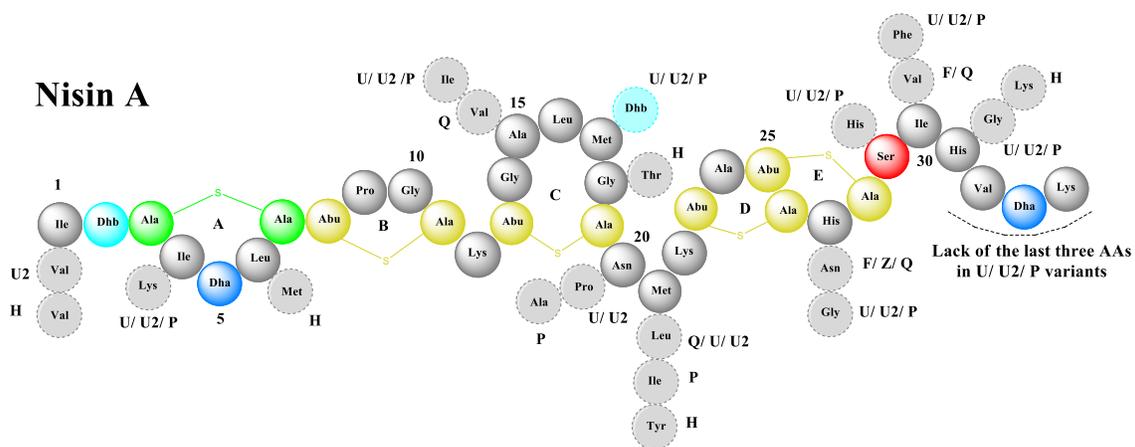


Figure A.1.1: Nisin A and variants. Dashed residues denote the different exchanged positions.

Organism	Variant	Exchanged residues in comparison to Nisin A	Molecular Mass
<i>Lactococcus Lactis</i>	Nisin A		3352 Da
	Nisin Z	His27Asn	3331 Da
	Nisin F	His27Asn/Ile30Val	3315 Da
	Nisin Q	Ala15Val/Met21Leu/His27Asn/Ile30Val/	3327 Da
<i>Streptococcus uberis</i>	Nisin U	Ile4Lys/Ala15Ile/Met17Dhb/Asn20Pro/Met21Leu/His27Gly/Ser29His/Ile30Phe/His31Gly	3029 Da
	Nisin U2	Ile4Lys/Ala15Ile/Met17Dhb/Asn20Pro/Met21Leu/His27Gly/Ser29His/Ile30Phe/His31Gly	3015 Da
<i>Streptococcus gallolyticus</i> subsp. <i>Pasteurianus</i> and <i>Streptococcus suis</i>	Nisin P	Ile4Lys/Ala15Ile/Met17Dhb/Asn20Ala/Met21Ile/His27Gly/Ser29His/Ile30Phe/His31Gly	2989 Da
<i>Streptococcus. Hyointestinalis</i>	Nisin H	Ile1Phe/ Leu6Met/ Gly18Thr/ Met21Tyr/His31Lys	3453 Da

Table A.1.1: Exchanged residue list in comparison with Nisin A.

A.1.1 Physical and Chemical Properties

Numerous studies reported that the biological activity and solubility, is strongly associated to the pH of the media. Nisin A is almost insoluble in neutral or alkaline conditions, it becomes inactive and unstable at high pH, while nisin Z shows improved solubility at neutral pH.^[64,93] It was also suggested that the differences in nisin activity in acid or alkaline conditions must be related to the polymerization of nisin or unfolding in case of a single molecule at higher pHs.^[64] This observation is associated to the fact that the dehydroamino acids can be potentially attacked by nucleophiles such as hydroxide ions, deprotonated amines or deprotonated hydroxyl groups present at high pHs.

A.1.2 Mode of Action

Nisin shows a broad spectrum of activity inhibiting strains of *bacilli*, *clostridia*, *corynebacterial*, *lactobacilli*, *micrococci*, *pediococci*, *streptococci*, and *actinomycetes*.^[64] Additionally, nisin also affects in a lower extent other pathogens such as: *Mycobacterium tuberculosis*, *Staphylococcus- pyogenes*, *- aureus*, *- epidermis* among others.^[64]

At first, the research effort was focused on the identification of key residues for the effective bactericidal Nisin's functionality. The following remarks were defined:

- Nisin was biologically inactive without Dha5.^[94]
- Dha33 was not crucial for antimicrobial activity.^[93]
- Modifications in the C_t region of the peptide showed a remarkable decrease impact in the nisin-membrane binding.^[95]
- Studies with different N_t fragments (N1-12 and N1-20) showed minor affinity for membranes in comparison with the full chain.^[96]
- A and B moieties were responsible for antimicrobial activity.^[64] Moreover, the opening of the first lanthionine ring promoted loss of activity.^[97]

Additionally,

- through the estimation of the nisin molecular mass^[80] was confirmed that this lanthipeptide form stable dimers or tetramers.^[64]
- It was also determined how the nisin, at micromolar concentrations, permeabilizes the membrane inserting itself between the head groups of the negatively charged bacterial lipids membranes. It was proposed that in a specific threshold concentration, the peptides could build a short-lived pore-like structures. The relaxation of those pore-like structures allowed the translocation of the peptides across the membrane.^[98,99] This mechanism was also proposed for other AMPs.^[100]

8. Bacteriocins

- Moreover, early studies suggesting that the nisin inhibited the cell wall synthesis^[101] were supported by Reisinger *et al.*^[102], who confirmed that nisin binds to lipid II (the cell wall synthesis precursor - see section 6.3.2). Additionally, it was also proposed that the pore formation by nisin in lipid membrane surfaces could be upheld by lipid II.^[103]

Consequently, in order to clarify the last two remarks, numerous studies of nisin binding, insertion, orientation and pore formation were performed.^[97] From these studies, two different mechanisms for membrane permeabilization were suggested:

- In a micromolar concentration range, nisin perturbs the lipid membrane in model experiments.
- In a nanomolar concentration range, nisin efficiently kills bacteria.

These observations were reconciled by experiments confirming that nisin has a two-fold mode of action by the use of lipid II as 'docking molecule' to form pores in bacterial membranes. Indeed, it was also verified that the stability of the pore was increased in presence of receptor than in its absence. Additionally, it was also proofed that the lipid II is a constituent of the pore in a stoichiometry four lipid II-eight nisin molecules. In this regard, molecular insights were also provided:^[79]

- NMR studies of the nisin Z peptide with a lipid II variant confirmed that the phosphates moiety of the peptidoglycan subunit are the binding motif used by nisin.^[104] Concisely, five intermolecular hydrogen bonds between the backbone amides of N_T-region of nisin with the pyrophosphates group build a characteristic cage architecture for the binding motif. These results supported the experimental observations concerning the nisin indistinguishable linkages to lipid I and to lipid II, where the difference between both is placed in the GlcNAc moiety.

A.2 Nisin Lantibiotic: Molecular Dynamic Simulations of Nisin Natural Variants produced by *Lactococcus lactis*

Test simulations with the nisin natural variants A, Z and Q produced by the *Lactococcus lactis* were carried out (structural details in previous section A-Nisin Lantibiotic, figure A.1.1/ table A.1.1). Nisin F was excluded due to its high similarity to natural variant Q. The nisin Z structure, (PDB code-1WCO)^[104] was used as template to develop the required starting geometries for the variants A and Q.

The focus of this project was to analyse i) the conformational dynamics of the nisin Z in solution compared to the experimentally well-studied nisin A together with the natural variant Q and ii) critical analysis of the complex nisin Z-lipid II in a biologically relevant environment.

Beforehand, it was required to validate the CHARMM-compatible FF for the dehydroamino acids Dha and Dhb.

A.2.1 Validation of the CHARMM-Compatible FF for Dehydroamino Acids.

With the aim of investigating the structural and dynamic properties of the lantibiotics studied in this thesis (natural variants of Nisin and two-component Lantibiotic Lichenicidin) by means of MD simulations, a CHARMM compatible FF for dehydroamino acids was required. The FF for these AAs was initially described by Thormann and Hofmann^[105]. Concisely, their dehydroamino acids CHARMM-FF was generated using the force field for standard L-amino acids with a slight correction of the torsional potential term for the rotation around ϕ , consisting of the increase of the torsional potential from 0.25 to 0.48. (section 2-MM eq. (2.1.1.1)). With this modification, they could assess reproducibility of the structural ab-initio together with the experimental data for the dehydroalanine peptide. Recently, a new CHARMM-FF for Dha and Dhb was reported by Turpin *et al.*^[106]. Despite that this latest work was developed using a well-established procedure based on ParamChem interface (www.paramchem.org)^[107] together with the fact that the parameters involving the $C_{\alpha}=C_{\beta}$ double bond were reoptimized with the help of quantum mechanical calculations at the HF/6-

31G* and MP2/6-31G* levels, several inconsistencies were found after a close analysis of this parameter set such as:

- for the Dhb, exclusively the E-isomer was evaluated as target molecule, although it is the Z-isomer which is obtained after the posttranslational modifications of threonine.^[54,71]
- essential FFs parameters defining the atom types CG2DC1 and CTD1, among others, were missing or inconsistent.

Along with the uncertainty of this published force field, a new derived parameter set was developed. Herein, Paramchem.org server was used for generating CHARMM-compatible force field parameters for Dha and Dhb by analogy to other molecules, in similarity to Turpin's work^[106] procedure. The target molecules used were built with an acetylated N-terminus and N-methylamide at the C-terminus (figure A.2.1.1) Furthermore, in contrast to Turpin's parameterization procedure, the Z-isomer was chosen as detected so far for all AMPs due to the posttranslational modifications.^[54,71]

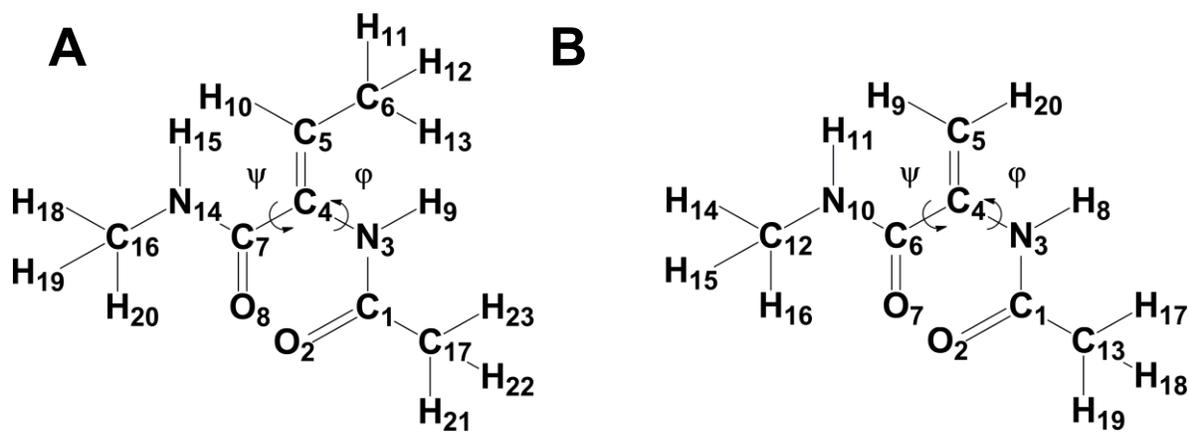


Figure A.2.1.1: Target molecules used for the validation of the CHARMM-compatible FF. A- Target molecule to study Dhb AA. **B-** Target molecule to study Dha AA.

According to the Thormann and Hofmann results, the force constant for the torsional potential around the atom types X-NP-CUA1-X (associated to the C7-C4-N3-C1 dihedral – see figure A.2.1.1) was set to 0.48. The existing force fields parameters for Dha and Dhb were evaluated according to the potential energy curves for partially constrained model compounds computed at a quantum mechanical level followed by the comparison with those obtained at a molecular mechanical level by means of the different sets of force field

parameters (see figure A.2.1.1). For the quantum mechanical calculations, the geometry of the target molecules was initially optimized at the MP3/ HF (3-31G) / HF (6-31G*) level. Subsequently, dihedral scans around the ψ dihedral N14-C7-C4-N3 (N14-C6-C4-N3) and ϕ dihedral C7-C4-N3-C1 (C6-C4-N3-C1) of Dhb (Dha) were performed in steps of 10° (Figure A.2.1.2, blue traces). The QM calculations were carried out using Gaussian 09 program.^[108] Identical molecular starting geometries were employed for the computation at the molecular mechanics level. The MM calculations were done with the NAMD Energy plugin in the VMD package.^[26] The potential energy curves for the ψ dihedral angles of Dha and Dhb are practically identical for both peptides, with clearly defined global minima at about 30° and less pronounced local minima at ca. 140° and 220° , in agreement with the ab initio calculations. Conversely, the ϕ dihedral angle showed relatively flat minima with two energy barriers at about -20° and 170° for Dhb, but only one barrier at about -10° for Dha. Independent of the parameter set used for the force field calculations, the height of the potential energy curves is overestimated compared to the corresponding MP2/6-31G* values. As MD simulations sample low energy states, emphasis was focused on reproducing the energy wells. Force field parameters are provided in the Appendix - Bacteriocins.

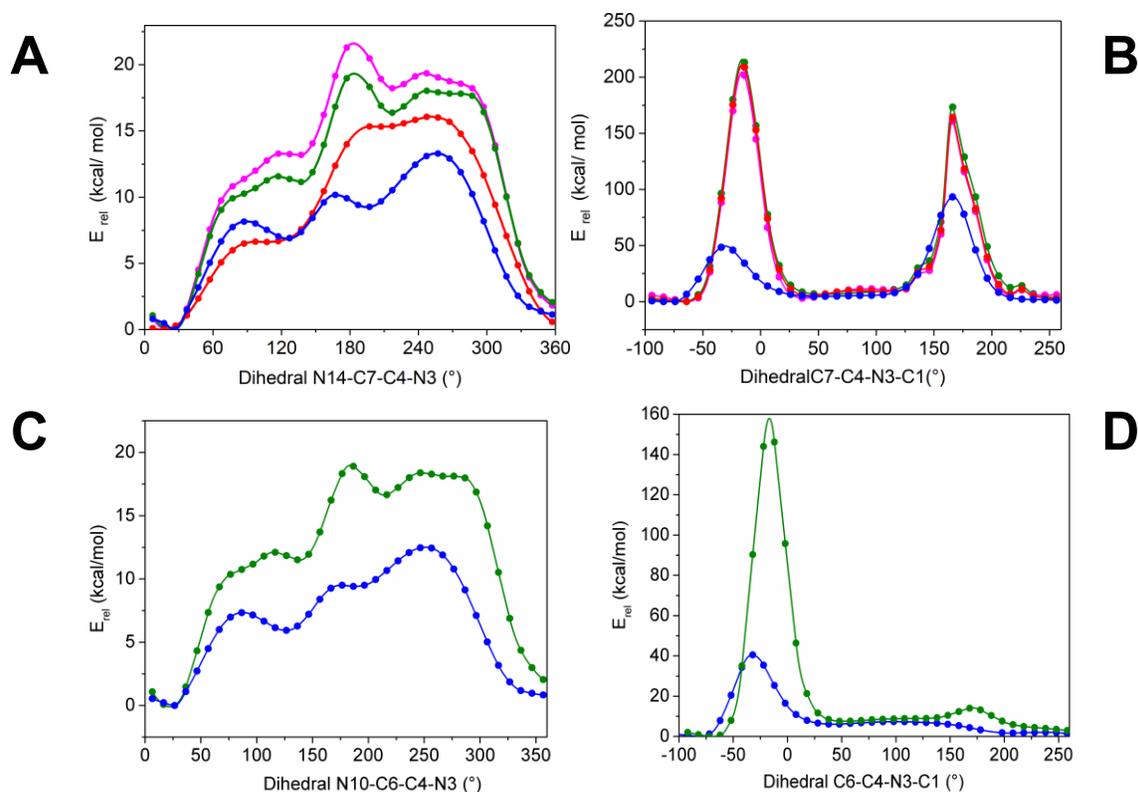


Figure A.2.1.2: Dihedral scan of ψ dihedral N14-C7-C4-N3 and ϕ dihedral C7-C4-N3-C1 of Dhb. **A-B-** Dihedral scan of ψ dihedral N14-C7-C4-N3 and ϕ dihedral C7-C4-N3-C1 of Dhb computed at MP2/6-31G* level (blue trace) and at MM level using force field parameters used for MD simulations (green trace). For comparison, the potential energy curves for ψ and ϕ predicted using exclusively ParamChem parameters (pink trace) and the force field parameters suggested by Turpin *et al.* [106] (red trace) are showed. **C-D-** Dihedral scan of ψ dihedral N14-C7-C4-N3 and ϕ dihedral C7-C4-N3-C1 of Dha computed at MP2/6-31G* level (blue trace) and at MM level using force field parameters used for MD simulations (green trace)

A.2.2 Conformational Dynamics of Nisin Z, A and Q.

After the successful validation of the CHARMM-compatible FF for dehydroaminoacids, the study of the conformational dynamics of the nisin Z and variants A and Q followed. In order to do so, the peptides were individually solvated in cuboid boxes of TIP3P water.^[109] The boxes were generated in such a way to assure a layer of water molecules 15 Å away from the peptides. Accordingly, boxes of 70 Å x 68 Å x 50 Å, 70 Å x 65 Å x 50 Å, 70 x 69 x 50 Å for the Z, A, Q variants respectively, were constructed. Afterwards, the system was ionized,

in agreement to the experimental studies of Rollema *et al.*^[93] at pH 7.0 and employing an ionic strength of 150 mM. For these purposes, the SOLVATE and AUTOIONIZED plugins of VMD program^[26] were employed. The MD simulations were run under periodic boundary conditions employing CHARMM27 for the canonical amino acids and the successfully validated CHARMM-compatible FF for the dehydroaminoacids.^[110–112] Particle Mesh Ewald Summation^[12] was used to compute the extended electrostatics while the Van der Waals (VdW) interactions were evaluated applying a standard cut-off of 12 Å. The SHAKE algorithm^[113] was used to constrain all bond lengths between heavy and hydrogen atoms. Previously, the energies of the three systems were minimized 20000 steps with the conjugated gradient integrator decreasing the harmonic constrain on heavy atoms from 25 to 5 kcal/ (mol Å²). Afterwards, the systems were heated during 100 ps using Langevin dynamics with a timestep of 2 fs and decreasing position restraints on the heavy atoms from 7.5 to 2.5 kcal/ (mol Å²). Subsequently, the equilibration of the peptides during 60 ps followed the heating. Indeed, the harmonic constrains on the peptides were also progressively released during the equilibration run until all atoms were allowed to move freely. Finally, dynamics for these models were simulated with a timestep of 2 fs for 200 ns at 300 K in an NPT ensemble under constant atmospheric pressure and temperature by means of the Langevin Piston method with a damping constant of 0.01 fs⁻¹.^[21] The MD simulations were performed in NAMD 2.9 program.^[16]

The statistical analysis of the structural properties of the different peptides were performed by means of RMSD, RMSF, RG combined with the monitoring of the structural changes in the simulation time (Figures A.2.2.1 and A.2.2.2). In the table A.2.2.1 is displayed the average values of the aforementioned properties. As it was expected, the variants A and Q behave very similarly to nisin Z. From the backbone-RMSD analysis (see A-figure A.2.2.2), a relative stability for A and Q peptides is detected while Z suffers more fluctuations during the simulation. Z and A variants exclusively differ in residue 27. The charged histidine 27 of nisin A is exchanged with a polar AA in nisin Z. Q variant shows in contrast, the same modification in residue 27 and three more exchanged amino acids Ala15Val/ Met21Leu/ Ile30Val. Mainly, due to the exchange of residue 21, the hydrophobic character of this peptide is slightly higher in comparison with the others. Because of that, the RMSD_{avg} obtained for this variant is the highest (table A.2.2.1).

Variant	$\langle \text{RMSD} \rangle / \text{\AA}$	$\langle \text{RMSF} \rangle / \text{\AA}$	$\langle \text{RG} \rangle / \text{\AA}$	$\langle \mu \rangle / \text{D}$
Nisin Z	10.2 ± 2.5	6.4 ± 1.4	11.6 ± 2.2	69.8 ± 30.8
Nisin A	9.4 ± 1.5	5.3 ± 1.5	12.0 ± 1.4	61.4 ± 25.7
Nisin Q	11.5 ± 1.3	5.7 ± 1.4	11.1 ± 0.8	75.0 ± 28.5

Table A.2.2.1: Average values of the RMSD, RMSF, RG and dipolar moment for the nisin Z and variants A, Q.

Moreover, the effect of exchanged AAs can be also observed when monitoring the secondary structures (see figure A.2.2.1). While in the N_t - and C_t -regions the behavior between the variants as well as around the thioether cross-linkages remain quite similar, major differences are found in the hinge region (residues 20 to 22).

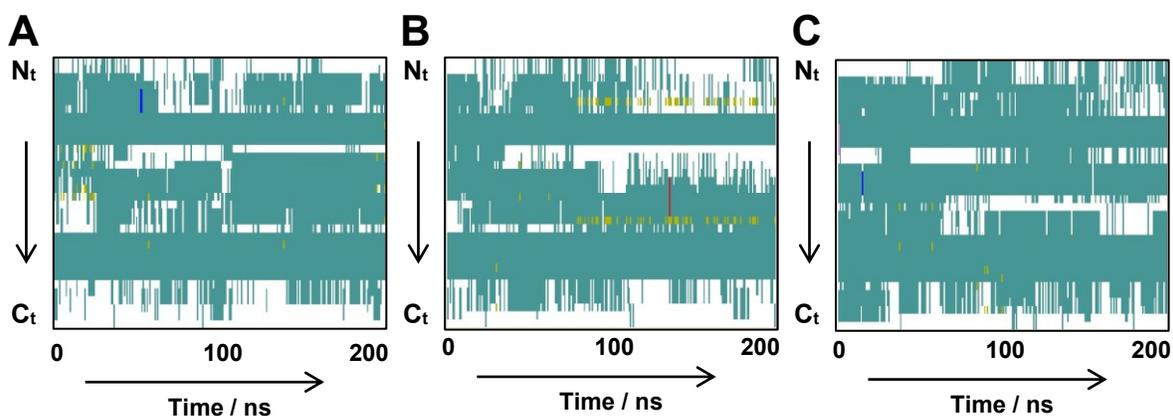


Figure A.2.2.1: Monitoring the secondary structural changes in the course of 200 ns of simulation. A- Nisin Z, B- Nisin A, C- Nisin Q. Color code associated to kind of structure observed in the course of the simulation (pink: α -helix, yellow: β -sheet, teal: turns, white: coil, green: isolated bridges, blue: 3_{10} helix, red: π -helix).

Additionally, as the thioether-cross pattern is identical in all these peptides, the lower fluctuations in these regions of the sequence was expected to be found. This fact was confirmed through the calculation of the C_{α} -RMSF (see B-figure A.2.2.2), ranging average values between 5.3 ± 1.5 to 6.4 ± 1.4 . The lanthionine (ring A: 3-7) and methylanthione (rings B: 8-11, C: 13-19, D: 23-26, E: 25-28) showed lower fluctuations, while the highest-RMSF values were associated to the biggest loop (ring C) and to the hinge region (residues 20 to 22) connecting the rings D and E with the rest of the sequence. This observation is also in agreement with the mechanism of action proposed for nisin: while the N_t -region attaches to lipid II, through the flexible hinge region, the C_t -section could insert itself in the membrane and generate the pore.^[104]

Interestingly, despite the fact that nisin A charge is one unit higher than the rest of the studied variants (see previous section A- Nisin Lantibiotic) at physiological pH, all variants showed the same grade of compactness $\sim 12 \text{ \AA}$ (see C-figure A.2.2.2) keeping the same tendency during the simulation. In opposite, the computed averaged dipolar moments reflected that the exchange of amino acids promoted some differences. Nisin A resulted to be the variant with the lower dipolar moment $\sim 62 \text{ D}$. This fact is related principally to the presence of a charged residue at the 27th position (where the other two variants show a polar non-charged residue) generating a different contribution to the total dipole vector. While the same position in the variants Z and Q share the same chemical character resulting in an equivalent dipolar moment $\sim 70 \text{ D}$ (table A.2.2.1).

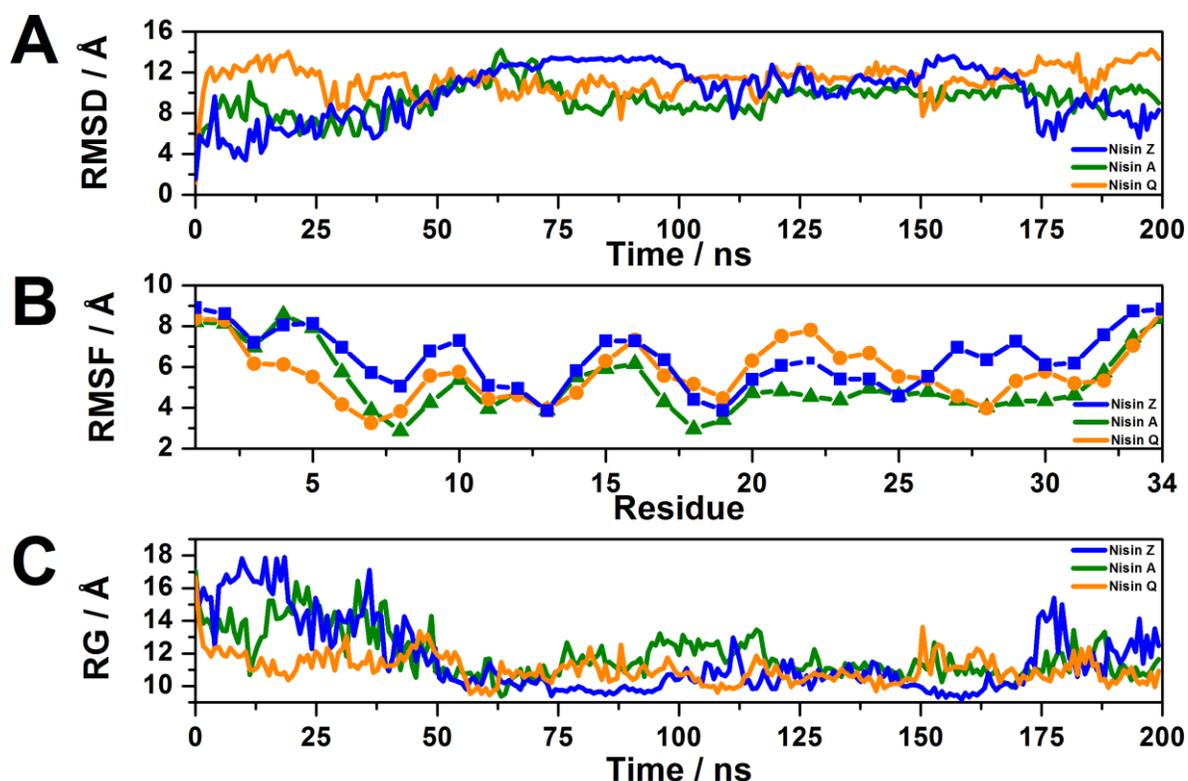


Figure A.2.2.2: Analysis of trajectories by means of A-RMSD, B- RMSF, C- RG. Nisin Z (blue line), nisin A (green line) and nisin Q (orange line).

A.2.3 Complex Nisin Z-Lipid II Immersed in a Membrane Mimicking a Gram-Positive Bacteria

After the evaluation of the conformational dynamics of the nisin Z in solution comparing it with the nisin natural variants A and Q, a critical analysis of the complex nisin Z-lipid II in a biologically relevant environment followed.

The main objective of this section was to evaluate the characteristic cage-architecture between the peptide nisin Z and the peptidoglycan subunit, lipid II. To do so, the stability analysis of the unique H-bonding pattern between both components in the course of the simulation resulted to be the main interest.

Firstly, CHARMM-GUI^[114] was used to develop a membrane, mimicking the surface of a Gram-positive bacteria, involving a stoichiometry 3:1 of POPG and POPE lipids, respectively. An 80 x 80 Å POPG/ POPE bilayer in the x- y plane was obtained. In 2004, the pdb-structure of nisin Z attached to short variant of Lipid II (namely 3LII, due to the presence of exclusively three isopren units) was published (PDB entry-1WCO).^[104] This structure was used and a manual modification followed with the purpose of obtaining the natural eleven isoprene chain. Additionally, the CHARMM-FF for the lipid II component was computed with ParamChem interface (www.paramchem.org)^[107] (CGenFF parameters and topology for lipid II are provided in Appendix – Bacteriocins).

Second, the complex Nisin Z- full lipid II was introduced in the POPG/ POPE (3:1) membrane. Once the complex was inserted in the upper layer of the membrane, 3 POPG and 1 POPE molecules were manually deleted, to keep the lipid ratio.

Third, the system was solvated in a water box size of 80 Å x 80 Å x 130 Å and ionized using the VMD plugins^[26] according with the experimental studies of Rollema *et al.*^[93] The protonation states of ionizable groups were used at neutral pH with an ionic strength of 150 mM. The MD simulation was run under periodic boundary conditions employing CHARMM27 for the canonical amino acids together with the successfully validated CHARMM-compatible FF for the dehydroaminoacids and CGenFF for lipid II component.^[110–112] The long-range electrostatics were computed employing Particle Mesh Ewald Summation^[12] and for the Van der Waals interactions a standard cut-off of 12 Å was applied. All water molecules were treated with the TIP3P model^[109]. To assure a timestep of 2 fs the SHAKE algorithm^[113] was

used to constrain all bond lengths between heavy and hydrogen atoms. The energy was minimized for 20000 steps with the conjugated gradient integrator employing harmonic constrains on heavy atoms from 25 to 5 kcal/ (mol Å²). Afterwards, the system was heated during 100 ps using Langevin dynamics with a timestep of 2 fs and decreasing position restraints on the heavy atoms from 7.5 to 2.5 kcal/ (mol Å²). Subsequently, the equilibration of the system during 80 ps was produced. During the equilibration run, the harmonic constrains on the system were progressively set free. Finally, the free dynamics using a NPT ensemble was run.^[19] The system was simulated employing a time step of 2 fs during time enough to assure the initial mentioned goal, concisely 170 ns at 300 K. Visualization of the system is provided in figure A.2.3.1.

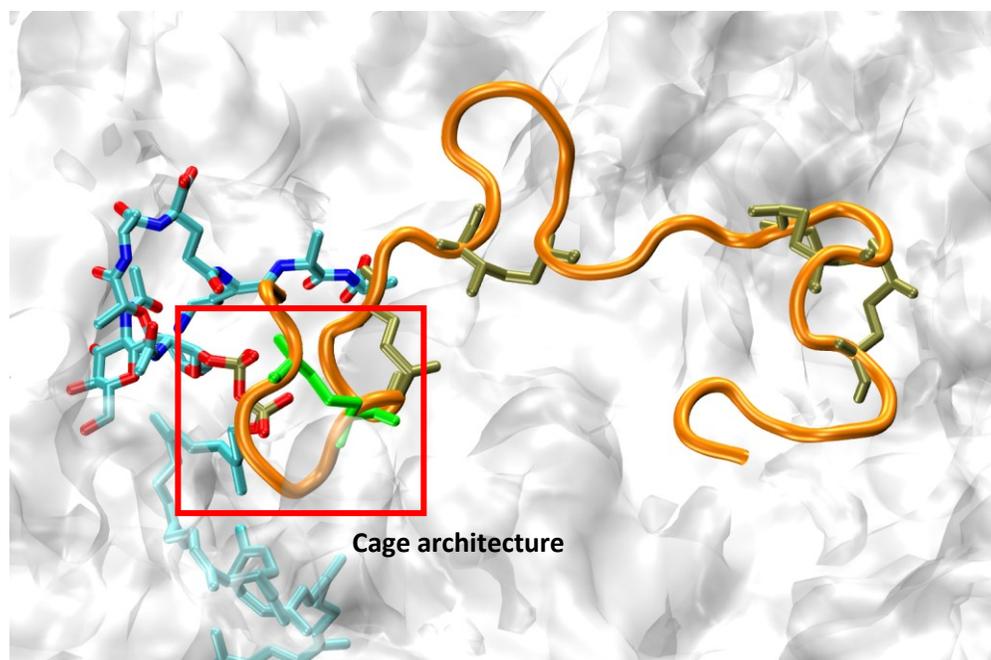


Figure A.2.3.1: Complex Nisin Z-Lipid II over the membrane surface. Water and ions are not shown for clarity.

It was previously reported the unique cage architecture between lanthipeptide nisin and the cell wall building block. Concisely, the cage is specifically composed by five H-bonds involving the ring A (Lan 3-7), Ile 4, Dha 5, Leu6 of the peptide to the phosphate moiety of the lipid II molecule. Herein, in the course of 170 ns of simulation, the stability of H-bonds cage was evaluated. As it was expected, at the beginning of the simulation several interactions between the side chains of the peptide and the membrane were detected, causing some fluctuations in the hydrogen bonding between both components (see figure

A.2.3.2). These fluctuations were diminishing after some nanoseconds obtaining an undisturbed behavior until the end of simulation. Interestingly, it was also noticed that the ring B also affects to the stability of the cage. This fact is also reflected in the variable distance found in the CYS 7.

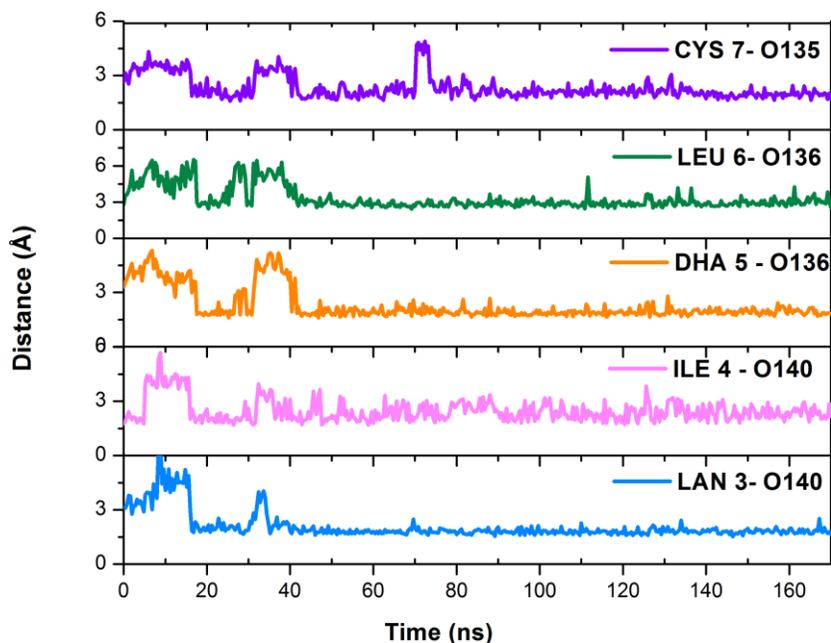


Figure A.2.3.2: Monitoring of the stability of main H-bond peptide-phosphate cage.

Furthermore, the lantibiotic showed a lower backbone-RMSD_{avg} 6.156 ± 0.998 than its analogue in solution, as a consequence of the reduction of degrees of freedom. Therefore, lower deviations with respect to the initial structure are noticed during the trajectory. The RDF of the phosphates of the lipid II component relative to the phosphates of the POPG and POPE lipids during the trajectory were also computed (figure A.2.3.3). RDF-plot shows how the lipid density varies as a function of distance between the phosphates groups and those of lipid II used as reference point. This calculation also contributed to the understanding of the experimental observables regarding the nisin capacity of recruit the negatively charged lipids. This calculation also provides useful information of the proposed pore in the membrane by a complex defined by 8 Nisins-4 lipid II.^[79]

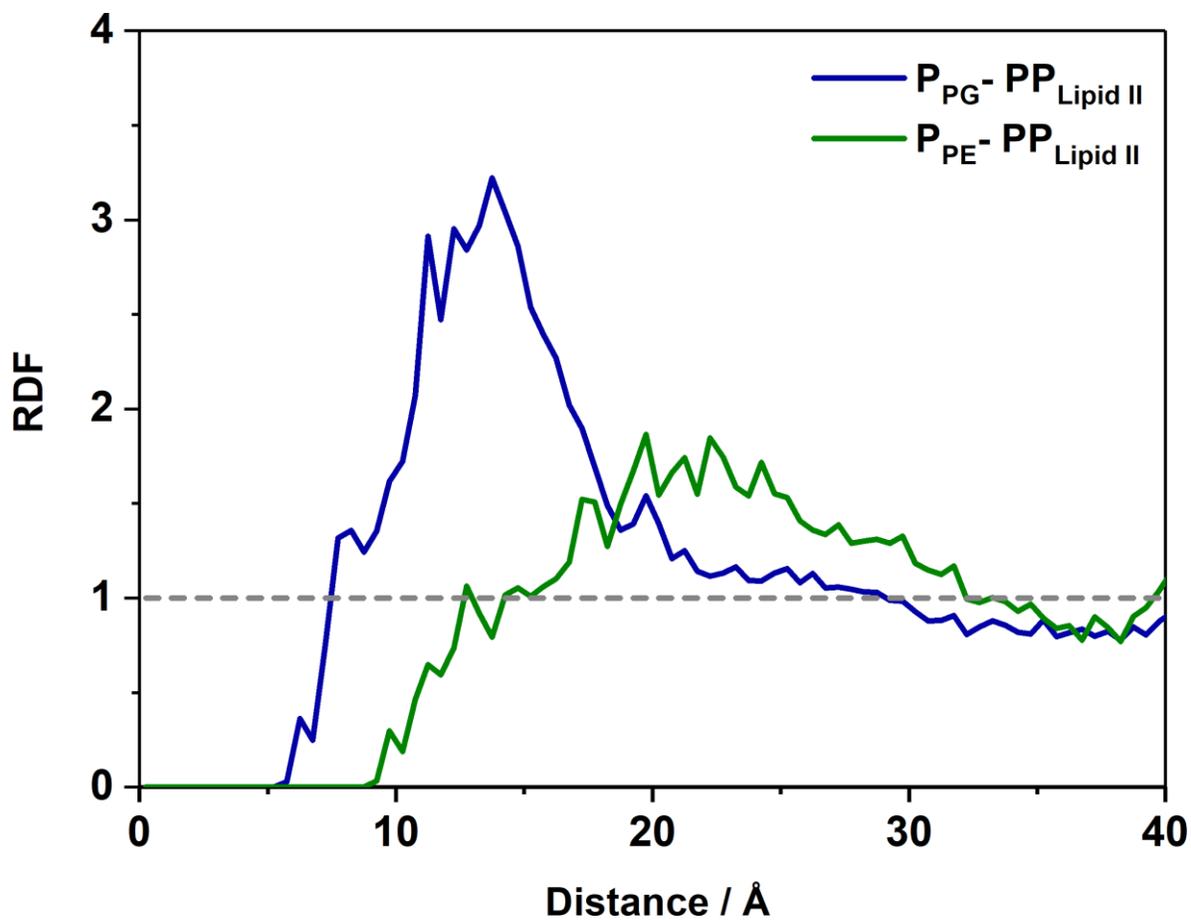


Figure A.2.3.3: RDF for the phosphates group lipid II relative to the phosphates groups of the surrounded lipids. POPG (blue line) and POPE (green line). RDF was computed for the 170 ns of simulation. Dashed line denotes the RDF=1.

A.2.4 Concluding Remarks of the all-atom MD Simulation of Nisin Z, Natural Variants A and Q, and Evaluation of the Cage Architecture of the Complex Nisin Z-lipid II

After the successful validation of the CHARMM-compatible FF for the dehydroaminoacids, conformational analysis of nisin Z by means of all-atom MD simulations was carried out. The study consisted in i) the comparison with the natural variants A and Q, ii) the evaluation of the H-pattern to the lipid II component. From the analysis of the trajectories, it is possible to conclude that:

- i. The **exchange** of few **AAs** in the nisin variants (see figure A.1.1, table A.1.1 in section A- Nisin Lantibiotic) **does not promote significant differences in the dynamics** of all of them. This similarity (reflected with comparable values of RMSD, RMSF, RG, and secondary structural evolution) can also be extrapolated for the understanding of the bactericidal effects of nisin Z, A, and Q which at the experimental level display equivalent activities against a huge variety of tested bacteria.^[83]
- ii. The presence of **Lan** and **Mlan rings** constraining the structures provide **mechanic stability** (see B-figure A.2.2.2). Absence of these structural motifs provide a **remarkable flexibility in** comparison with the full chain around the residues 20-22, known as **hinge region**, affecting the mobility of the surrounding residues. The free movement associated to this region supports its suggested important role for the insertion of the nisin in the surface of the membrane^[95] (see section A- Nisin Lantibiotic, A.1.2).
- iii. In agreement with the results of Koch *et al.*^[115], the **stability of the H-bond cage characterizing nisin Z and lipid II** was detected. Some fluctuations in the H-bonding at the beginning of the sampling were observed, due to the stabilization of the peptide on the membrane surface. Furthermore, through the theoretical approach, it was possible to probe an additional experimental evidence: nisin recruits the PG lipids of the membrane surface (see figure A.2.3.3).

8.1.1.1-B Subclass II Lanthipeptides: Bifunctional Lanthionine Synthetase

A single bifunctional enzyme, LanM is responsible of the dehydration and cyclization reactions to yield the thioether cross-links. Some examples belonging to this class are: Mersacidin^[116], Lacticin 481^[117] and a newly class II lanthipeptides involving two-component lantibiotic. This last group compromises two post-translationally modified peptides that synergistically show strong antibacterial activity. Interestingly, it has been noticed that generally in an individual evaluation a lack or strong decrease of activity is displayed. During the last decades several two-component lantibiotics have been discovered and sequenced: lichenicidin^[118,119], plantaricin W^[120], enterocin W^[121,122], Smb^[123], Bht^[124], staphylococcin C55^[125,126], lacticin 3147^[127], haloduracin^[128], bicereucin^[129], thusin^[130], formicin^[131], and cytolysin^[132–135] which is characterized by a high grade of similarity between both components. Furthermore, it has been recently discovered that the two-component Lantibiotic Flv System^[136] compromises a set of flavecins consisting of a single α -component and 8 different β -peptides. Interestingly, only some of which displayed synergistic antimicrobial activity which is the reason of why in the figure 8.1.1.1.1 only a single β -peptide is displayed (see table 8.1.1.1.1 and figures 8.1.1.1.1-2).

Two-component Lantibiotic	Producer species
Plantaricin W	<i>Lactobacillus plantarum</i>
Enterocin W	<i>Enterococcus faecalis</i> NKR-4-1
Smb	<i>Streptococcus mutans</i> GS5
Bht	<i>Streptococcus rattus</i> BHT
Staphylococcin C55	<i>Staphylococcus aureus</i> C55
Lacticin 3147	<i>Lactococcus lactis</i> DPC3147
Haloduracin	<i>Bacillus halodurans</i> C-125
Lichenicidin	<i>Bacillus licheniformis</i> VK21
Bicereucin	<i>Bacillus cereus</i> SJ ₁
Thusin	<i>Bacillus thuringiensis</i> BGSC
Formicin	<i>Bacillus paralicheniformis</i> APC 1576
Cytolysin	<i>Enterococcus. faecalis</i>
Flavecins	<i>Ruminococcus flavefaciens</i> FD-1

Table 8.1.1.1.1: List of two-component Lantibiotic discovered up-to-date.

8. Bacteriocins

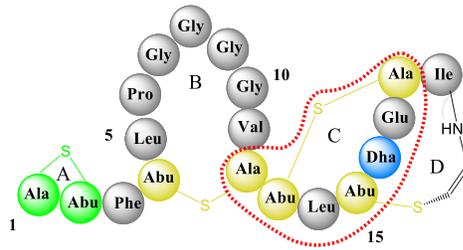
With the exception of cytolysin and bicereucin, which are characterized by lower amounts or none (such as the case of Bsja) lanthionine or methyllanthionine bridges, all other up-to-date discovered two-component lantibiotic can be considered as mersacidin-like peptides. This designation is produced because of the structural analogies of the α -components to the aforementioned lantibiotic, which are highlighted in figure 8.1.1.1.1.

It has been reported that Mersacidin inhibits the peptidoglycan synthesis by targeting lipid II component.^[79,116] Concisely, throughout a mechanism involving its ring C/ D to the phosphate cage, *N*-acetylglucosamine (GlcNAc) and *N*-Acetylmuramic acid (MurNAc) of lipid II. This attachment pathway differs from those proposed for the glycopeptide antibiotic Vancomycin which attaches to DAL6 and DAL7 of the lipid II side chain. As well as, for that proposed for Ramoplanin which attaches to the phosphate cage together to the *N*-Acetylmuramic acid (MurNAc) of lipid II.^[79]

Due to the structural similarities before mentioned regarding the α -peptides and mersacidin, a similar mechanism of interaction is expected for the α -lipid II attachment.

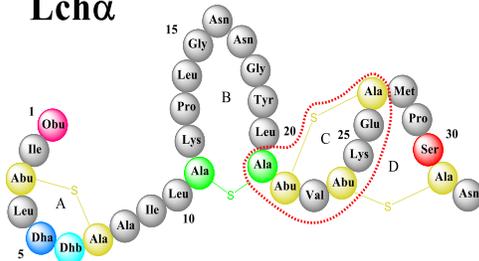
Mersacidin

Mrs

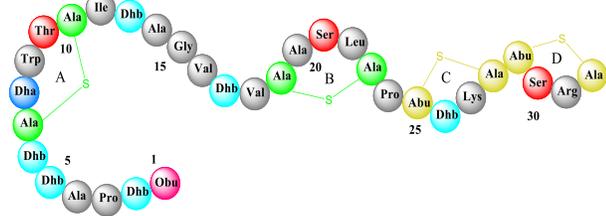


Lichenicidin

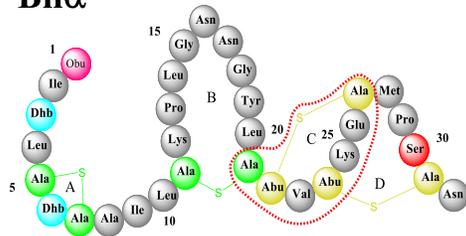
Lch α



Lch β

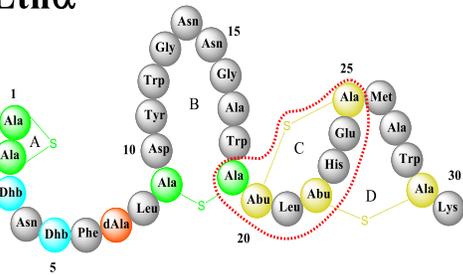


Blic α

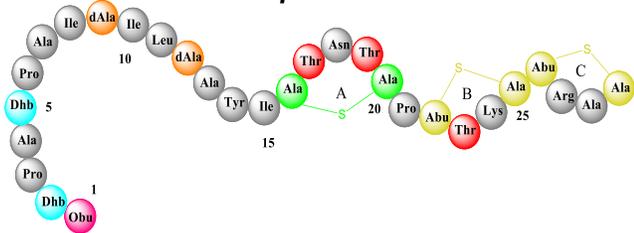


Lacticin 3147

Ltn α

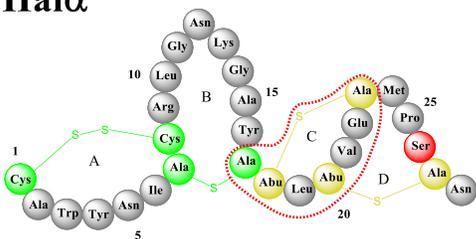


Ltn β

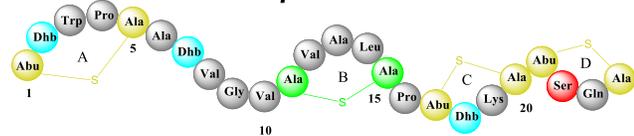


Haloduracin

Hal α



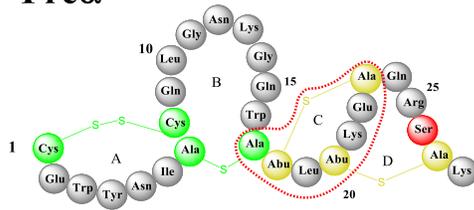
Hal β



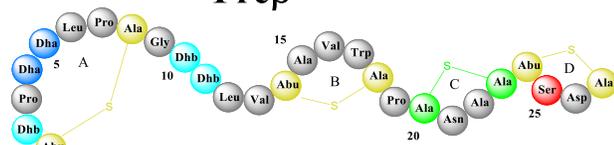
8. Bacteriocins

Formicin

Frc α

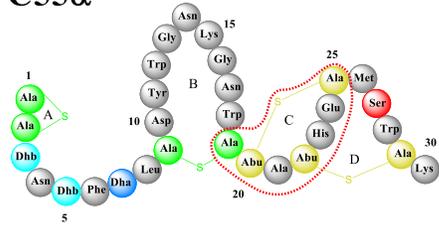


Frc β

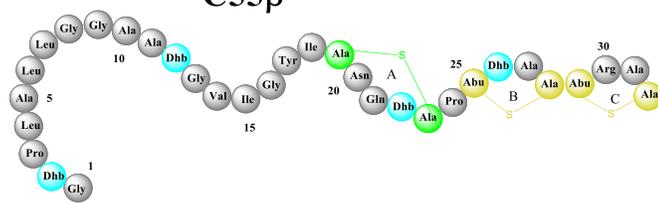


Staphylococcin C55

C55 α

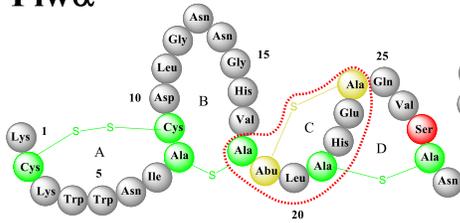


C55 β

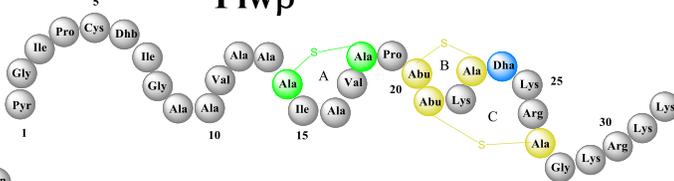


Plantaricin W

Plw α

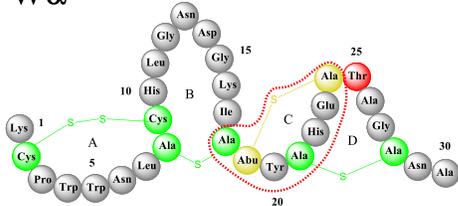


Plw β

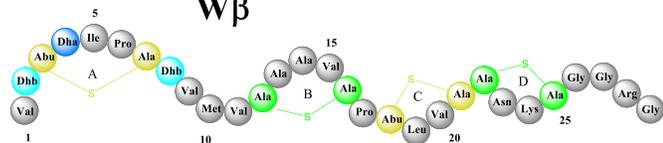


Enterocin W

W α



W β



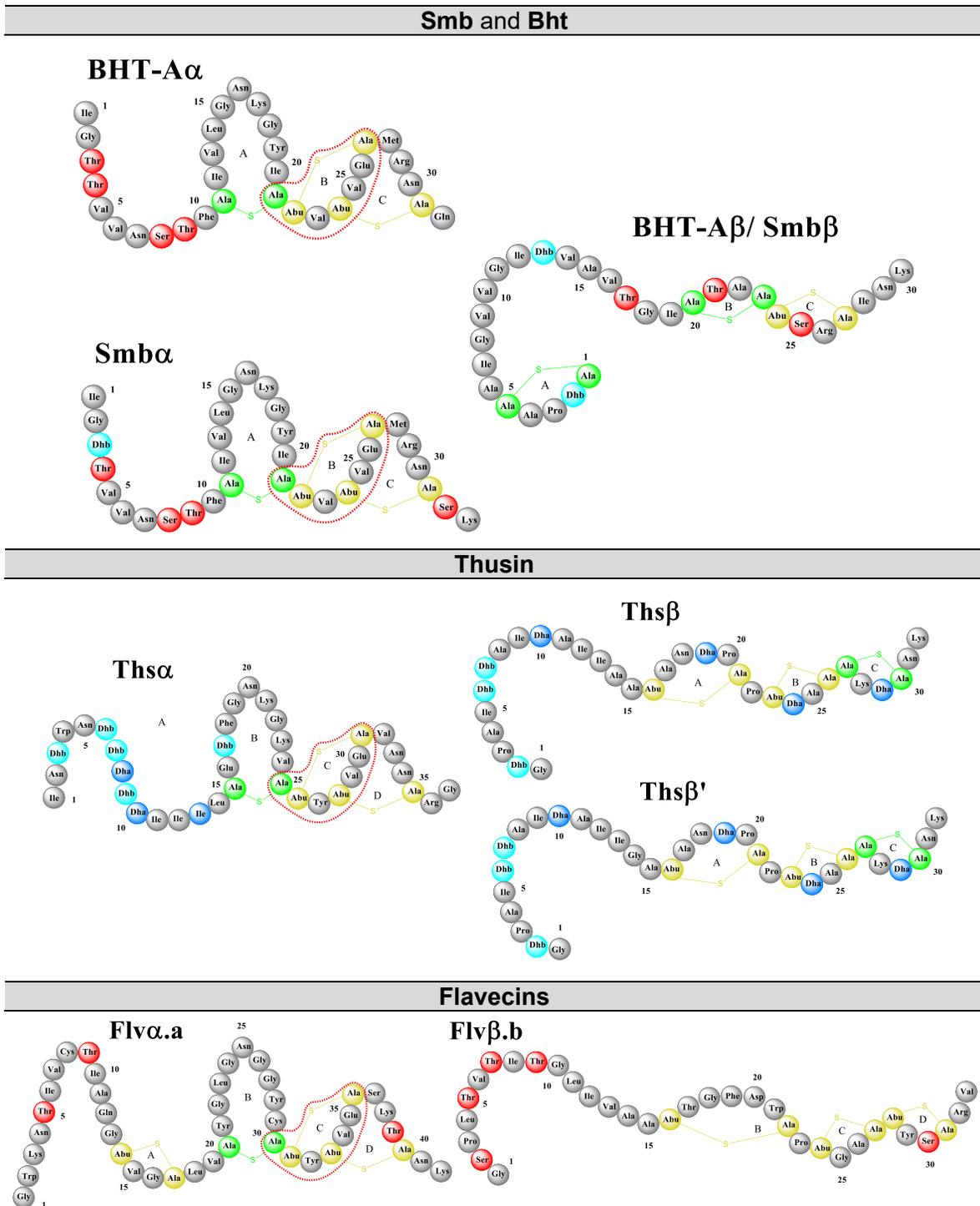


Figure 8.1.1.1: Two-component lantibiotic peptides Mersacidin-like peptides. Two-component lantibiotic Lichenicidin, Lacticin, Haloduracin, Formicin, Staphylococcin C55, Plantaricin W, enterocin W, Smb and Bht, thusin and flavecin (the latter one only a single β -peptide is displayed) belongs to the group denote as Mersacidin-like peptides due to the structural analogies found with Mersacidin.

8. Bacteriocins

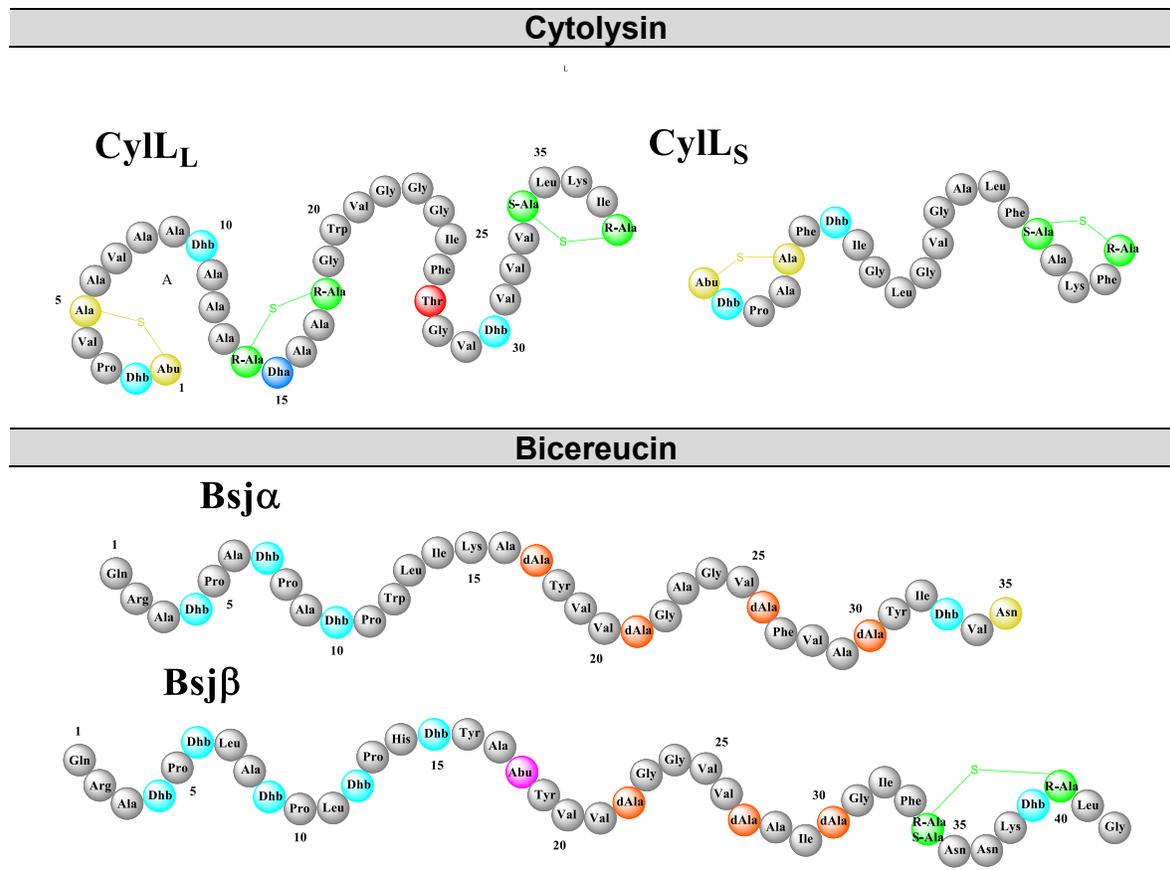
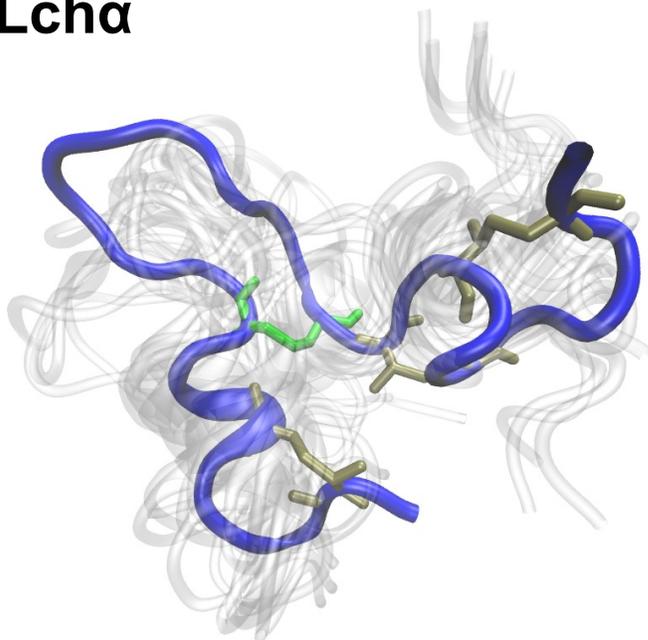


Figure 8.1.1.2: Two-component lantibiotic peptides. Cytolysin and Bicereucin do not keep structural analogies with Mersacidin.

B- Two-component Lantibiotic Lichenicidin

Lichenicidin Lantibiotic is a component of the not so extensive studied two-component lantibiotic family (Section 8.1.1.1-B) produced by the bacteria *Bacillus Licheniformis*. As it can be found in the NCBI database^[137], several strains of *B. Licheniformis* have been revealed as Lichenicidin producers, such as the isogenic strains ATCC 14580 and DSM 13^[138] to the VK21 strain^[118]. So far, only for the mature lantibiotic peptides generated by *B. Licheniformis* VK21, named as Lch α and Lch β , the 3D-structures are available. Concisely, the corresponding structures were solved by NMR spectroscopy in methanol solution (pdb-entries: 2KTN and 2KTO)^[118]. According to these structures, Lch α consists of a flexible loop connecting the N- and C-terminal domains which is stabilized by a thioether bridge between residues 11 and 21. Conversely, Lch β folds into an α -helix with highly mobile N- and C-terminal domains (figure B.1).

Lch α



Lch β

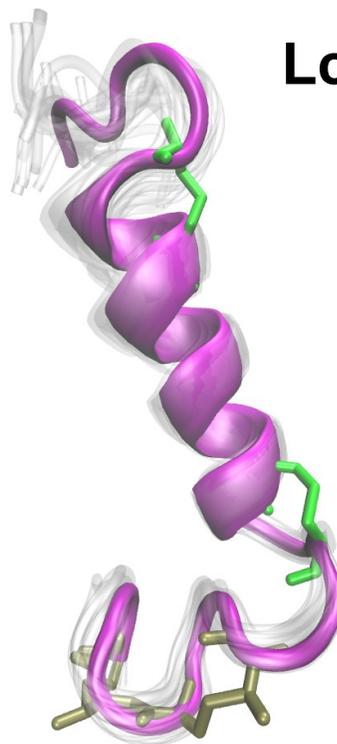


Figure B.1: 3D-structures of Lch α and Lch β peptides. The Lan and Mlan ring are displayed in green and olive color, respectively.

8. Bacteriocins

Interestingly, the considerable signal doubling observed in the NMR spectra is a strong indication for conformational exchange in the millisecond time range. This is particularly true for the Lch α , for which a RMSD of the atomic positions of 6.45 ± 1.79 Å has been estimated.^[118] Moreover, since the NMR experiments were performed in methanol which is known to enhance secondary structure formation^[139] the structural properties of these peptides in aqueous solution remain unsolved.

Furthermore, a following study discovered that the *B. Licheniformis* I89 was also a Lichenicidin producer.^[140] These mature peptides were named Bli α and Bli β . On basis of MS/MS spectroscopy,^[119] the amino acid sequence for the α -peptide was proposed to be different than that detected in the earlier cases. The sole difference between the primary structures of Lch α and Bli α lies on the description of the A-ring, where the methylanthionine bridge between residues 3 to 7 of Lch α is exchanged by a lanthionine bridge compromising the residues 5 and 7 in Bli α (figure B.2). Nevertheless, no differences in the amino sequence between Bli β and Lch β were found. Currently, the 3D-structure of the Bli α peptide has not been solved.

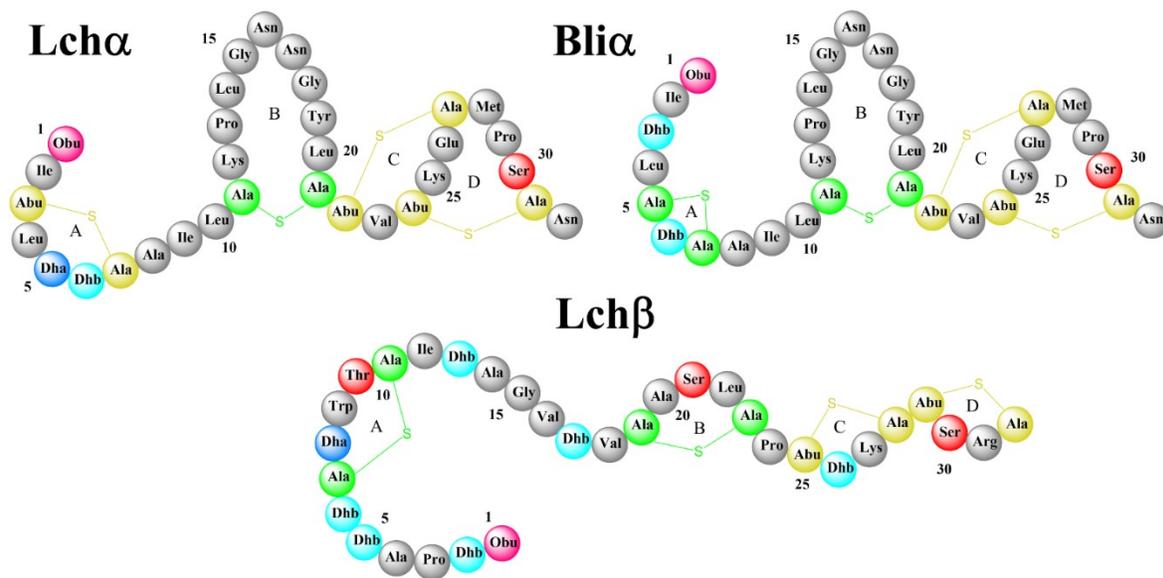


Figure B.2: Amino acid sequence of two-component lichenicidin produced by *B. Licheniformis* strains VK21 and I89. While the α -peptides differ in the A-ring, β -peptides are identical.

B.1 Conformational Dynamics of Two-Component Lantibiotic Lichenicidin in Solution

B.1.1 Model Building

From the Protein Data Bank^[15] were extracted the starting configurations for Lch α and Lch β (figure B.1) (PDB-entries 2KTN and 2KTO, respectively).^[118] Both structures were characterized from NMR measurements. The atomic coordinates of the representative conformer were chosen for both peptides. In contrast, in silico transformations maintaining the correct stereochemistry using the software Gaussview^[141] were necessary for developing the structure of Bli α using the atomic coordinates of Lch α as template:

- Aminobutyrate (Abu3) and the dehydroalanine (Dha5) were converted into dehydrobutyrate (Dhb3) and alanine (Ala5) residues, respectively.
- Ala5 was connected to Ala7 via a thioether bridge and further transformed into a lanthionine residue (residues 5 and 7).

Lch β comprises a short peptide chain of 31 amino acids with two lanthionine bridges placed between the residues 7-11 and 19-23 closing rings A and B, respectively and two methyllanthionine bridges between the residues 25-28 and 29-32, giving rise to the so-called rings C and D, respectively. Furthermore, the peptide is characterized by an α -helix enclosing the positions 9 to 18 together with a N-terminal 2-oxobutyryl group. Lch β showed slightly positively charged, (+1). In agreement with the β -partner, Lch α also constitutes a short peptide of 31 amino acids, which is characterized by the presence of a single lanthionine bridge between residues 11-21 (closing ring B) and three methyllanthionine bridges between the residues: 3-7 (ring A), 22-27 (ring C) and 24-31 (ring D). In Bli α , in contrast, the lanthionine bridge on ring A is found between the residue 5 and 7. Beyond residue Dhb6, both α -peptides exhibit the same sequence of amino acids. In similarity to β -peptide, both α -components harbour a 2-oxobutyryl group at the N-termini. Although most lantibiotics and antimicrobial peptides are positively charged, Lch α and Bli α exhibit a net charge of zero. The same electric property has been previously reported for the α - component of lactacin or mersacidin.^[127]

B.1.2 All-Atom MD Simulation 2 μ s-long for Lichenicidin Lantibiotic

All-atom MD simulations were carried out in order to investigate the structural flexibility of the α - and β -components of Lichenicidin produced by *B. Licheniformis* VK21 (Lch α , Lch β) together with those by *B. Licheniformis* I89 (Bli α , Bli β). As the exclusive difference between these pairs of lantibiotics is placed in the ring A of the α -component, the MD simulations were delimited to Lch α , Lch β and Bli α . In order to do so, according to the experimental studies of Mendo *et al.*^[140] (at pH 7.0 and with an ionic strength of 160 mM of NaCl) the three polypeptides were individually solvated in cuboid boxes of water molecules and ionized through the use of the SOLVATE and AUTOIONIZE plugins of VMD^[26]. Concisely, the simulation systems for Bli α and Lch α were solvated in water boxes (52 Å x 52 Å x 50 Å) contained about 4800 water molecules, 7 Na⁺ and 7 Cl⁻ ions, while for Lch β , due to its elongated shape, a larger water box (52 Å x 56 Å x 55 Å) harboring 6622 water molecules, 9 Na⁺ and 10 Cl⁻ ions was required.

The MD simulations were run with NAMD 2.9 program^[16] under periodic boundary conditions. For Van der Waals (VdW) interactions and real space electrostatics a cut-off of 12 Å was applied, while long-range electrostatics was calculated with the Particle Mesh Ewald Summation.^[12] The canonical amino acids were described with the CHARMM27^[110-112] force field while the dehydroamino acids Dha and Dhb were modeled using the CHARMM-compatible force field described in this work (section A- Nisin Lantibiotic). All water molecules were treated with the TIP3P model.^[109] To satisfy a time step of 2 fs, the SHAKE algorithm^[113] was used to constrain all bond lengths between heavy and hydrogen atoms. According to the MD protocol explained in 3.7 Simulations Protocol, the energies of the three systems were minimized for 20000 steps with the conjugated gradient integrator stepwise decreasing harmonic constrains on all heavy atoms (from 25 to 5 kcal/(mol Å²)). Posteriorly, the systems were heated during 25 ps using Langevin dynamics with a time step of 0.5 fs and decreasing position restraints on the heavy atoms from 5 to 2.5 kcal/(mol Å²). After heating, the systems were equilibrated for 60 ps. The harmonic constrains on the peptides were progressively released during the equilibration run until all atoms were allowed to move freely. Subsequently, the dynamics for the three models were simulated with a time step of 2 fs for 2 μ s at T=300 K in an NPT ensemble under constant atmospheric pressure and temperature using the Langevin Piston method with a damping constant of

0.01 fs^{-1} .^[19] In all production runs, the energies were evaluated at intervals of 0.5 ps while coordinate trajectories at intervals of 2 ps.

The trajectories were analyzed by monitoring the secondary structure evolution with STRIDE program^[25] (section 3.7.3.1.5) implemented in VMD package,^[26] the RMSD change of the peptide backbone in comparison with the initial structure and RMSF. The dynamical behavior of the three peptides displayed in the monitoring of the secondary structure predicted by the MD simulations, differ from each other (figure B.1.2.1). The three peptides showed fundamentally flexible turns and random coil conformations (white and turquoise regions in figure B.1.2.1) and have the tendency to form sporadically β -sheet structures (yellow lines in figure B.1.2.1) and have the tendency to form sporadically β -sheet structures (yellow lines in figure B.1.2.1). Interestingly, only Lch α and Lch β possess the ability to adopt, in addition, α -helical structures in the nanosecond time scale, as indicated by the pink regions in figure B.1.2.1. In contrast, this evidence is not predicted for Bl α for which a single short lasting 3_{10} -helix resulting from sporadic hydrogen bonded interaction between Abu24 and Glu26 was formed.

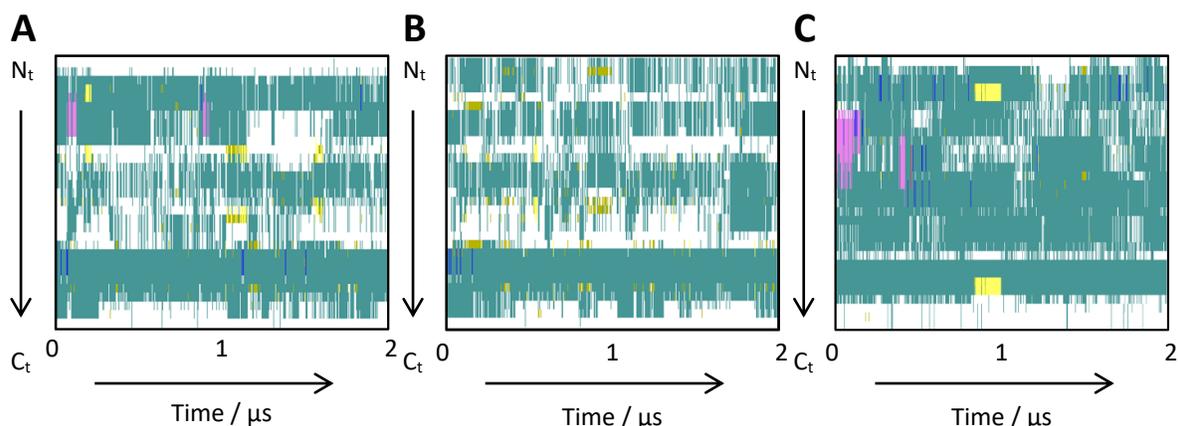


Figure B.1.2.1: Monitoring of the secondary structure during 2 μs of simulation. A- Lch α peptide. **B-** Bl α peptide. **C-** Lch β peptide. Color code for the designation of the possible secondary structures: teal (turns), yellow (extended configurations), green (isolated bridges), pink (α -helix), blue (3_{10} helix), red (π -helix), white (coil).

The higher rigidity at the N-terminal region of Lch α compared to Bl α is most likely a consequence of the methyllanthionine bridge connecting residues 3 and 7 present in the former and absent in the latter peptide. Indeed, the RMSF of the α -carbons averaged over the entire peptide of Bl α is $5.16 \pm 1.51 \text{ \AA}$, are slightly higher compared to the RMSF predicted

for Lch α of only 4.77 ± 1.32 Å. In the case of Lch β , the lanthionine bridge between residues 7 and 11, serve as stabilizing structural element.

For the three peptides, from the analysis of the RMSD (see figure B.1.2.2), large conformational changes of the peptides backbone were obtained in comparison with the initial conformations, ranging values between 4 to 10 Å for the two α -components and from 2 to 13 Å for the β -peptide. In agreement with the experimental evidences, where a RMSD average value of 6.45 ± 1.79 Å estimated via NMR spectroscopy^[118] was depicted, the computed average RMSD value of about 6.7 ± 1.0 Å was obtained. However, a different scenario was found for the Lch β peptide. The average RMSD value of about 9.0 ± 1.7 Å computed for the β -component differ from the experimental value of 3.1 ± 1.0 Å^[118]. The difference between the computed and experimental RMSD values most likely originate from solvent effects. It is already well-known that methanol, the solvent used in the NMR experiments, stabilizes secondary structures.^[139]

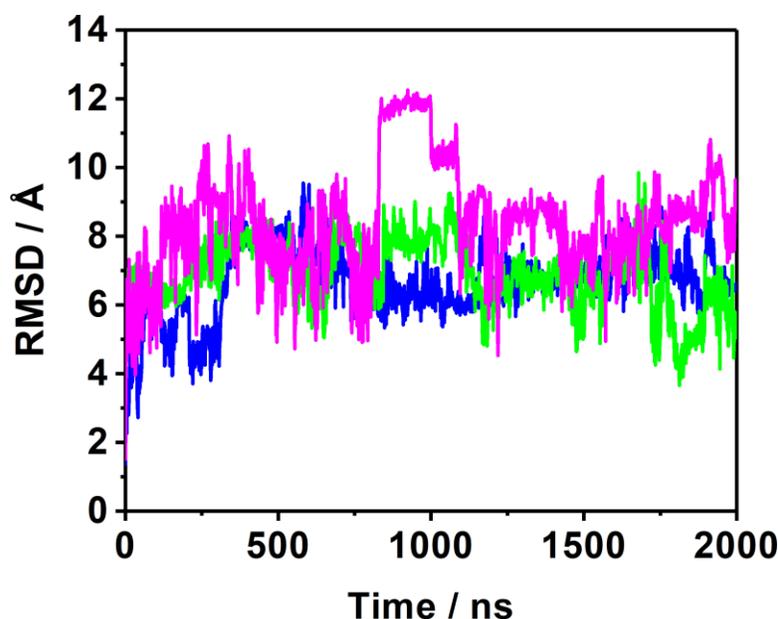


Figure B.1.2.2: Root Mean Square Deviation of Lichenicidin. RMSD computed with respect to initial configuration of the backbone atoms for Lch α (blue color); Bli α (green color); Lch β (magenta) peptides is plotted. Terminal residues were excluded for the evaluation.

The high flexibility exhibited by the peptides in solution, exemplified via the complex RMSD traces obtained, difficult the identification of relevant kinetic states. Therefore, MSM were developed for the three peptides in order to achieve an essential description characterizing the kinetics properties and the native dynamical behaviour of this lantibiotic.

B.1.3 MSMs for Lichenicidin Lantibiotic

This work was carried out in collaboration with Dr. Rafael Tapia-Rojo from Columbia University. The equilibrium ensemble of each peptide was described by building the corresponding MSM from the MD equilibrium trajectories.^[27,142] The MSM was constructed according to protocol explained in section 4- MSMs:

- calculating the microstate network (Conformational Markov network^[30]) - performing a geometric discretization of the state space of the system-
- lumping these microstates into kinetically significant clusters, from which physical insight could be obtained.

Time-lagged Independent Analysis (TICA)^[29] was performed on the C_α of the MD trajectories, defining the state space as the three first Independent Components (TICs), providing a kinetically meaningful dimension. Each TIC was discretized into 30 bins of equal volume. The microstate network was built from the MD trajectories, counting the occupation of each state π_i and calculating the transition matrix T_{ij} , which measures the probability of going from state i to state j within time τ (here $\tau = 1$ ns). For the three analyzed peptides, the obtained transition matrix was ergodic and fulfilled detailed balance (microscopic reversibility). Lumping of the microstates was followed with the aim of providing a coarse-grained description of the free energy landscape of the system. Therefore, Stochastic Steepest Descent algorithm^[30] was applied obtaining at the end a coarse-grained representation with a new transition matrix.

Characterization of the MSMs of each peptide was performed by means of i) its population (π_i) or weight of each state, this property can be understood as the “depth” of the associated free energy well (in the classical surface representation of free energy landscapes) as well as how probably the state can be visited ii) the average RMSD of all C_α employing the initial conformation as reference iii) the average escape time t_e iv) the free energy differences (ΔG_{ij}) between state i and the most occupied state v) the corresponding entropy (S_i) vi) the strength of average dipole moment (μ). These values are listed in table B.1.3.1. Additionally, the graphic representations of the transition matrices (T_{ij}) are given for each peptide in the Appendix – Bacteriocins, in order to provide a clear picture of the connectivity network and transition probabilities. Furthermore, for the three peptides the

Chapman-Kolmogorov test was also computed in order to validate the quality of the MSMs^[142]. Concisely, in Appendix - Bacteriocins the implied timescales are displayed (or relaxation timescales), plotted for lag times of 1, 2, 3, 5 and 10 ns. Undoubtedly, all timescales levelled, which proved the validity of the used MSMs with a lag time of 1 ns.

B.1.3.1 MSMs for Lch α Peptide

The initial microstate network consisted of 3144 nodes connected through 8452 links. After applying the SSD algorithm, 35 macrostates interconnected through 345 links defined the network. The microstate network was refined by eliminating nodes with an occupation $\pi_i < 10^{-4}$ avoiding pathological as well as extremely infrequent states. The final network description for the Lch α peptide was composed by 14 configurations connected through 54 allowed transitions including auto-links (figure B.1.3.1.1).

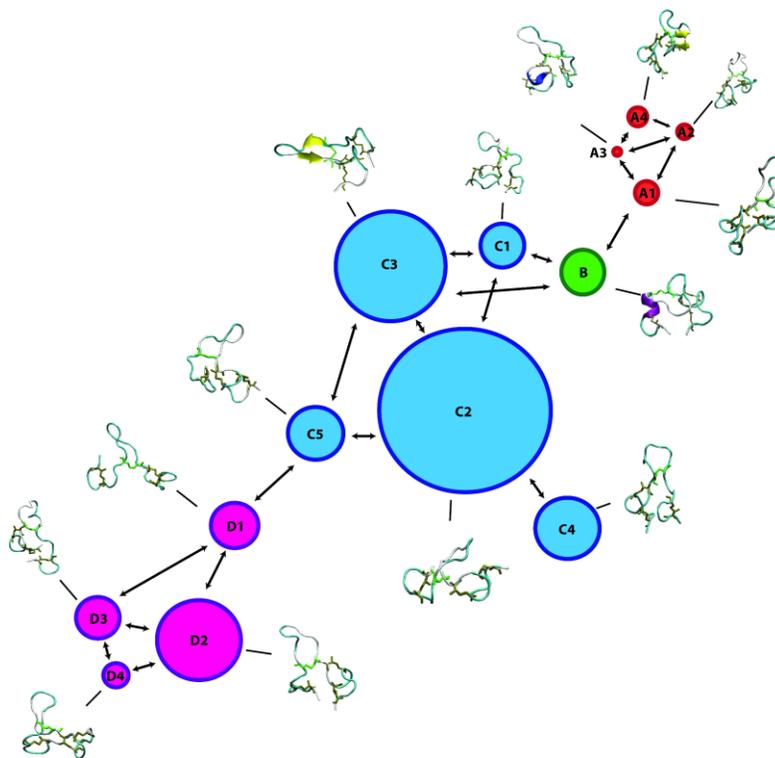


Figure B.1.3.1.1: MSM of Lch α peptide. The 14 states are depicted as beads with size proportional to their occupation, arrows represent allowed transitions between states according to the corresponding transition matrix in Appendix - Bacteriocins.

The network was characterized by three well differentiated regions with different kinetic properties categorized as A, C and D, with B playing the role of a transition state between regions A and C. Clustering of the MSM in four communities was performed with a modularity algorithm.^[143] States in A were denoted to the first relaxation stage, herein the peptide suffers different fast conformational changes before entering in the C region. Low occupations together with low RMSD values in comparison to the initial configuration, see table B.1.3.1 and Figure B.1.3.1.1 mainly characterize this region. A-states defined exclusively the 10% of the network population. Additionally, the state A3 was characterized by a 3_{10} -helix structure involving ring C (residues 22 to 27), while states A1 and A2 were classified as open and closed random-coil conformations. Contrary, state A4 was defined by an antiparallel β -sheet involving the residues 4-7 and 10-14. Since relatively high values of the transition matrix (see transition matrix in Appendix - Bacteriocins) were obtained, low energy barriers interconnecting these four communities were assumed.

However, within this subgroup of states the largest fluctuations were found around the residues defining the lanthionine central loop (ring B -residues 12 and 20), as it can be displayed in the figure B.1.3.1.2 with RMSF values of up to 5 Å. Contrary, due to the presence of the methyllanthionine bridges, the rings A, C and D were stabilized and consequently lower RMSF values of ca. 2 Å were obtained.

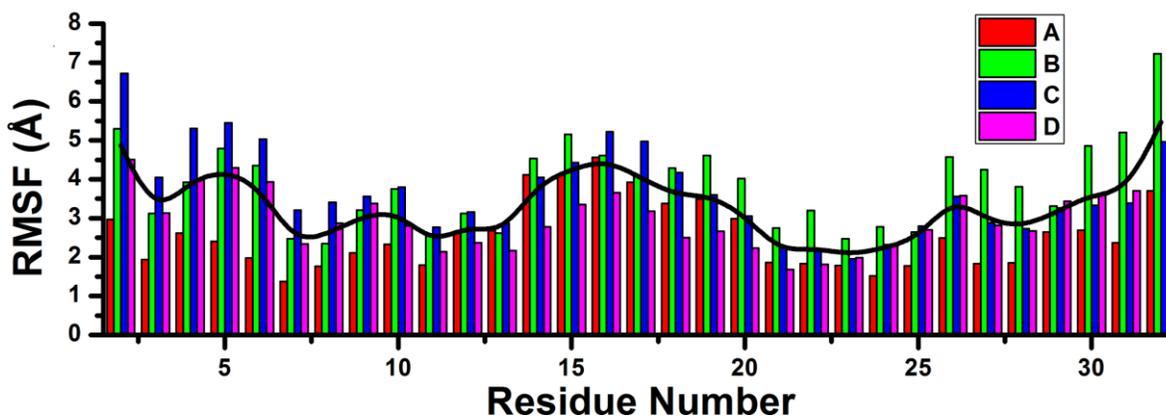


Figure B.1.3.1.2: Average C_{α} -RMSF value per residue per state for Lch α peptide.

States in C was sub-classified in five macrostates. Altogether represented the main region of the equilibrium ensemble characterized by a 60% of the population. Herein, the peptide adopts principally a coil structure, important evidence because this type of structure is usually found for most AMP in water solution.^[54] Furthermore, exclusively the state C3

showed an antiparrallel β -sheet conformation involving the central B-loop, concisely the residues 12-13 and 19-20. Moreover, as it was observed by the non-negligible transition matrix elements (see transition matrix in Appendix - Bacteriocins), confirmation of the strong connectivity between the states in C were found and consequently relatively low free energy barriers. In this network, C2 is predicted with lowest the free energy and highest entropy, therefore constitutes the main basin connecting all other states together. Additionally, all C-states are characterized by a highly mobile N-terminal region, this fact is directly reflected in RMSF-values between 4 – 7 Å for residues 1 to 7 which differ from state A.

The key role in the peptide kinetic was represented by State B. Principally, because this state established a bridge interconnecting A and C communities, as it was clarified through the non-negligible transition probability value (A1, B), (C1, B) and (C3, B) in Appendix - Bacteriocins. Indeed, elimination of B from the network would provoke the disconnection between A and C states. Interestingly, B state was characterized by an α -helical structure covering residues 6 to 11 and representing exclusively the 5% of the network. Moreover, the C-terminal region of the state B was found to be very flexible, according to that RMSF-values arising 7 Å were found.

Concluding with the states in D, which represented the second largest equilibrium ensemble with 25% of the network population. Herein, four macrostates, characterized by a random coil structures, could be differentiated on the basis of their kinetic properties. Comparatively to the random coil structures distinguished in A, B, and C states, the coil structures defining state D resulted to be more rigid, concisely in the central loop (ring B region) where the RMSF-values scarcely reach 4 Å. Additionally, the four D-states resulted to be very interconnected, as it is displayed in the table 1 due to the low time escapes values, finding D1 as the bridge to C-states.

B.1.3.2 MSMs for Bli α Peptide

For this peptide, the starting microstate network was defined by 1424 nodes connected through 3560 links. Subsequently, following the protocol deeply explained in chapter I, section 4. Markov Models MSM, it was reduced to 38 nodes connected through 95 transitions. After the refinement stage, the final network for Bli α peptide consisted of 9 configurations interconnected through 48 transitions (figure B.1.3.2.1).

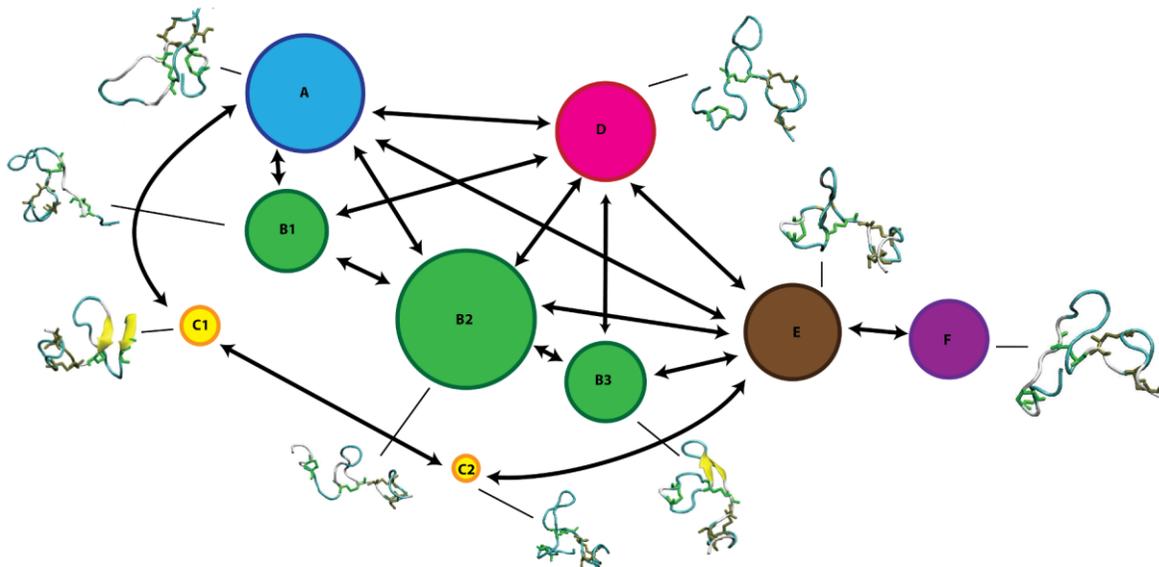


Figure B.1.3.2.1: MSM of Bli α peptide. The 9 states are depicted as beads with size proportional to their occupation, arrows represent allowed transitions between states according to the corresponding transition matrix in Appendix - Bacteriocins.

Six communities defined the Bli α network. Furthermore, the general structure of the equilibrium ensemble of this peptide differ significantly from those characterizing Lch α component. Short lived, highly connected and well-populated states described the free energy landscape (figure B.3.2.1 and Appendix - Bacteriocins).

State A, characterized by a coil structure, constituted the first stage of the simulation, which resulted to be visited several times afterwards, due to the non-zero transition probabilities to states B1, D, C1 and E (see transition matrix in Appendix - Bacteriocins). This initial state comprises only 15% of the network population. Low free energy barrier and consequently, fast transitions to the neighbouring states are attributable to the predicted low escape times. Moreover, this state resulted very flexible according to the RMSF-values ranging 6 – 7 Å.

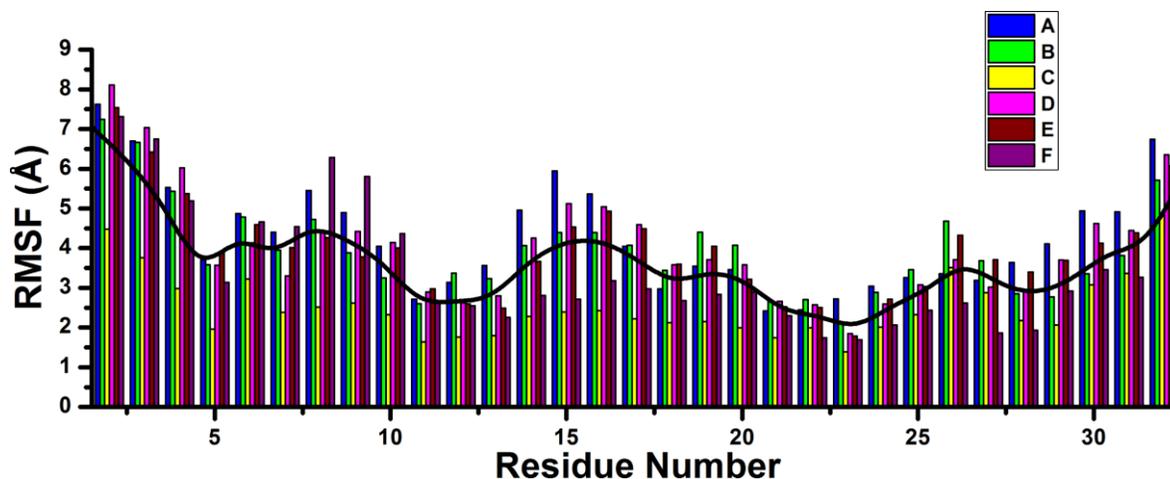


Figure B.1.3.2.2: Average C_{α} -RMSF value per residue per state for Bli α peptide.

In contrast to state A, state B community was constituted by three macrostates namely B1, B2 and B3. Altogether contributed to 37% of the network population. A random coil structure was found for the macrostates B1 and B2, while a short β -sheet was detected involving the residues 11-13 and 18-20 (ring B) in the macrostate B3 in analogy to the structure found in state C3 of Lch α . Interestingly, B community was characterized by an elongated N-terminus region attributable to the exchange of the methylanthionine bridge connecting residues 3 and 7 (found in Lch α) by a shorter lanthionine between positions 5 and 7. Within the B-states, B2 constituted the central state in the kinetics due to the lowest free energy predicted. In line with the C2 state of Lch α , the B2 state constitutes the basin from which transitions to the A, B1, B3, D and E states were allowed. Additionally, this community resulted very connected, as it could be distinguished as a result of the relatively low escape time (below 8 ns).

C1 and C2 macrostates constituted the peripheral communities. Both macrostates played a remarkable role in the kinetics connecting macrostates placed far away from each other in the continuous space state. C1 state was characterized by a parallel β -sheet conformation in the proximity of the N-terminal together with half of the central loop of ring B (G18-L20), while C2 state adopted a closed coiled conformation. The escape time related to C1 (16.7 ns) was slightly higher in comparison to C2 (8.1 ns), this fact could be understood as higher energy barriers for the C1 state compared to C2. Additionally, this community was the most rigid one according to the RMSF plot (figure B.1.3.2.2) characterized by an average RMSF-value of ca. 3 Å. The C-community together with B1 state resulted to possess the

highest free energies, which are at least $1 K_B T$ higher than that of the reference state B2 (see table B.1.3.1).

State D maintained some similarities to macrostate A, due to be characterized by relatively closed random coil conformation. Indeed, the RMSF-values resulted to be very similar to each other (figure B.1.3.2.2), in line with previous evidences, highest values were found at the N- and C-terminal region together with the central loop or ring B region. State D is also well connected to the other macrostates by relatively low energy barriers as indicated by the non-negligible transition matrix elements (see transition matrix in Appendix - Bacteriocins) and the short escape time of only 6.6 ns (see table B.1.3.1).

Concluding with the states E and F, which were in line with the evidences described for state D. In both states, Bli α adopts a closed coil structure (as in state D) constituting the 15% of the population network. The E and F states resulted to be highly mobile, as it is depicted in the RMSF plots (figure B.1.3.2.2). Concisely, in the state F the peptide's flexibility involving residues 7 and 10 was significantly large compared to all other states. High thermal stability in both states was attributable to the values obtained for the escape times. However, it could be also determined that the state F was poorly connected to the main network despite of its low energy and high thermal stability. Just a weak transition to state E was predicted (see transition matrix in Appendix - Bacteriocins).

B.1.3.3 MSMs for Lch β Peptide

An initial microstate network of 725 nodes connected through of 1805 links was generated. After the same procedure used for the previous peptides, a final network consisting of 7 configurations and 25 transitions was obtained (figure B.1.3.3.1).

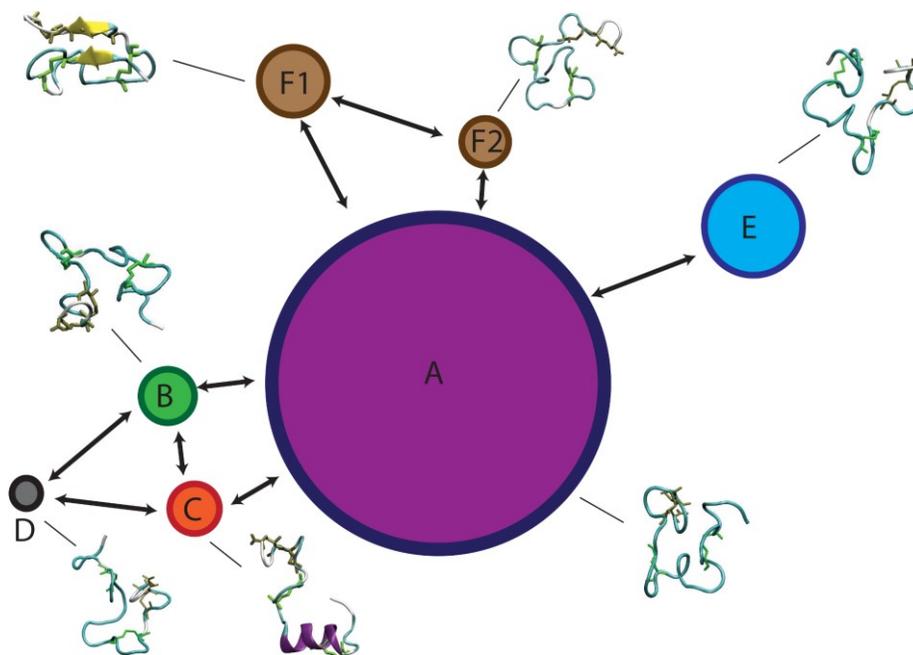


Figure B.1.3.3.1: MSM of Lch β peptide. The seven states are depicted as beads with size proportional to their occupation, arrows represent allowed transitions between states according to the corresponding transition matrix (see transition matrix in Appendix - Bacteriocins)

Six communities involving seven states defined the MSM network for the β -component. Consequently, differences in comparison with the α -peptides were found. Lch β network was mainly characterized by a central macrostate A, interconnected to the surrounded peripheral states.

State A was characterized by a coil structure where the highest population of the system (53%) was found. Herein, the escape time was low enough (< 8 ns) allowing numerous transitions to neighboring states causing a well-interconnected network. Low free energy barriers were determined as a result of the escape time values combined with a high population. According to that, state A could be considered as a wide shallow basin with low free energy and high entropy. Additionally, state A was determined to be the lowest in free energy ($2/K_B T$ lower than other states), which is why was afterwards considered as reference state for computing free energy differences. The peptide showed in this state a high flexibility, concisely in the ring A region (involving the residues 7 to 10) showing RMSF-values up to 7 Å. In agreement with the α -partners, the terminal regions displayed a high mobility (figure B.1.3.3.2)

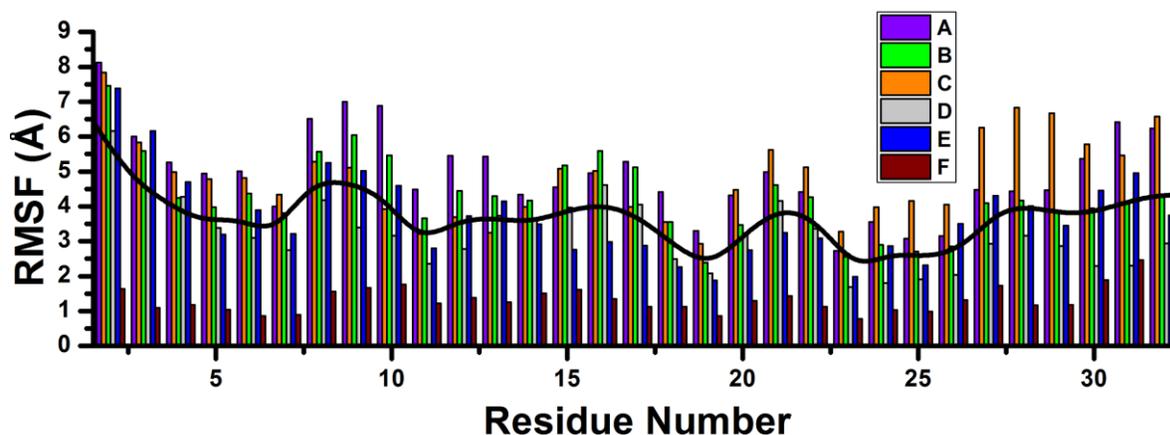


Figure B.1.3.3.2: Average C_{α} -RMSF value per residue per state for Lch β peptide.

Interestingly, only the macrostates C and F1 were found to be the states where the peptide adopts well-defined secondary structures. Concisely, in state C, an α -helix was built comprising the residues 7 to 16, while in the F1 state, a parallel β -sheet structure was detected between residue 5-6 and residues 26-27. Furthermore, although these two states were practically isoenergetic, (resulted to be connected to each other only via the central basin -state A-) their kinetic properties differ. From the table B.1.3.1, it can be interpreted that the α -helical state C, characterized by low energy barriers as a consequence of very low mean escape times allowed fast transitions to the surrounding states A, B and D while in contrast, state F1 resulted to be very confined showing a mean escape time of 173 ns. Additionally, despite that the F1-free energy does not differ meaningfully from the other states, this state resulted to be the most rigid of the entire network, as it can be interpreted from the RMSF-plot (figure B.1.3.3.2).

The peptide in the macrostates B, D and E were defined by random coil structures characterized with different kinetics and energetics properties. Interestingly, state B was predicted to be unstable due to the low mean escape time (5 ns), low occupation (8%) showing moderate transition rates as well. Indeed, it resulted to be almost isoenergetic with the α -helix state C ($\Delta G \sim 1.9 K_B T$). These two states showed low energy barriers, as it was reflected from the transition matrix elements TBC and TCB of 0.1 ns^{-1} and 0.2 ns^{-1} . Additionally, both of them connect the state A (main basin) with the peripheral state D. Interestingly, despite that D state showed a very low population (4%), the escape times were considerably high (40 ns) which allowed to predict that once the transition until state D takes place, the peptide remains more time inside of it.

8. Bacteriocins

Macrostate	π_i	RMSD /Å	t_e /ns	$\Delta G_i/k_B T$	S/k_B	$\langle \mu \rangle / D$
Lchα						
A1	0.047 ± 0.005	5.3 ± 1.3	30.0 ± 4.0	1.51 ± 0.02	0.194 ± 0.010	86.3 ± 22
A2	0.017 ± 0.003	2.3 ± 0.5	11.0 ± 0.4	2.55 ± 0.14	0.068 ± 0.009	104.0 ± 17
A3	0.009 ± 0.002	3.4 ± 1.7	2.3 ± 0.01	3.16 ± 0.01	0.042 ± 0.07	57.8 ± 24
A4	0.024 ± 0.003	5.2 ± 0.9	6.0 ± 0.6	2.18 ± 0.01	0.095 ± 0.008	82.1 ± 18
B	0.053 ± 0.005	5.1 ± 1.2	11.8 ± 0.7	1.39 ± 0.02	0.188 ± 0.021	92.7 ± 18
C1	0.054 ± 0.005	5.8 ± 1.0	4.5 ± 0.1	1.38 ± 0.02	0.157 ± 0.01	63.0 ± 27
C2	0.212 ± 0.010	7.3 ± 0.9	25.1 ± 6.0	0.00	0.552 ± 0.006	74.2 ± 21
C3	0.136 ± 0.08	7.1 ± 0.9	27.8 ± 0.2	0.64 ± 0.04	0.362 ± 0.008	67.3 ± 21
C4	0.072 ± 0.006	2.8 ± 0.9	143.0 ± 50.0	1.09 ± 0.03	0.189 ± 0.010	95.4 ± 15
C5	0.088 ± 0.007	4.9 ± 0.5	59.0 ± 12.0	0.88 ± 0.03	0.214 ± 0.010	75.1 ± 16
D1	0.051 ± 0.005	6.4 ± 1.0	5.0 ± 0.1	1.42 ± 0.02	0.240 ± 0.01	53.4 ± 23
D2	0.099 ± 0.007	6.5 ± 0.8	4.0 ± 0.1	0.77 ± 0.03	0.297 ± 0.009	49.8 ± 17
D3	0.056 ± 0.005	5.9 ± 0.8	3.0 ± 0.1	1.34 ± 0.02	0.189 ± 0.009	55.7 ± 19
D4	0.028 ± 0.004	4.3 ± 0.8	55.0 ± 12.0	2.04 ± 0.02	0.094 ± 0.01	70.8 ± 15
Blia						
A	0.147 ± 0.009	8.0 ± 1.2	4.7 ± 0.2	0.11 ± 0.05	0.501 ± 0.008	66.3 ± 21
B1	0.130 ± 0.008	6.1 ± 1.1	8.0 ± 0.5	1.45 ± 0.05	0.389 ± 0.008	67.9 ± 22
B2	0.164 ± 0.009	4.3 ± 0.9	7.2 ± 0.4	0.00	0.414 ± 0.007	81.0 ± 20
B3	0.097 ± 0.007	6.7 ± 0.8	5.4 ± 0.8	0.53 ± 0.04	0.292 ± 0.009	82.1 ± 17
C1	0.052 ± 0.005	6.5 ± 0.5	16.7 ± 1.4	1.15 ± 0.03	0.186 ± 0.010	58.6 ± 18
C2	0.029 ± 0.004	4.2 ± 1.0	8.1 ± 0.3	1.73 ± 0.02	0.121 ± 0.010	59.0 ± 19
D	0.119 ± 0.008	5.6 ± 0.8	6.6 ± 0.3	0.32 ± 0.04	0.185 ± 0.009	65.3 ± 22
E	0.147 ± 0.009	5.0 ± 1.2	11.1 ± 1.0	0.11 ± 0.05	0.381 ± 0.008	75.0 ± 17
F	0.148 ± 0.009	6.2 ± 0.9	40.0 ± 12.0	0.10 ± 0.05	0.517 ± 0.008	90.3 ± 16

Lch β						
A	0.534 ± 0.006	7.2 ± 0.9	7.8 ± 0.8	0.00	1.142 ± 0.006	74.3 ± 24
B	0.078 ± 0.009	6.9 ± 0.8	4.9 ± 0.1	1.92 ± 0.03	0.238 ± 0.009	70.3 ± 20
C	0.074 ± 0.005	7.2 ± 1.4	4.5 ± 0.1	1.98 ± 0.01	0.258 ± 0.014	69.7 ± 16
D	0.041 ± 0.009	5.5 ± 1.6	41.0 ± 7.0	2.57 ± 0.09	0.160 ± 0.011	92.2 ± 18
E	0.149 ± 0.009	8.8 ± 1.1	37.0 ± 10.0	1.28 ± 0.02	0.507 ± 0.008	57.9 ± 19
F1	0.087 ± 0.007	8.8 ± 0.4	173.1 ± 89.0	1.81 ± 0.01	0.254 ± 0.01	122.8 ± 13
F2	0.039 ± 0.004	1.4 ± 0.4	78.3 ± 23.0	2.62 ± 0.01	0.126 ± 0.009	123.5 ± 15

Table B.1.3.1: MSM properties of the Lch α , Blia and Lch β peptides. Occupation of each state π_i ; root mean square deviation, RMSD; average escape time, t_e ; free energy differences between states, (ΔG_{ij}) calculated with respect to the most occupied state; entropy, S average of the dipolar moment, $\langle \mu \rangle$. Standard errors associated with each quantity are indicated in gray.

B.1.4 Concluding Remarks of Conformational Dynamics of Lichenicidin Lantibiotic

From the combination of the 2 μ s-long all-atom MD simulation and the generation of the MSM models of Lichenicidin peptides Lch α , Blia and Lch β peptides, it is feasible to determine a:

- i. **Different kinetic and dynamic behavior of Lch α and Blia peptides:**
 - a. Although the sole difference between Lch α and Blia was placed exclusively in the N terminal region, the MSM network differ significantly from each other. In general, both peptides adopt most of the time random coil conformations in aqueous solution which is consistent with experimental observations.^[118,119,140] Furthermore, it was predicted that exclusively Lch α was able to adopt an α -helical conformation in the vicinity of ring A, while both of the them could fold in β -sheet conformations. Interestingly, the α -helical conformation observed for Lch α peptide seemed to be stabilized by the methylanthionine bridge involving the

residues 3-7, which is not present in the Bli α component. According to that, the undergoing secondary transition from α -helical to a β -sheet, and *vice versa*, was enabled exclusively for the Lch α peptide in aqueous solution, which may result from deficiencies of the CHARMM27 force field because it over-stabilizes α -helix folds.^[144]

ii. **Lch β peptide undergoes secondary structure transition:**

b. In analogy to the facts described for the Lch α peptide, the α -helical Lch β structure observed from NMR spectroscopy was able to turn to a β -sheet conformation involving the N-terminal residues Dhb5-Dhb6 and the C-terminal Lys27-Ala28. Both states resulted to be connected through transitions promoted via transient random coil conformations. It is important to mention that CHARMM force field tends to over-stabilize the α -helical content^[144] which combined with the effect of the solvent (methanol) in the NMR studies,^[118] promote an α -conformation for this peptide. In line with these facts, it has to be highlighted that the residence time in the β -sheet conformation (state F1) is one order of magnitude longer than in the α -helical state (state C). Consequently, it can be concluded that the structural properties together with the dynamic behavior of this lantibiotic in aqueous solution undoubtedly diverge from that observed via NMR spectroscopy in methanol solution.

iii. **Structural versatility is a key property of AMPs and lantibiotics for an efficient adaption in various biological environments:**

c. It has been reported that the presence of a membrane or any other reaction partner can perturb the delicate equilibrium between random coil conformations, α -helical structure and β -sheet fold (see next section B.2). These secondary structure transitions in aqueous solution have been observed for other AMPs in the past.^[54,145,146]

d. According to these results, it is feasible to extrapolate it to the lantibiotics. Indeed, the folding of random coil structures into α -helices or β -sheets upon interaction with membranes has been considered as an important process determining their mode of action.^[54,145,146]

B.2 Mechanism of Action of Two-Component Lantibiotic Lichenicidin by Means of All-Atom MD Simulations

MD simulations have been used to study more common antimicrobial peptides such as maganine or alamethicin in a lipid environment providing significant insights to their respective modus operandi.^[147,148] More challenging even, from an experimental or theoretical point of view, is the study of the family denoted as two-component lantibiotics where the Lichenicidin belongs. The presence of two components is an extra difficulty in order to understand and find out how these AMPs exert their bactericidal effects. In this regard, Lacticin 3147 (Ltn α and Ltn β peptides)^[127] is the best-experimentally studied system. It has been suggested that the peptides split the activities of targeting to lipid II and pore formation in the lipid surface. It has been proposed a three-step mechanism of action i) the α -peptide must attach to lipid II ii) a specific spatial arrangement between α and lipid II must be achieved to iii) allow the interaction of β peptide and promote the pore formation. Furthermore, according to the structural similarities found of α -components to Mersacidin, it has been suggested an analogue interaction pathway (section 8.1.1.1-B).

However, although several efforts such as the successful evaluation of the dynamics of lipid II component^[149] or the stability of the already built complex nisin-lipidII^[115] through MD simulations have been made, atomistic details regarding the unbiased attachment of a two-component lantibiotic to the lipid II docking molecule are not available, and consequently, several open questions regarding the suggested mechanism of action remain unanswered.

In order to shed light to this interesting paradigm, after the evaluation of the dynamics of the peptides in solution, the unbiased atomistic study of the mode of action of two-component lantibiotic lichenicidin by MD simulations was performed. The study was focused on the variant generated by *B. Licheniformis* VK21, Lch α and Lch β peptides (figure B.2) and the evaluation of the interaction pathways to Lipid II component.

B.2.1 Model Building and Set-up of the All-Atom MD simulation of Two-Component Lantibiotic Lichenicidin in a Biological Relevant Environment

To mimick a Gram-positive bacteria membrane, the protocol used in the previous study (A-Nisin Lantibiotic section A.2.3) was followed. Herein, just succinctly summarized.

In order to generate a Gram-positive model membrane CHARMM-GUI^[114] was used. The membrane consisted of POPG and POPE lipids in a 3:1 proportion, respectively. A 80 x 80 Å POPG/ POPE bilayer in the x- y plane was obtained. In parallel, the short 3LII variant (with three isoprene units) of lipid II- PDB entry-1WCO^[104] was used as template and manually extended in Gaussview program^[141] with the purpose of obtaining the natural eleven isoprene chain. Afterwards, the full lipid II was inserted in the upper layer of the membrane, followed by the manual deletion of 3 POPG and 1 POPE molecules in order to keep the ratio.

Solvation and ionization of the system was done using the VMD plug-ins^[26] according to the same experimental conditions used for the study of the peptides in solution (160 mM of NaCl, pH 7).^[140] The system full lipid II-POPG/ POPE membrane was energetically minimized, heated at 300 K, equilibrated decreasing progressively the constraints with further production free run stage of 400 ns. For this, the same MD-protocol described in Nisin Z lantibiotic and lipid II component was employed. After this preparatory step, the equilibrated lipid II-membrane was extracted to be used for the next stage including the peptides.

Subsequently, in the model consisting of an equilibrate lipid II-membrane a pair of lichenicidin Lch α and Lch β peptides was added, in such a way that α -component was placed 10 Å over the lipid II component while β -peptide 10 Å over α . The peptide structures used were related to the markov model states C2 and C for α and β peptides, respectively. These two states corresponded to the structures of the peptides after 400 ns of simulation in solution, concisely characterized by a coil structure (central basin) and an α -helical conformation for α and β peptides, respectively. Afterwards, the lichenicidin-lipid II-membrane model was replicated in the in the x- y- plane obtaining a final system consisting of 9 Lch α : 9 Lch β and 9 full lipid II embedded in a POPG/ POPE membrane. Once this big system was generated, the spatial arrangement of the peptides was manually modified in

order to provide different starting orientations (see figure B.2.1.1). Subsequently, the final system was solvated and ionized in VMD package^[26] at the same experimental conditions described before.^[140] The all-atom MD simulation was run in the NAMD 2.9 program^[16] under periodic boundary conditions. For Van der Waals (VdW) interactions and real space electrostatics a cut-off of 12 Å was applied, while long-range electrostatics was calculated with the Particle Mesh Ewald Summation.^[12]

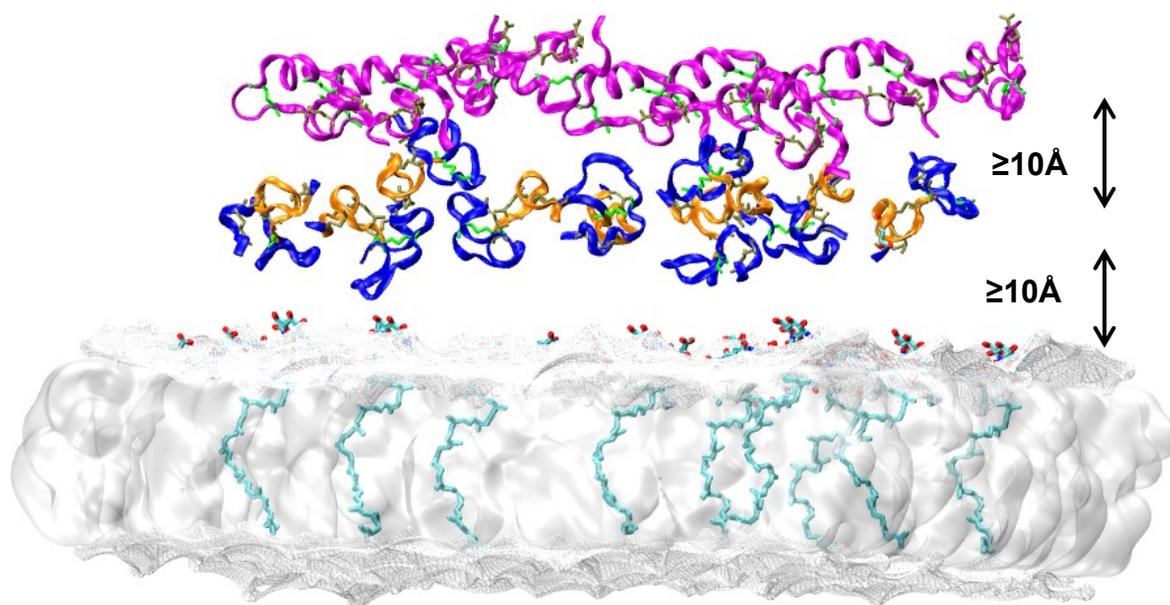


Figure B.2.1.1: Initial scenario for the study of the mechanism of action of two-component Lantibiotic Lichenicidin. System composed of 9 Lch α : 9 Lch β and 9 full lipid II embedded in a POPG/ POPE membrane. In the peptides, Lan and Mlan bridges are represented by green and olive colors, respectively. Additionally, the ring C/ D in α -components is highlighted with orange color. The peptides were placed 10 Å over the surface while the β -components were 10 Å over the α -ones.

Furthermore, the water molecules were described with the TIP3P model.^[109] In order to assure a timestep of 2 fs, the SHAKE algorithm^[113] was used to constrain all bond lengths between heavy and hydrogen atoms. Furthermore, according to the MD protocol explained in the chapter I section 3.7, the energy was minimized for 20000 steps with the conjugated gradient integrator employing an harmonic constrain on heavy atoms from 25 to 5 kcal/ (mol Å²). Afterwards, the system was heated during 100 ps using Langevin dynamics with a timestep of 2 fs and decreasing position restraints on the heavy atoms from 7.5 to 2.5 kcal/ (mol Å²). Subsequently, during the equilibration run, the harmonic constrains

on the system were progressively released. Concluding with the production run, where free dynamics during 1.3 μs at $T=300\text{ K}$ in an NPT ensemble under constant atmospheric pressure and temperature using the Langevin Piston method with a damping constant of 0.01 fs^{-1} ^[19] were simulated. The simulation was performed using the CHARMM27 force field for the canonical AA and membrane, the successfully validated CHARMM-compatible force field for the dehydroAAs together with the CGenFF for lipid II component.

This long-time simulated system was analysed by means of i) evaluation of the peptide-lipid II interactions through the number of contacts between the peptides and the docking molecules, ii) interaction energies, iii) evaluation of the dynamics of the polyisoprenoid anchor C55, iv) study of the aggregates-stabilization in order to identify the peptide key functional regions. For this point, the time-dependent contact maps (TdCM) per aggregate through the combinations α - α , β - β , α - β were generated, and further assessment of the membrane role was followed. Concluding with v) ring pattern analysis and vi) elucidation of Lys25 and Glu26 roles in the mechanism of action.

Through this unbiased 1.3 μs -long all-atom MD simulation, three aggregates (see figure B.2.1.2) in different peptides stoichiometries were achieved. Indeed, a free attachment of two α -components to distinct lipid II molecules interaction pathways was distinguished (further details section B.2.2.1).

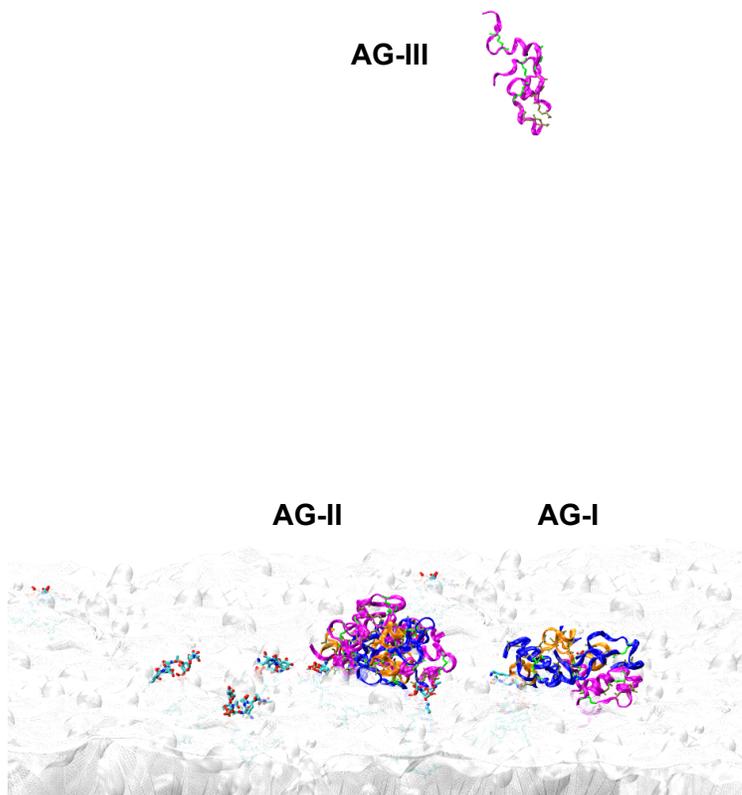


Figure B.2.1.2: Final scene after 1.3 μ s of all-atom MD simulation. Three aggregates are distinguished by their different stoichiometries: AG-I consisting of 4 α and 2 β peptides, AG-II stabilized by equal number of peptides 5 α :5 β while AG-III involved exclusively 2 β partners.

B.2.2 Analysis of the 1.3 μ s-long of All-Atom MD Simulation

B.2.2.1 Interaction with Lipid II, Cell-Wall Precursor

The aggregate AG-I is composed by $lch\alpha_A$, $lch\alpha_B$, $lch\alpha_C$, $lch\alpha_D$ and β_R , β_S peptides interacting intermittently with the surrounding lipids II: LII₁, LII₂, LII₆. The stabilization of this complex over the surface of the lipid membrane is mainly aided by the interaction produced between $lch\alpha_C$ to LII₂. Additionally, the remaining peptides adjacent to LII₁, LII₆ and very sparse presence of LII₇ contribute as well.

8. Bacteriocins

Three different phases can be distinguished during the trajectory: i) approach (first 500 ns), ii) reorientation (between 500 to 900 ns) and iii) stabilization (from 900 to the end). These three stages can be exemplified by the anchoring of I α _C to LII₂ (see figure B.2.2.1.1). During the first 500 ns the most solvent-exposed fragments of lipid II (mainly DAL7 and the sugar moiety NAG, which consequently influence the exposure of MUR) allowed the peptide approach. The Mlan 24-31 approaches to the sugar NAG triggering the access of Lys25, causing the interaction with the MUR moiety as well. Just at the beginning of this stage, an interaction produced between Lys12 and Lys14 to phosphate cage, NAG, MUR and FGA was detected. Interestingly, a more stable contact between Lys12 and hydrophobics Ile9, Leu10 and to DAL 7 was observed until the end of the stage. Furthermore, random interactions between Glu26-php's cage, Val23 and Mlan24-31-to ALA3 and FGA4, as well as Lys25 to FGA4 and Lys10 to LYS5 were also found. Interestingly, DAL6 does not play any role during this period (see figure B.2.2.1.1).

During the reorientation stage (500 to 900 ns) interactions between I α -rings B, C/ D to the lipid II-sugars took place, concisely involving the residues Leu14, Mlan24-31 and Lys25 to NAG combined with the interactions of Lys12, Leu14, Mlan24-31, Lys25, Glu26 to MUR. Due to the spatial rearrangement noticed during this stage, the Lys12 started to interact with ALA3, FGA4 and LYS5. This interaction caused the approach of the hydrophobic residues Ala8, Ile9 and Leu10 to the same fragments of lipid II. Interestingly, during this stage no interaction with the phosphate moiety and the terminals DAL6 and DAL7 were noticed (see figure B.2.2.1.1).

The peptide arrives at the stabilization stage after 900 ns. From this point to the end, the interaction produced between Lys25 and the phosphate cage is crucial. This interaction is also influenced by the proximity of the Mlan24-31 to the MUR moiety. It was also noticed that the approach of Glu26 to LYS5 and DAL7 triggered contacts of Lys25 with FGA4 and DAL6 (to the latter one because of proximity). Furthermore, around 1 μ s the hydrophobic residues Ile9 and Leu10 showed close contacts with ALA3, FGA4, DAL6 and DAL7.

Mainly produced by the 'salt-out' effect,^[150] the region consisting of the hydrophobic residues Ala8, Ile9 and Leu10, provides a favourable mobility between I α -rings A and B. It can be suggested that this hydrophobic region constitute the N_T-hinge region of the I α -component. A flexible N_T-hinge region has been also detected for nisin peptide, which is composed by the three AAs of distinct nature: Asn20, Met21 and Lys22 (see figure A.1.1).

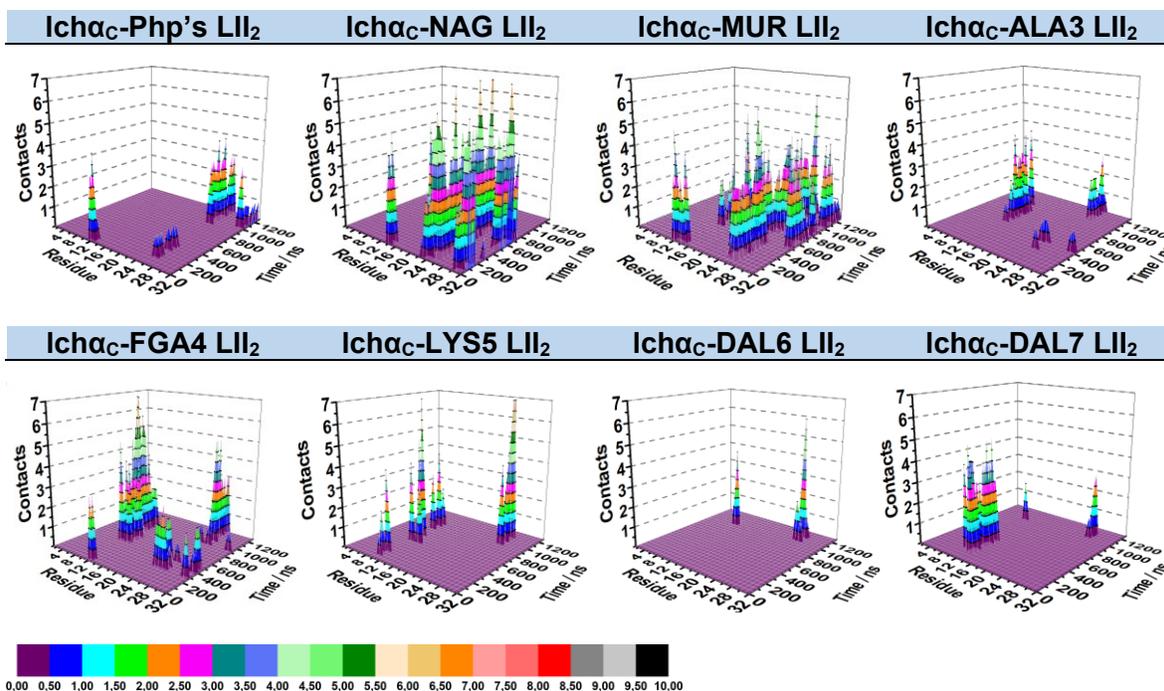


Figure B.2.2.1.1: Interaction pathway followed by Icha_C to LII₂. Monitoring of the number of contacts per residue of Icha_C to the structural regions of the lipid II in the course of 1.3 μs of all-atom MD simulation.

The interaction involving Icha_B-LII₁ resembles the previously explained approach. Only few and weak interactions involving the ring A (residues 3 to 7) together with Ala8 and Ile9 were sequentially found to i) sugars moieties, ii) LYS5, DAL6 and DAL7, iii) MUR and FGA at the end of the simulation. While no interactions were detected with ALA3 or with the phosphate cage (see figure B.2.2.1.2).

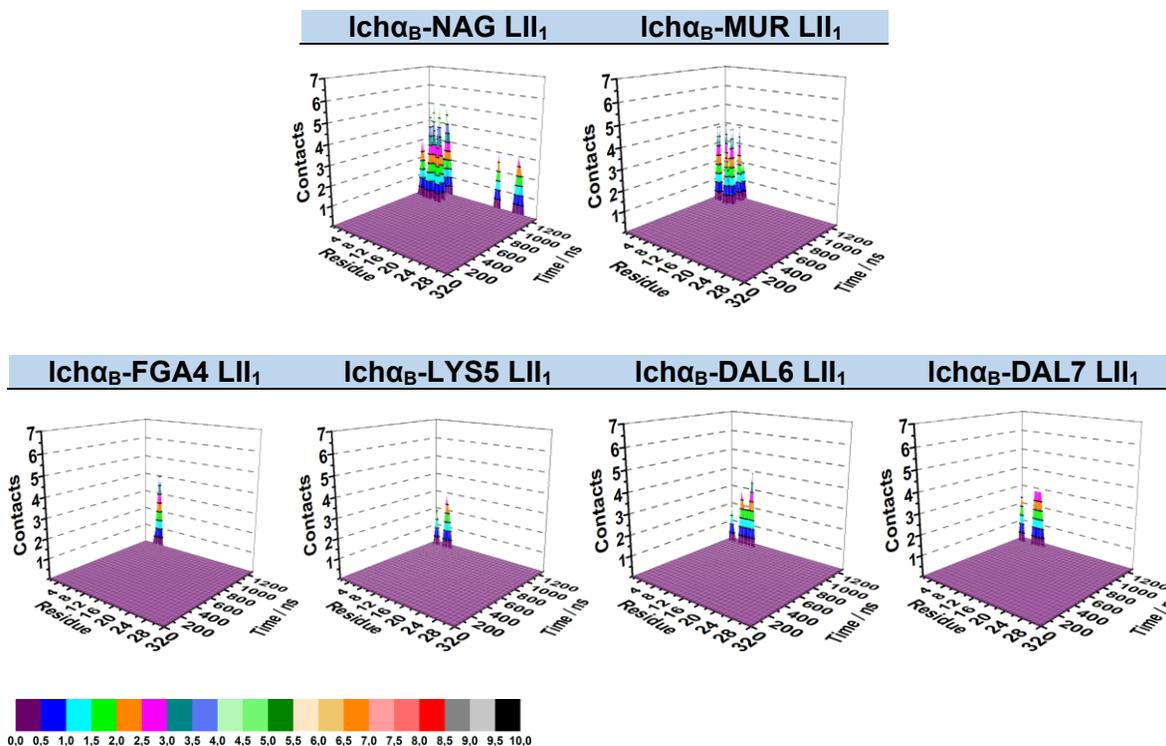


Figure B.2.2.1.2: Interaction pathway followed by Icha_B to LII₁. Monitoring of the number of contacts per residue of Icha_B to the structural regions of the lipid II during 1.3 μs of all-atom MD simulation.

In general, during the approaching stage (first 500 ns) an unstable interacting pathway was described by peptides in the surrounding of the lipids II constituents:

- Asn16 and Tyr19 of Icha_A showed short interaction exclusively to the NAG sugar of LII₂ (see figure B.2.2.1.3).

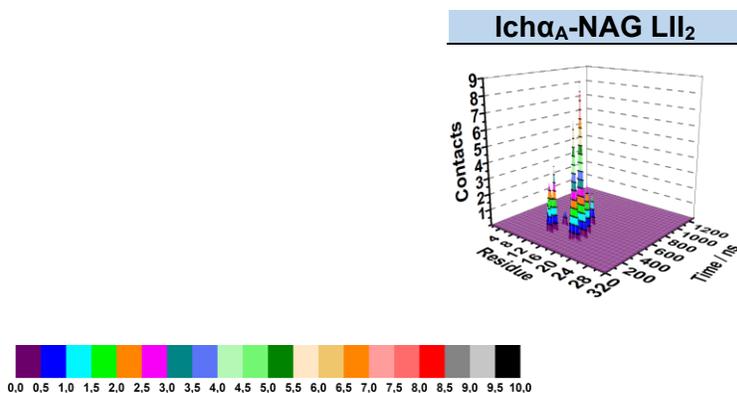


Figure B.2.2.1.3: Interaction pathway followed by Ich α _A to LII₂. Monitoring of the number of contacts per residue of Ich α _A to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

- Similar behaviour was observed regarding the hydrophobic Ile2 and Leu4 of Ich α _B to LII₆ (see figure B.2.2.1.4).

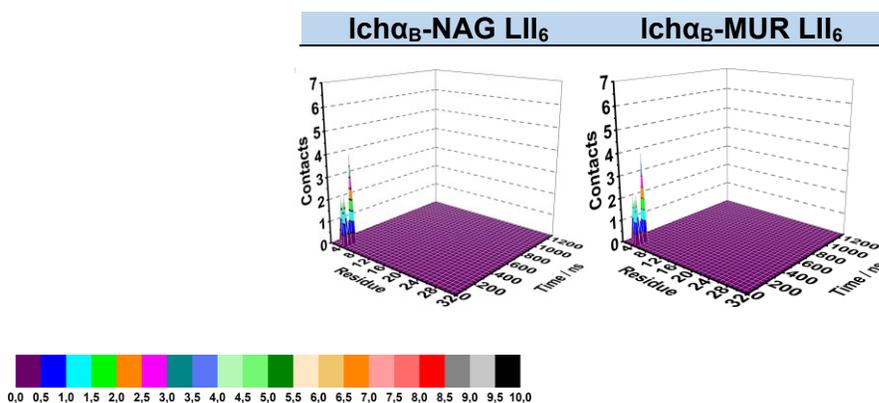


Figure B.2.2.1.4: Interaction pathway followed by Ich α _B to LII₆. Monitoring of the number of contacts per residue of Ich α _B to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

On the other hand, different findings were detected for the Ich β -peptides. In general, Ich β _R and Ich β _S showed weak and unstable interactions with LII₁. For example, Ich β _R-residues Dhb6, Dha8, Trp9, Thr10, Ile12, Lys27 and Mlan29-32 showed the first interactions with DAL6 (only Dhb6 and Mlan 29-32) and DAL7 after 500 ns of simulation. Afterwards, only few contacts between Lys27 with the phosphate cage and the sugars, with a further influence of Dhb6 and Mlan29-32 were detected (figure B.2.2.1.5).

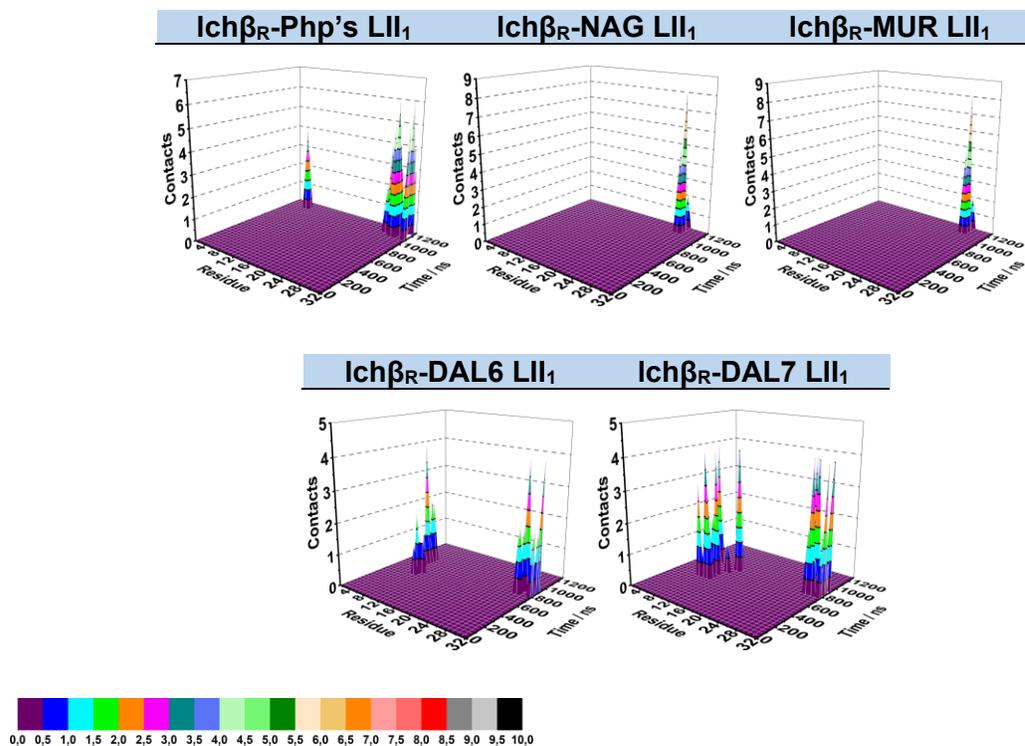


Figure B.2.2.1.5: Interaction pathway followed by $Ich\beta_R$ to LII₁. Monitoring of the number of contacts per residue of $Ich\beta_R$ to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

For $Ich\beta_S$, the same scenario is found: Obu1 and Dhb2 generated random interactions in the following sequential order to NAG, DAL7, LYS5, and MUR (see figure B.2.2.1.6).

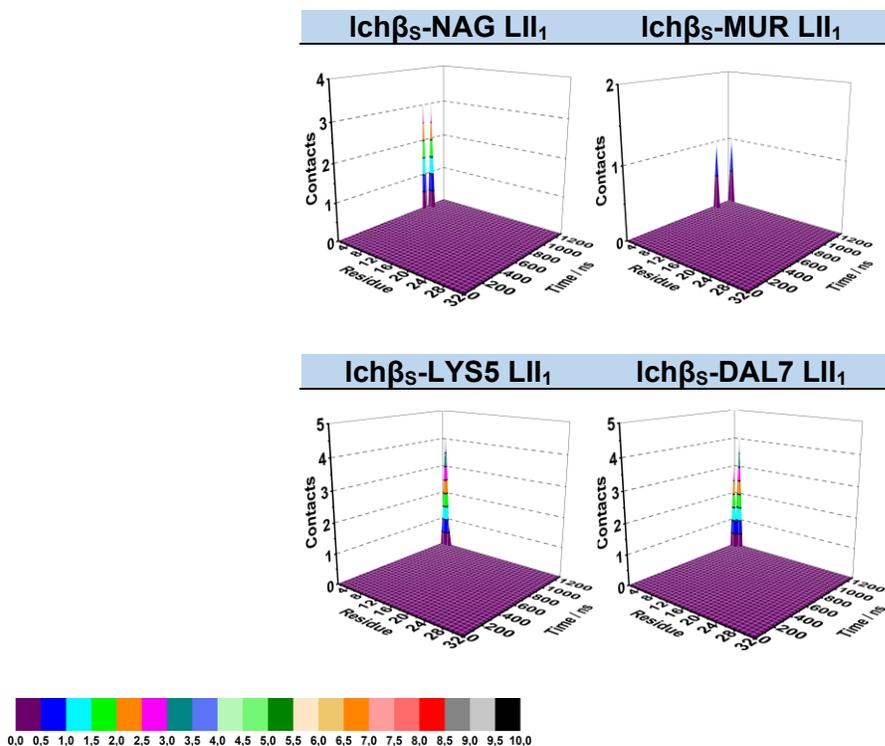


Figure B.2.2.1.6: Interaction pathway followed by Ich β_S to LII $_1$. Monitoring of the number of contacts per residue of Ich β_S to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

Interestingly, some differences are found for the Ich β_S in the surrounding of LII $_6$ (see figure B.2.2.1.7). During the first 500 ns, a first approach to php's moiety and FGA4 by Lys27 is noticed. This fact triggers the approximation of Dhb5 to MUR causing in a parallel manner the approach of Ala4 and Dhb6 to the phosphate moiety and to ALA3, respectively. The strong interaction produced by Lys27-php's promoted the immersion of Lys27 in the membrane establishing contacts with the two first isoprene moieties and, subsequently, promoted a favourable interaction of Mlan29-32 with ALA3 and FGA4, which was also affected by the interaction produced between Arg31 and DAL7 and due to proximity with DAL6. This observation was also aided by electrostatics interactions detected between Glu26 and LYS5 having as side effect the approach of Mlan29-32 to LYS5. At 400 ns, only minor interactions were noticed involving: Dhb5 to NAG; Dhb2 and Dhb5 to MUR2 which was followed same nanoseconds later by the contacts detected between Dhb5 to ALA3, Dhb6 to FGA4 and LYS5. No interactions were detected after that point between the peptide and LII $_6$, with the exception of a random contact between Dhb2 and ALA3 at 770 ns.

8. Bacteriocins

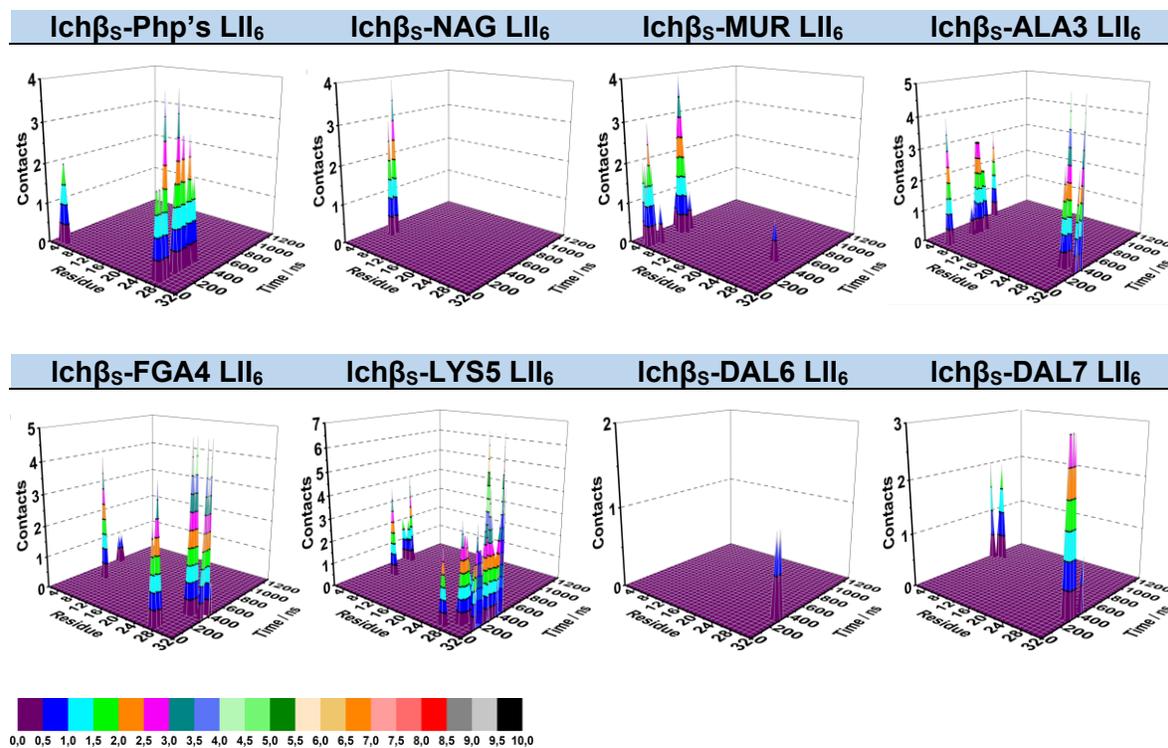


Figure B.2.2.1.7: Interaction pathway followed by Ich β_S to LII $_6$. Monitoring of the number of contacts per residue of Ich β_S to the structural regions of the lipid II during 1.3 μ s of all-atom MD simulation.

A general overview of the peptides interaction with lipid II molecules is provided in the figure B.2.2.1.8

AG-I		Approach	Reorientation	Stabilization
LII ₁	α_B			
	β_R			
	β_S			
LII ₂	α_A			
	α_C			
LII ₆	α_B			
	β_S			

Random
 Moderated
 Strong

Figure B.2.2.1.8: Schematic overview of the interaction pathway followed by Lch α and Lch β AG-I components with the surrounding Lipid II molecules. The lipid II molecules involved in this aggregate (LII₁, LII₂, LII₆, LII₇) showed seldom interactions with Lch α and Lch β peptides during all different stages of the simulation time (approaching, reorientation and stabilization). Only a single Lch α peptide (lch α_C) was able to keep a stable electrostatic interaction with a docking molecule (LII₂).

The complex AG-II was composed by 5 lch α -components (α_E , α_F , α_H , α_J , α_G) and 5 lch β -peptides (β_N , β_Q , β_V , β_U , β_X) which showed some random contacts to the lipids II LII₃, LII₄, LII₅ and LII₉. Analogous to AG-I, despite the elevated number of α and β components, this aggregate only showed a single stable lch α -lipid II interaction, specifically, produced between lch α_H and LII₅. However, it is noteworthy that the random interactions observed between the remaining lch α -peptides with the closest cell wall subunits used the same regions as in AG-I, consisting of the lch α -rings A, C/ D together with N_I-hinge region.

The stable interaction produced between lch α_H and LII₅ involved predominantly Mlan3-7 combined with some weak contributions of Mlan22-27 and Glu26 to NAG, as well as some hydrophobic influence of Ile2 to MUR was observed. This behaviour triggered the electrostatic interaction of Glu26 with the phosphate-cage. As it can be followed in the figure B.2.1.1.9, once the electrostatic interaction between Glu26 with the NAG took place, the approach to MUR followed, promoting the stabilization of the complex lch α -lipid II. This step was also aided by the phosphates. Interestingly, as a consequence of the spatial arrangement gained by the lch α -peptide, immersion of the N_I-hinge region, Ile9 and Leu10 in the membrane was noticed. This fact provoked some interactions with the isoprene units 3-8 and 11 (figures B.2.2.1.9).

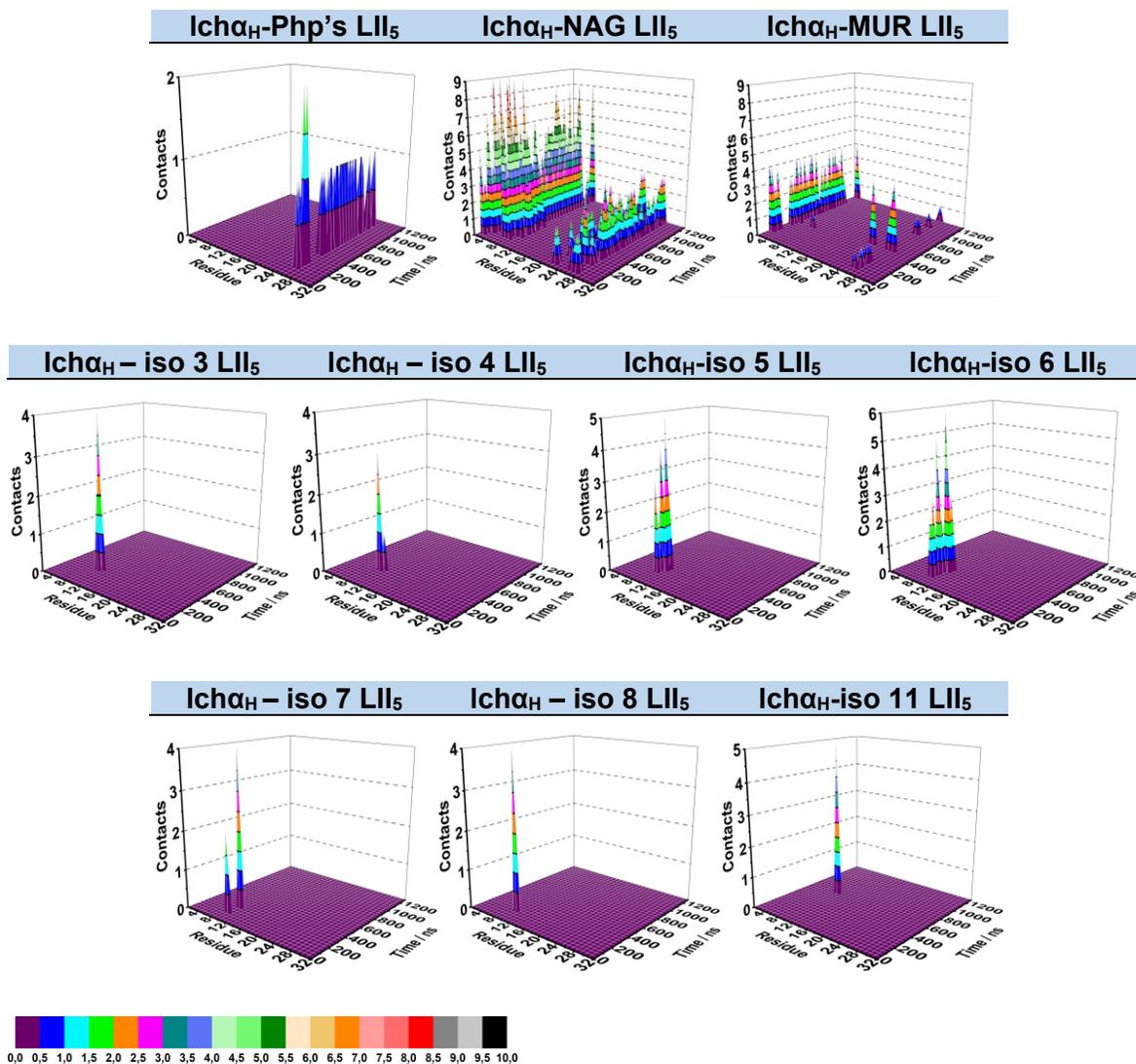


Figure B.2.2.1.9: Interaction pathway followed by Icha_H and LII₅. Monitoring of the number of contacts per residue of Icha_H to the structural regions of the lipid II in the course of 1.3 µs of all-atom MD simulation.

The rest of interactions detected between Icha_J-LII₃ (figure B.2.1.1.10), Icha_H-LII₄, Icha_E-LII₅, Icha_F-LII₉, (see figures B.2.1.1.11) resulted to be random. Interestingly, in none of them the lipid II phosphate moiety was involved, which also contributes to determine the importance of the phosphate cage for the stable formation of the Icha_α-lipid II complex. It is particularly remarkable that the lipid II regions sensitive to interact nearby of AG-II can be extrapolated from a peptidoglycan subunit to other. In most of the possible interactions earlier mentioned, the interaction to the highly solvent exposed NAG sugar seems to be aided by the contact with MUR. Additionally, the AG-II peptides also approached to the DAL6

and DAL7. This alternative approach to the lipid II components seems to be unfavourable, due to the unstable contacts within Lch α peptide and lipid II complexes. These periodic interactions were detected randomly during the MD simulation, which encourage to hypothesize the role of the side chain lipid II terminals as recognizer elements.

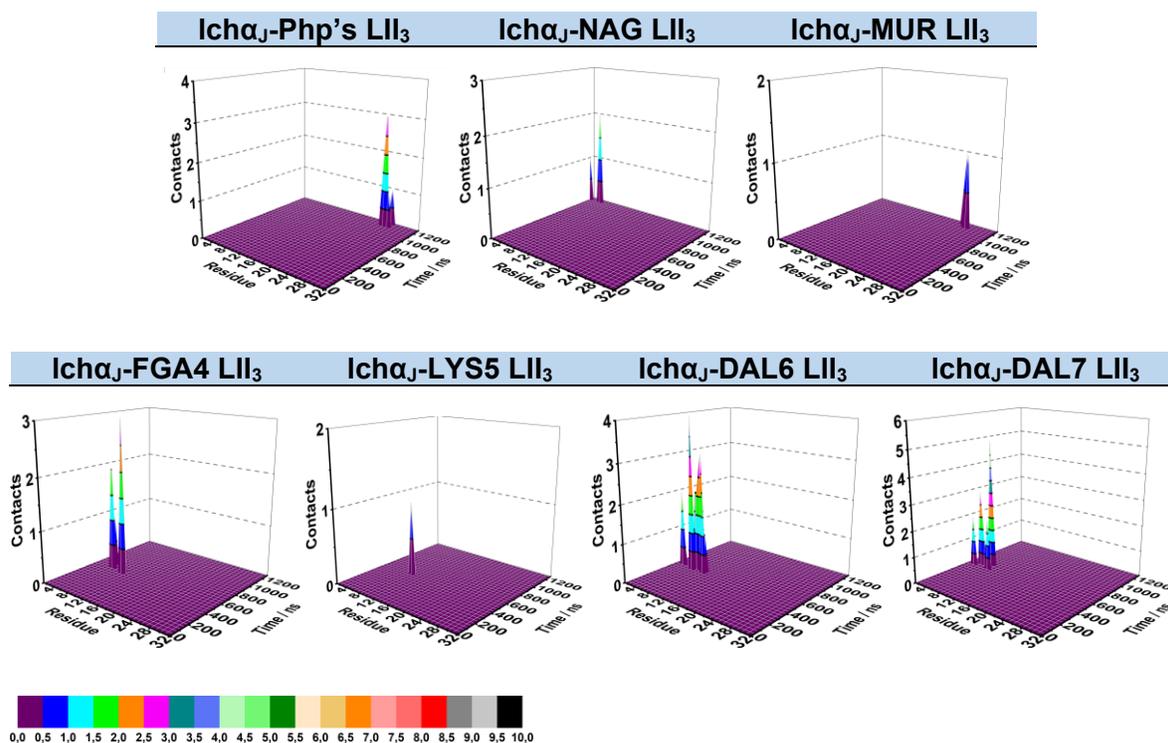


Figure B.2.2.1.10: Interaction pathway followed by Lch α _H-LII₃. Monitoring of the number of contacts per residue of Lch α _H to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

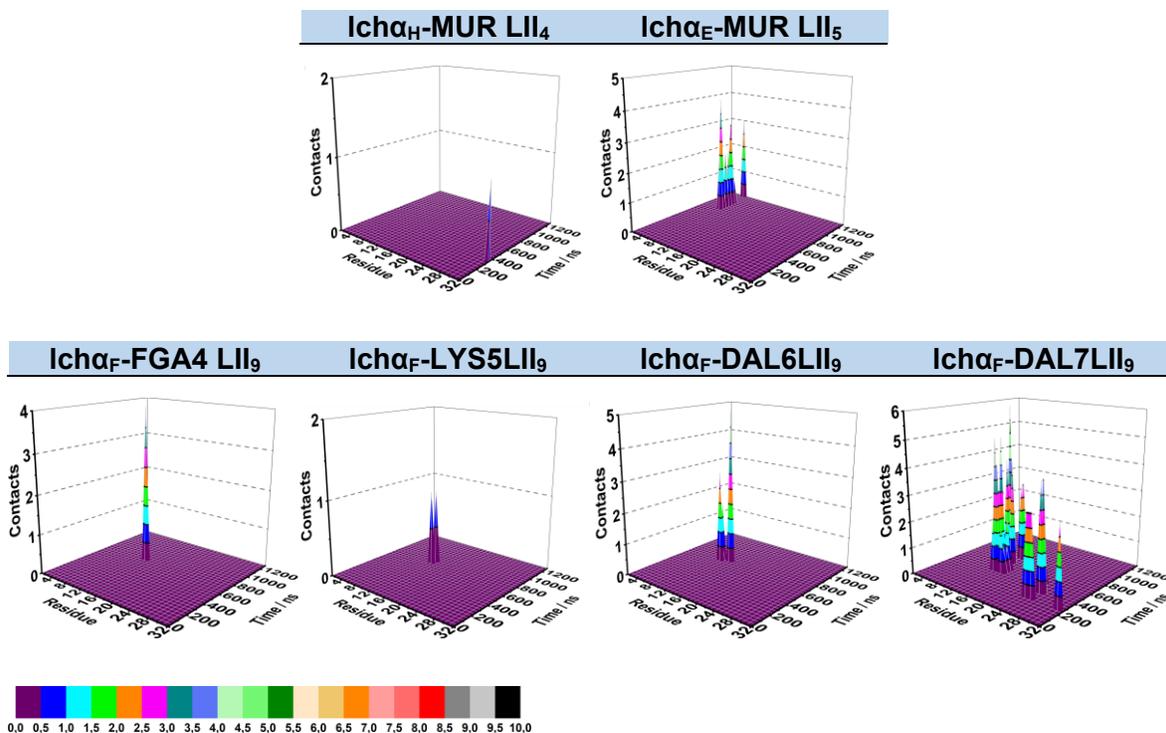


Figure B.2.2.1.11: Interaction pathway followed by Icha_H-LII₄, Icha_E-LII₅, Icha_F-LII₉. Monitoring of the number of contacts per residue of Icha_H, Icha_E and Icha_F to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

Furthermore, the Icb β -peptides showed also very few occasional interactions with the lipid II molecules. The pair of possible contacts Icb β _Q-LII₄ and Icb β _V-LII₅ (figure B.2.2.1.12), Icb β _Q-LII₅ (figure B.2.2.1.13), Icb β _V-LII₉ (figure B.2.2.1.14), Icb β _X-LII₃ (figure B.2.2.1.15)–shared the same part of lipid II to interact with, the sugar moiety NAG. Additionally, due to sporadic proximity of Icb β peptides to MUR, FGA4, LYS5, DAL6 and DAL7, some contacts were detected but in very minor extent. From all those possible peptide-lipid combinations, Icb β _Q-LII₅ was found to be the most stable (figures B.2.2.1.20-21). In analogy with the behaviour observed for the Icb β -peptides in AG-I, the AG-II Icb β -residues involved in the occasional interactions with lipid II are similar, involving predominantly the flexible N_T region, ring A, Ser 21, Leu22, Lys27 and regularly the whole ring C.

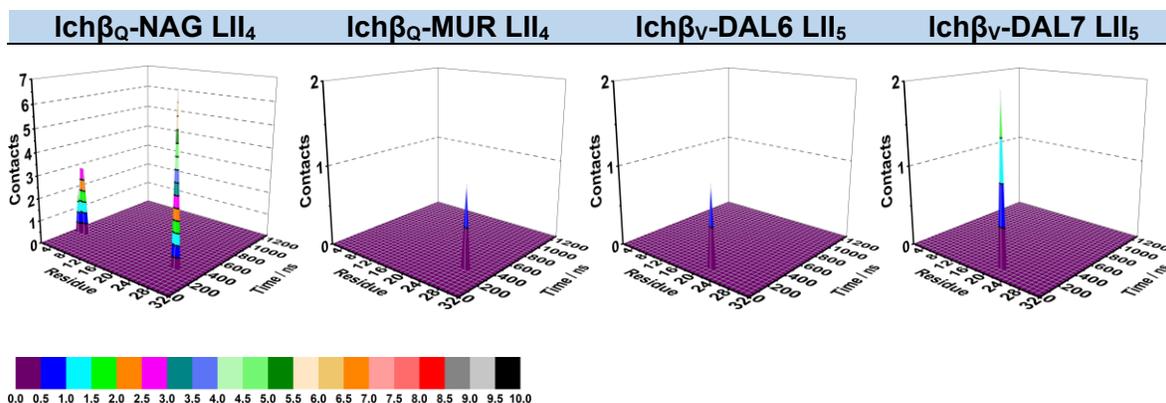


Figure B.2.2.1.12: Interaction pathway of $Ich\beta_Q$ -LII₄, $Ich\beta_V$ -LII₅. Monitoring of the number of contacts per residue of $Ich\beta_Q$ and $Ich\beta_V$ to the structural regions of the lipid II during 1.3 μ s of all-atom MD simulation.

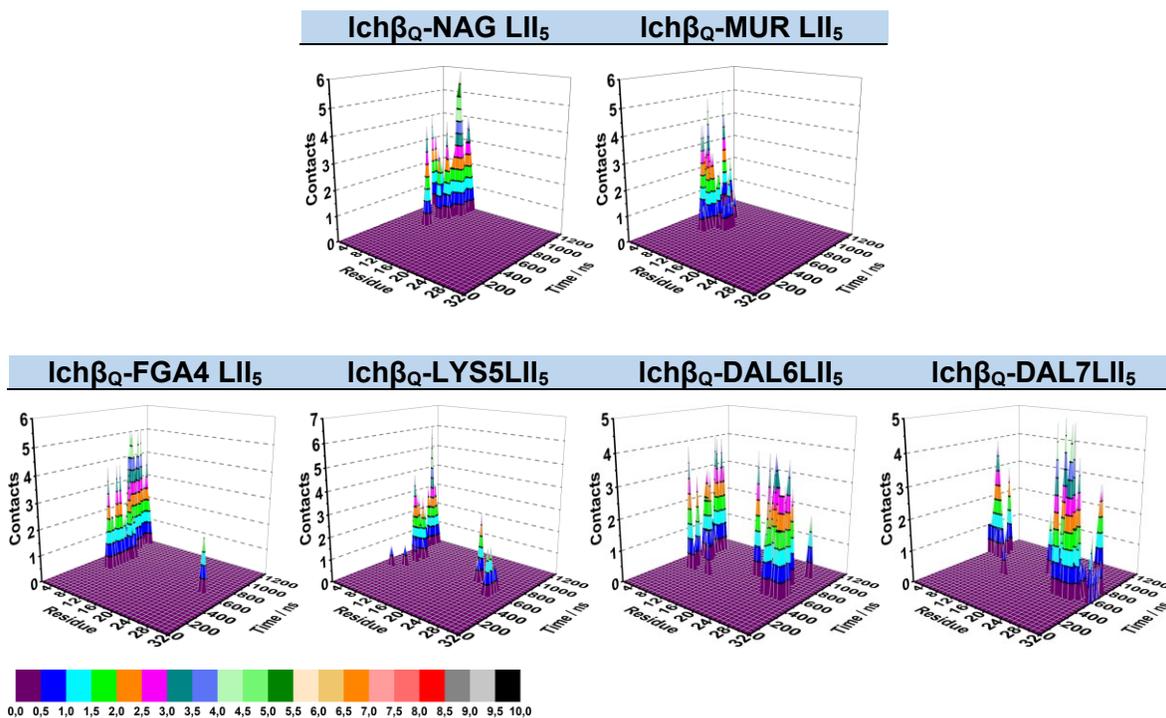


Figure B.2.2.1.13: Interaction pathway of $Ich\beta_Q$ -LII₅. Monitoring of the number of contacts per residue of $Ich\beta_Q$ to the structural regions of the lipid II in the course of 1.3 μ s of all-atom MD simulation.

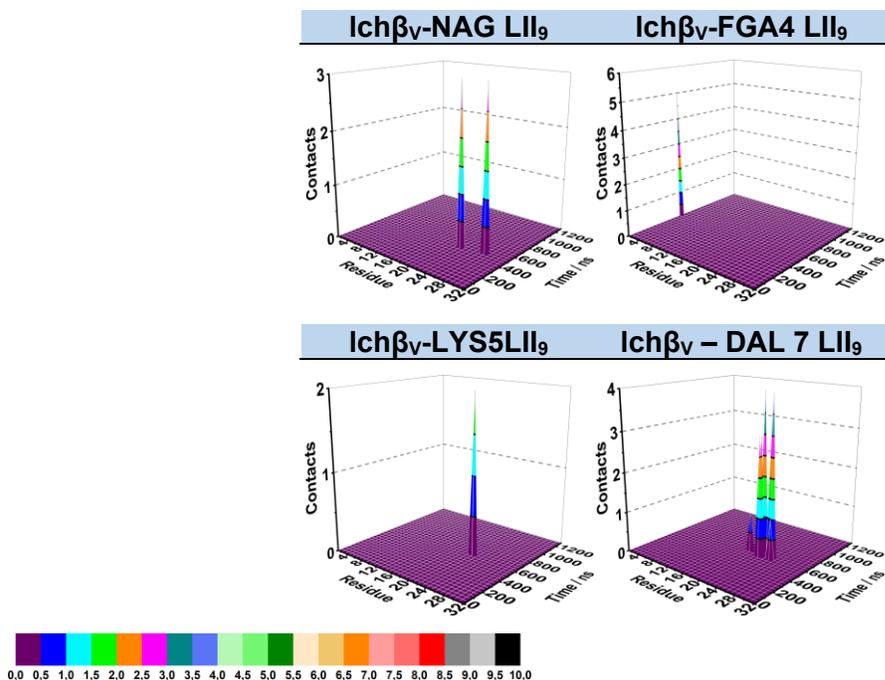


Figure B.2.2.1.14: Interaction pathway of $Ich\beta_V$ - LII_9 . Monitoring of the number of contacts per residue of $Ich\beta_V$ to the structural regions of the lipid II in the course of 1.3 μs of all-atom MD simulation.

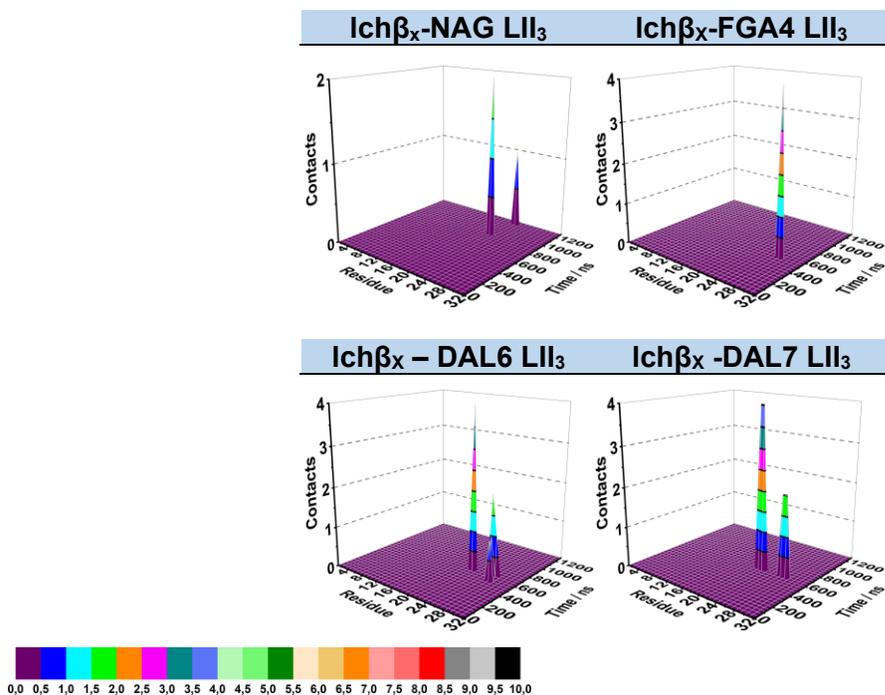


Figure B.2.2.1.15: Interaction pathway $Ich\beta_X$ - LII_3 . Monitoring of the number of contacts per residue of $Ich\beta_X$ to the structural regions of the lipid II during 1.3 μs of all-atom MD simulation.

A general overview of the peptides interaction with lipid II molecules is provided in the figure B.2.2.1.16.

AG-II		Approach	Reorientation	Stabilization
LII ₃	α_J			
	β_X			
LII ₄	α_H			
	β_Q			
LII ₅	α_E			
	β_V			
	β_Q			
	α_H			
LII ₉	α_F			
	β_V			

Random
 Moderated
 Strong

Figure B.2.2.1.16: Schematic overview of the interaction pathway followed by Lch α and Lch β AG-II components with the surrounding Lipid II molecules. The lipid II molecules involved in this aggregate (LII₃, LII₄, LII₅, LII₉) showed seldom interactions with Lch α and Lch β peptides during all different stages of the simulation time (approaching, reorientation and stabilization). Analogue to AG-I, only a single Lch α peptide (α_H) was able to keep a stable electrostatic interaction with a docking molecule (LII₅).

A third aggregate, namely AG-III, composed exclusively by two lch β -peptides, concisely lch β_T and lch β_O , was also found and contrary to the other two complexes does not interact with any lipid II.

B.2.2.2 Interaction Energies

Additionally, the interaction energies of the peptides which showed a stable interaction with lipid II molecules were evaluated, concisely $I\alpha_C$ -LI₂ and $I\alpha_H$ -LI₅.

The figure B.2.2.2.1 shows the monitoring of the interaction energies for $I\alpha_C$ -LI₂ and $I\alpha_H$ -LI₅, A- and B-, respectively. According to these results, the interactions were mainly of electrostatic nature (red trace) with a bipolar character between α -peptide and the cell wall structural subunit. Contribution of the VdWs (blue trace) interactions is low to the total interaction energy (black trace).

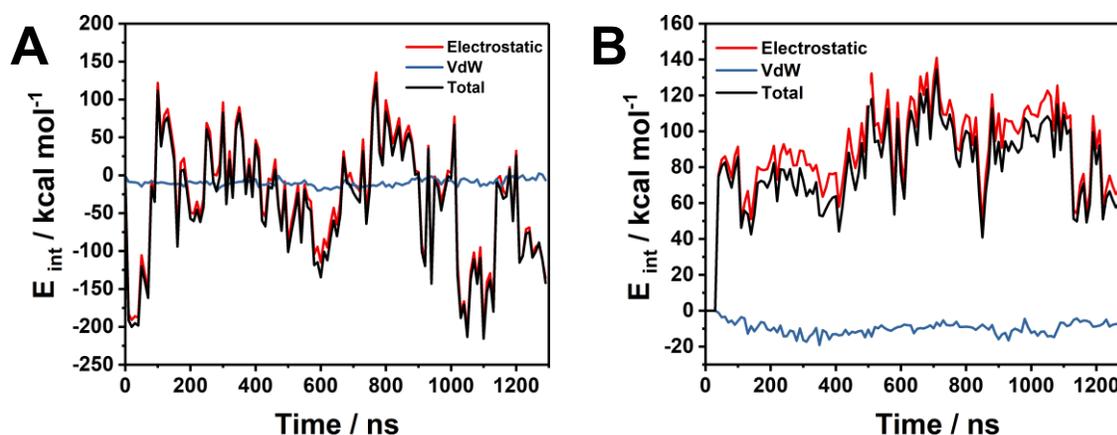


Figure B.2.2.2.1: Interaction Energies between $I\alpha$ peptides and lipid II component. Black trace denotes the total interaction energy, blue trace the Van der Waals while the red trace displays the electrostatic contributions, respectively. Mainly, both peptides approach to lipid II PGN precursor through electrostatic interactions.

The α -rings C/ D, concisely the residues Lys25 and Glu26, mainly accommodated the electrostatic interaction with the phosphate cage (see figure B.2.2.2.1). These results are in line with experimental evidences because it was suggested that the charged amino acids control the binding to lipid II component, instead of hydrophobic interactions.^[131]

B.2.2.3 Dynamics of the Undecaprenyl Pyrophosphate

All lipid II molecules showed the same behaviour. An ongoing equilibrium between two different configurations, named as extended configuration-E and V-shaped configuration (see A-figure B.2.2.3.1) was found. In order to capture the transition configurations, the

distance between the first carbon atom of the C55-isoprene chain (C72) to the last principal carbon of the C55-isoprene chain (C283) was monitored (see lipid II structure in Appendix - Bacteriocins). During the examination of the lipid II molecules interacting with AG-I (see B-figure B.2.2.3.1) and for those in contact with AG-II (see C-figure B.2.2.3.1), as well as the LI₈ (the only one without any interaction with lichenicidin), the distance between the aforementioned atoms fluctuated between two extremes of about 5 Å – 40 Å during the 1.3 μs of simulation. The flexibility exhibited by C55 during the simulation open diverse options to understand how the translocation of the lipid II takes place. This process is unknown, and does not belong to the scope of this work.^[151] The findings described here are in agreement with the theoretical evidences reported by Chugunov *et al.*^[149] and Koch *et al.*^[115]

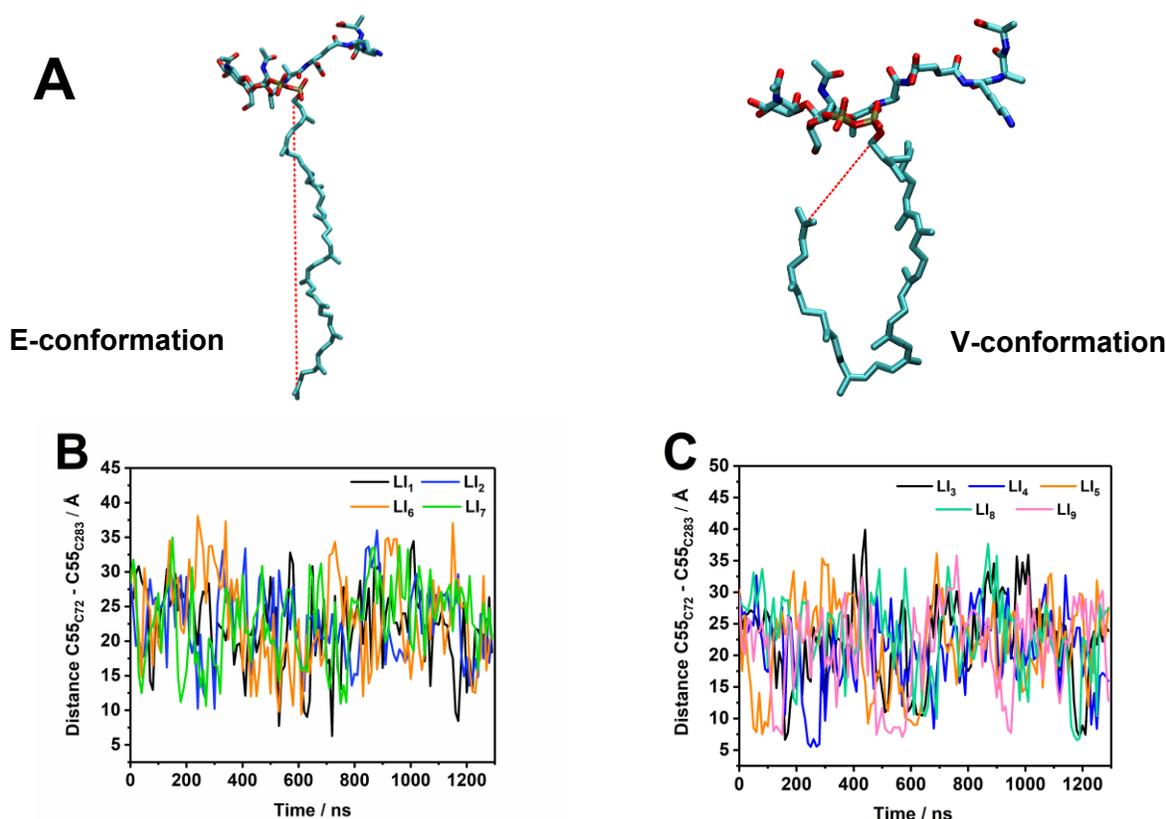


Figure B.2.2.3.1: Ongoing equilibrium of the undecaprenyl pyrophosphate. A- Graphical representation of the two-detected lipid II configurations: (upper right) E-extended shape, (upper left) V-configuration. **B-** Monitoring of the C55-isoprene chain for the lipid II molecules involved in AG-I. **C-** Monitoring of the C55-isoprene chain for the lipid II molecules involved in AG-II and lipid LI₈.

B.2.2.4 Aggregate-Stabilization

The stabilization of the three aggregates was studied through the TdCMs evaluating all possible interactions between $I\alpha_A$ - $I\alpha_A$, $I\alpha_A$ - $I\alpha_B$, and $I\alpha_B$ - $I\alpha_B$ peptides of each aggregate.

B.2.2.4-A Interactions Inter-AG-I Peptides

AG-I: $I\alpha_A$ - $I\alpha_A$. The peptides $I\alpha_{A_A}$ and $I\alpha_{A_B}$ started to interact at the end of the approaching phase (first 500 ns) involving mainly the residues of the ring C/ D: Pro29 and Ser30 together with Asn32 of $I\alpha_{A_A}$ to the Lys25 and Asn32 of $I\alpha_{A_B}$ (see figure B.2.2.4-A.1).

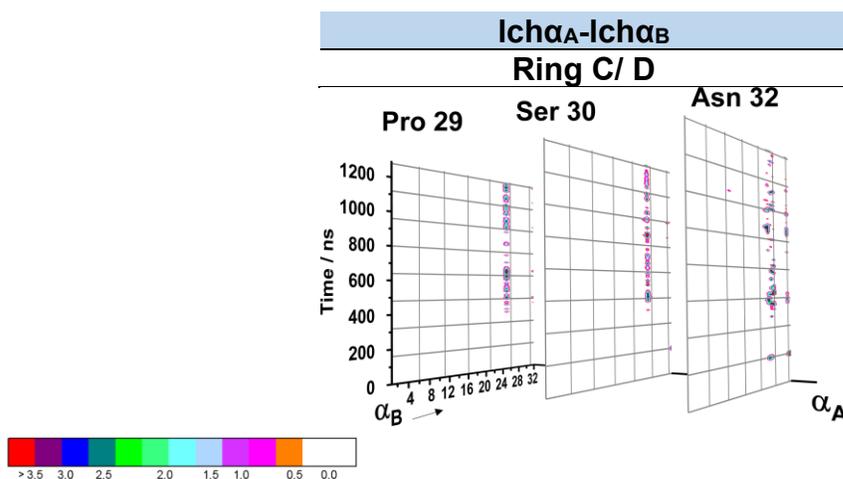


Figure B.2.2.4-A.1: TdCM of the interaction pathway followed between $I\alpha_A$ and $I\alpha_B$ peptides. From the end of the approaching phase until the end of the simulation, the rings C/ D (Asn32-Lys'25) are the main promoters of the electrostatic interaction.

Few nanoseconds after the beginning of the simulation are enough to detect the interaction between $I\alpha_A$ - $I\alpha_C$ peptides. In this case, $I\alpha_A$ uses the N_r-region (Ile 2) combined with the ring A (Leu4) promoting the first hydrophobic contacts to Mlan3-7 and Mlan22-27 of $I\alpha_C$ peptide. Additionally, contributions of Lan11-21, Tyr19 and Leu20 to the $I\alpha_C$ -rings C/ D were also noticed. Furthermore, the $I\alpha_A$ -Met28 and $I\alpha_A$ -Pro29 contribute through its hydrophobic interaction to the $I\alpha_C$ -Gly18, Met28 and Ile2. During reorientation, the interaction is mainly governed by hydrophobic interactions between $I\alpha_A$ -Leu4 and $I\alpha_A$ -Met28 to the same $I\alpha_C$ -residues observed in the preceding stage. Due to their proximity, contacts between $I\alpha_A$ -Tyr19 to the $I\alpha_C$ -Asn17 and rings C/ D were

also noted and attributed to sidechain H-bond formation. From 900 ns to the end, the peptides remained stabilized through the hydrophobic interactions predominantly produced between $\text{Ich}\alpha_A$ -Ile2, Tyr19 and Met28 to $\text{Ich}\alpha_C$ -Mlan3-7 and Mlan22-27, Asn17, Gly18 respectively. As it can be followed in the figure B.2.2.4-A.2, the surrounding side chains to Met28 also contributed to the interpeptide contacts.

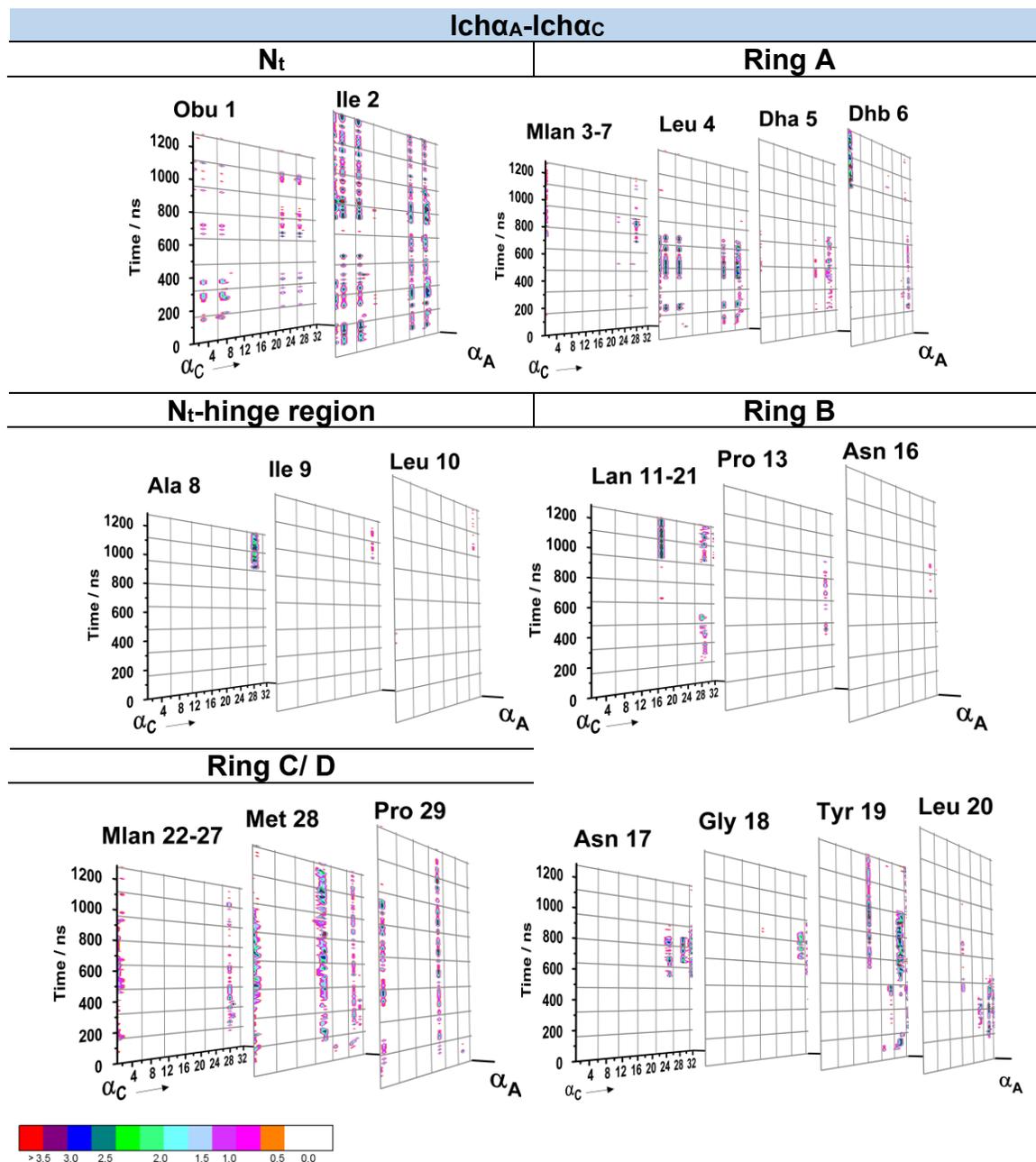


Figure B.2.2.4-A.2: TdCM of the interaction pathway followed between $\text{Ich}\alpha_A$ and $\text{Ich}\alpha_C$ peptides. Hydrophobic forces maintain the peptides in contact.

8. Bacteriocins

The electrostatic interaction found between $\text{I}\alpha_A$ - $\text{I}\alpha_D$ was established initially by $\text{I}\alpha_A$ -Lys25 to $\text{I}\alpha_D$ -Glu26 and $\text{I}\alpha_D$ -Asn32 (see figure B.2.2.4-A.3).

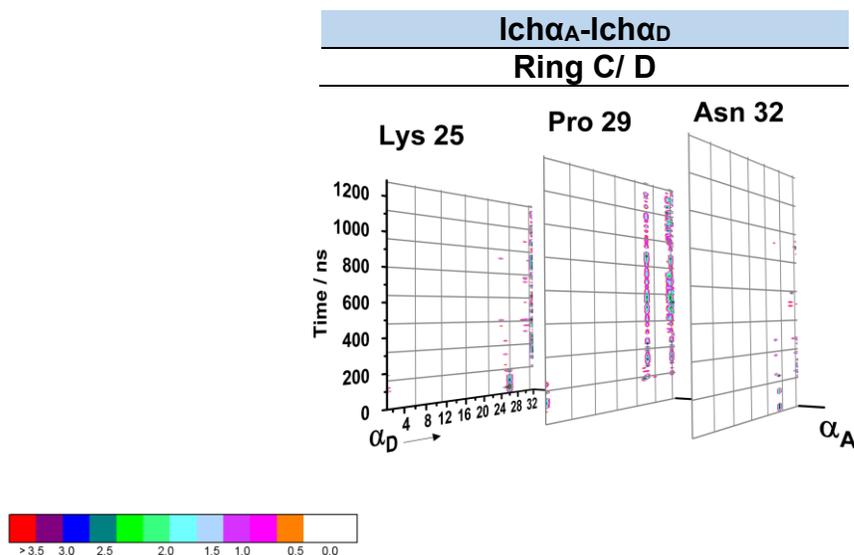


Figure B.2.2.4-A.3: TdCM of the interaction pathway followed between $\text{I}\alpha_A$ and $\text{I}\alpha_D$ peptides. Electrostatic interaction detected between both peptides Lys25-Glu'26/ Lys25-Asn'32.

In the course of the simulation, a hydrophobic approach was detected between $\text{I}\alpha_B$ -Leu14 to $\text{I}\alpha_D$ -Leu14, Gly15, Met28 and Pro29. Furthermore, due to proximity some contacts to Asn16, Asn17, and Asn32 (figure B.2.2.4-A.4) were also noticed. While some H-bonds involving $\text{I}\alpha_B$ -Tyr19 with $\text{I}\alpha_D$ -Lys12, and Asn16 were observed and stabilized by the proximal contacts with $\text{I}\alpha_D$ -Pro13, Leu14, Gly15, Asn16. Due to the same cause, Asn32 remained in continuous contact to the N_I -hinge region, as well as the residues Lan11-21, Mlan22-27, Val23, and Pro30 of $\text{I}\alpha_D$. (see figure B.2.2.4-A.4).

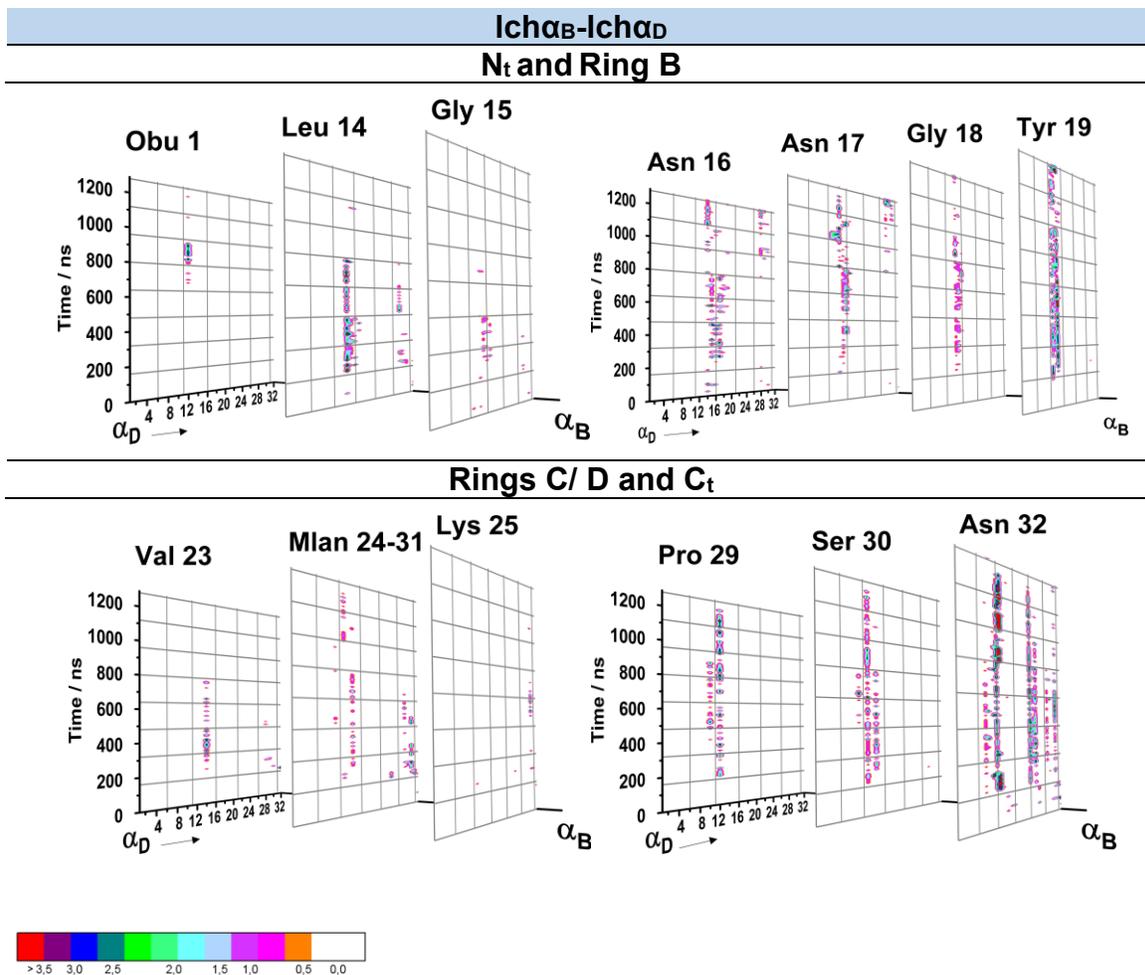


Figure B.2.2.4-A.4: TdCM of the interaction pathway followed between Ich α _B and Ich α _D peptides. Hydrophobic interaction stabilized the peptides.

The hydrophobic interaction produced between the Ich α _C-Ich α _D is mainly located between the Ich α _C-Ile2 and Ile4 to Ich α _D-Mlan3-7 and Mlan24-31, however, some influences of the surrounding residues were also noticed. Additionally, three more active regions were found: i) Ich α _C-Pro13 to Ich α _D-Mlan3-7, Dhb6 and N_t-hinge region; ii) Ich α _C-Gly15, Asn16, Asn17 and Gly18 to Ich α _D-Ile9 and iii) Ich α _C-Tyr19 and Leu20 which are significantly in contact with the Mlan3-7 and Mlan24-31 (see figure B.2.2.4-A.5).

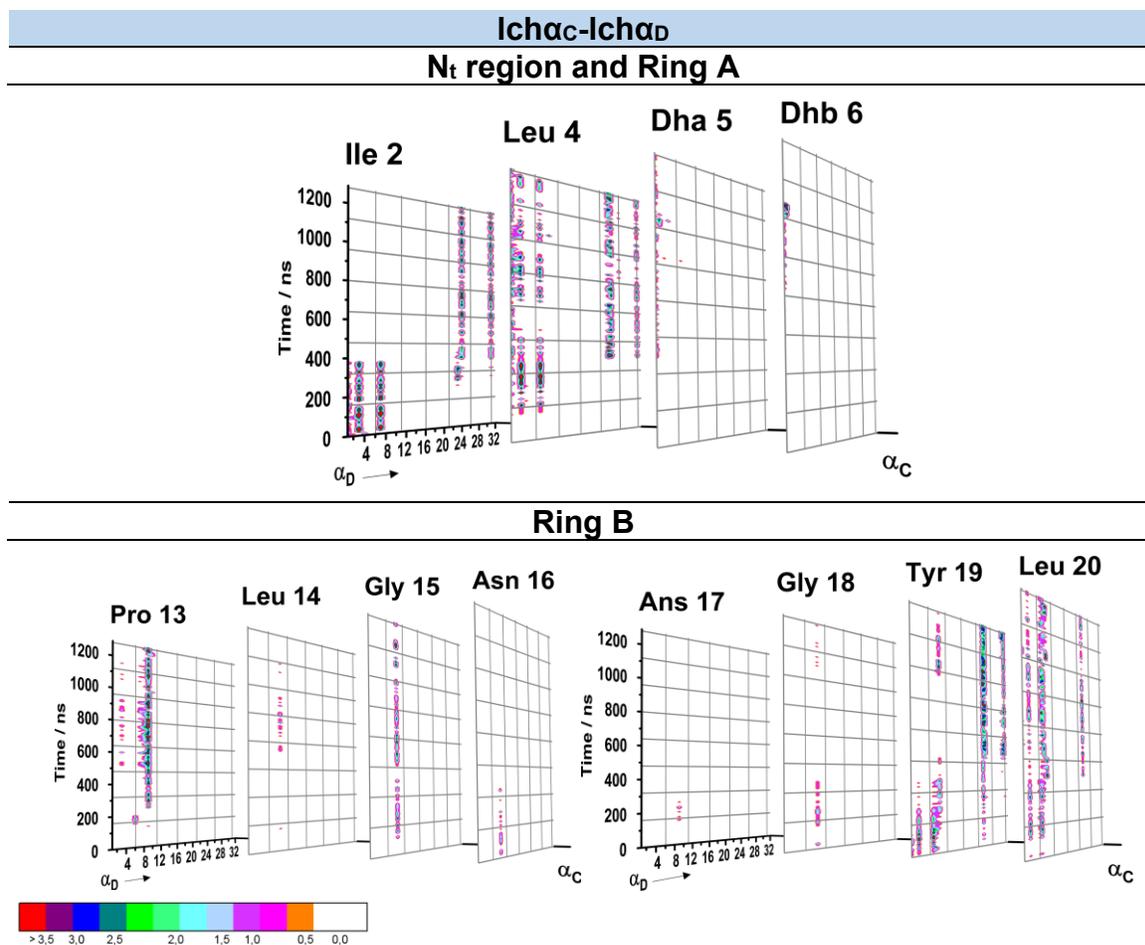


Figure B.2.2.4-A.5: TdCM of the interaction pathway followed between Ich α C and Ich α D peptides.

AG- I: Interactions Summary Ich α -Icha

The prevalent interaction pattern produced between AG-I-Ich α -components involved mainly the rings B and C/ D. If the interaction is produced through rings C/ D, the interpeptide hydrophobic contact is maintained until the end of the simulation. While the inclusion of the ring B is employed during the reorientation stage. However, the rings C/ D may also be involved in the interaction with lipid II, as it was detected in the case of Ich α C (figure B.2.2.4-A.6)

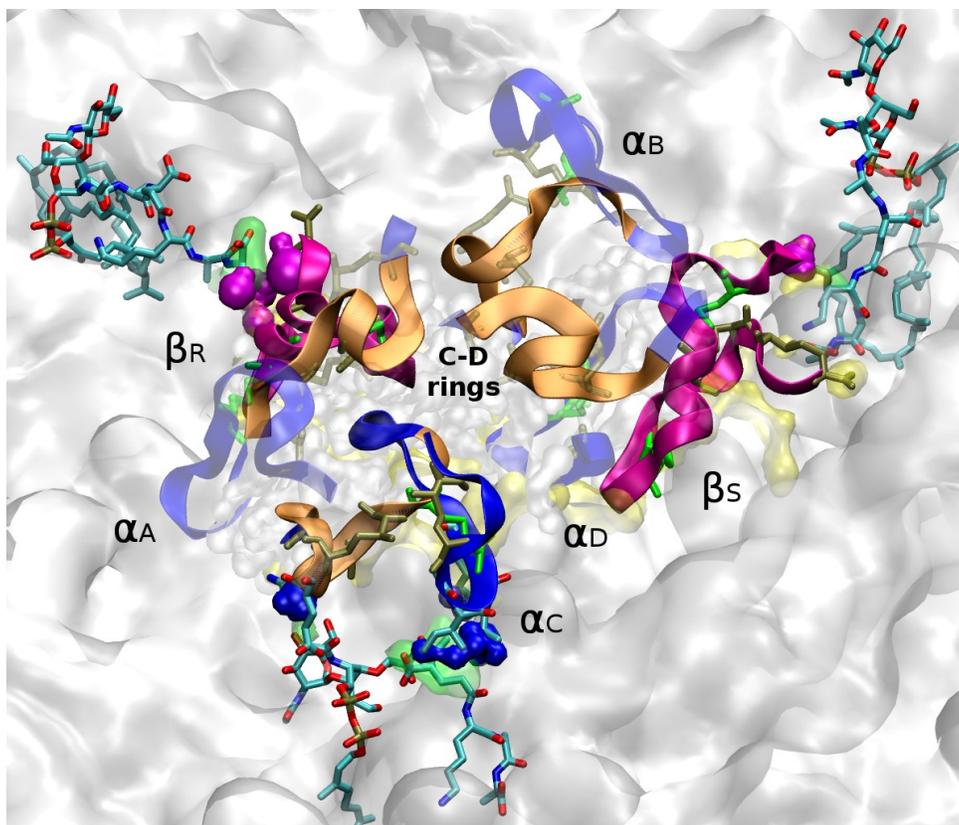


Figure B.2.2.4-A.6: Graphical representation of the 3D-arrangement of the AG-I in the surrounding of lipid II molecules. AG-I is composed by 4 Icha (rings C/D in α -components are highlighted with orange color) and 2 Ich β peptides. A single stable lipid II interaction was detected. Additionally, two lipid II molecules approached the aggregate through the recruiting β -role.

AG-I: Ich β -Ich β . A weak and random (less than 30 ns) hydrophobic contact between the Leu22 residues and Ich β_R -Lys27 to Ich β_S -Obu1 (10 ns) were noticed (data not shown). The spatial arrangement of the AG-I hindered any possible contact between AG-I-Ich β -peptides (see figure B.2.2.4-A.6).

AG-I: Ich α -Ich β . During the simulation time, several interactions between the α/β -pair of the peptides were also observed. According to the results, despite that other regions of the Icha-peptides confirmed their contribution to the stabilization, the role of the Icha-ring B was of outstanding importance. It is feasible to suggest that the ring B of the Icha-peptides is used as a *rudder* to drive the Ich β -peptides to the desired spatial arrangement i) contributing to the stability of the Icha-lipid II complex through its anchorage role observed and ii) by generating additional pore perturbations.

8. Bacteriocins

For example, the hydrophobic interactions were detected between $Ich\alpha_A$ - $Ich\beta_R$. Concisely, as it can be observed in the figure B.2.2.4-A.7 the hydrophobic interactions of: i) $Ich\alpha_A$ -Leu20 to $Ich\beta_R$ -Dhb17 and Val18, ii) $Ich\alpha_A$ -Val23 to $Ich\beta$ -Val18, iii) $Ich\alpha_A$ -Met28 to $Ich\beta_R$ -Ala20 due to proximity stabilized by contact with Ser21, iv) $Ich\alpha_A$ -Ser30 to $Ich\beta_R$ -Ring B (mainly Lan19-23 and Ala20) retained the peptides in continuous interaction until the end of the simulation.

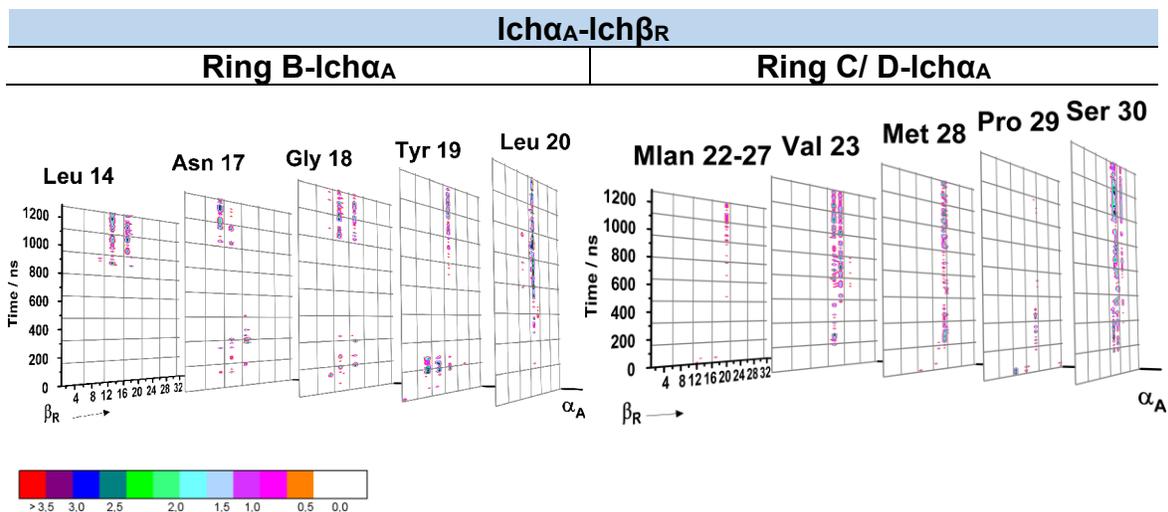


Figure B.2.2.4-A.7: TdCM of the interaction pathway followed between $Ich\alpha_A$ and $Ich\beta_R$ peptides. Mainly produced by hydrophobic interpeptide interactions.

Regarding the interaction between $Ich\alpha_B$ and $Ich\beta_R$, the scenario resulted similar. $Ich\alpha_B$ -Pro29 to $Ich\beta_R$ -Lan19-23, ring C and Arg31 also contributed to the stability (figure B.2.2.4-A.8). Additionally, some H-bonds between the rings C/ D of $Ich\alpha_B$, highlighting the role of $Ich\alpha_B$ -Lys25 which leads to a stable interaction to $Ich\beta_R$ -Ring B (Ser21) were observed.

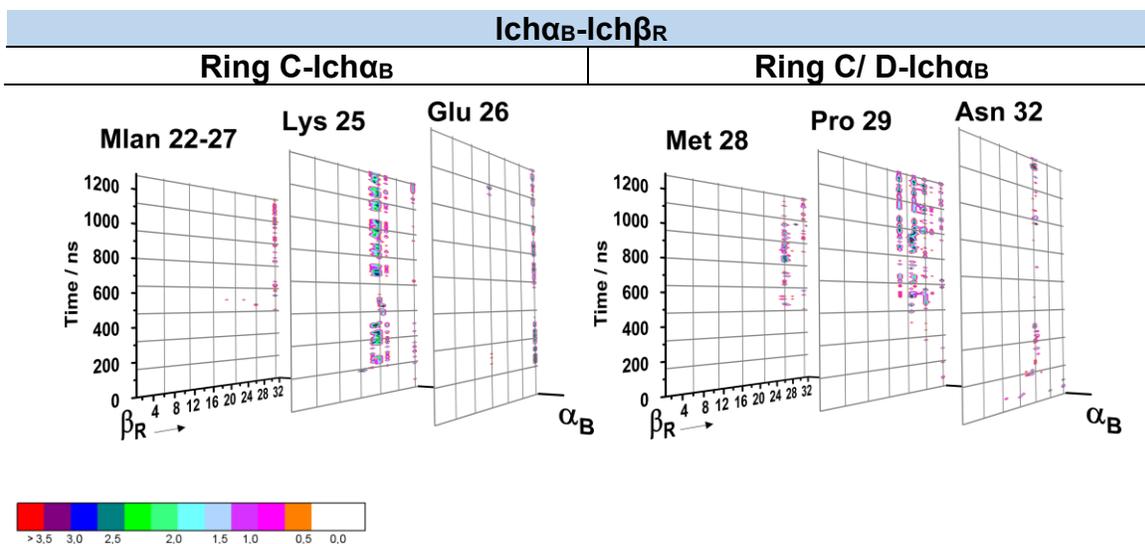


Figure B.2.2.4-A.8: TdCM of the interaction pathway followed between Ich α _B and Ich β _R peptides. Mainly produced by electrostatic and hydrophobic interpeptide interactions.

The prevalent interaction found between Ich α _C and Ich β _R is headed by Ich α _C- Asn16 and its neighbouring residues to the Ich β _R- Ala14, Val18, ring B and Pro24 (figure B.2.2.4-A.9).

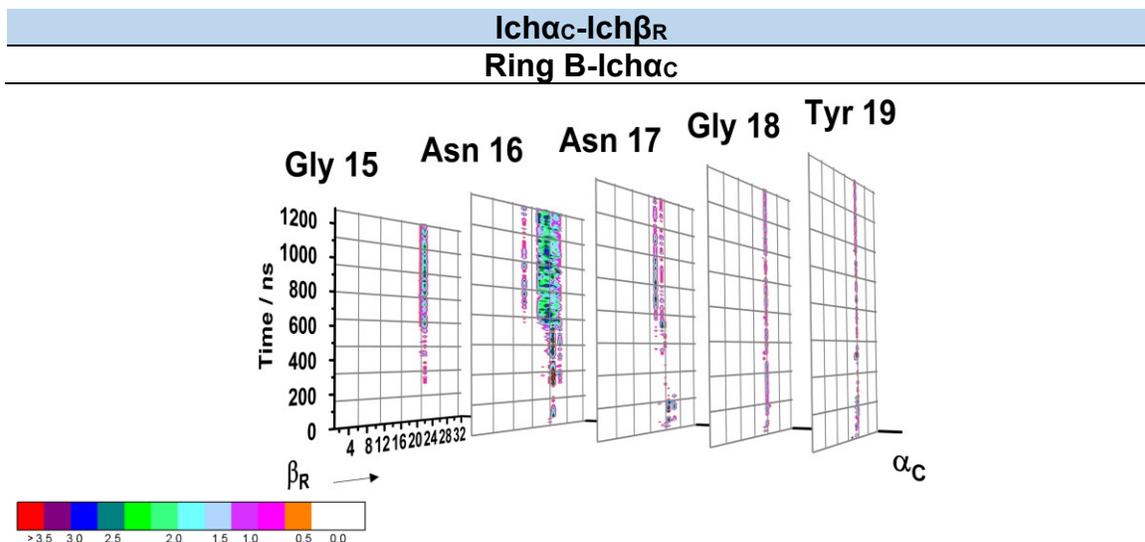


Figure B.2.2.4-A.9: TdCM of the interaction pathway followed between Ich α _C and Ich β _R peptides. H-bonds detected by Asn16.

Concluding with the interaction between Ich α _D and Ich β _R where the hydrophobic residues found in the N_F-hinge region of the Ich α _D-peptide which promote a stable interaction

8. Bacteriocins

to the $\text{Ich}\beta_{\text{R}}$ -ring B-Lan19-23 and Leu22. Furthermore, due to proximity, additional contacts with Ser21 were also noticed (figure B.2.2.4-A.10).

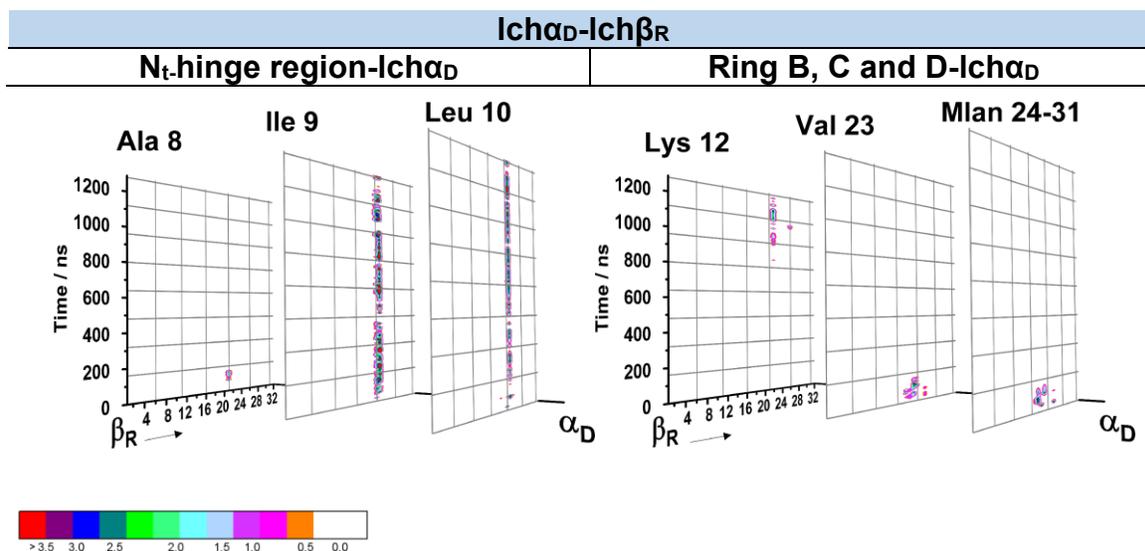


Figure B.2.2.4-A.10: TdCM of the interaction pathway followed between $\text{Ich}\alpha_{\text{D}}$ and $\text{Ich}\beta_{\text{R}}$ peptides. Mainly produced by hydrophobic interpeptide interactions.

In contrast to $\text{Ich}\beta_{\text{R}}$, which showed interactions with all $\text{Ich}\alpha$ -peptides, the $\text{Ich}\beta_{\text{S}}$ constituent was involved just in the interaction with two $\text{Ich}\alpha$ -peptides, $\text{Ich}\alpha_{\text{B}}$ and $\text{Ich}\alpha_{\text{D}}$.

The hydrophobic contacts were found between $\text{Ich}\alpha_{\text{B}}$ and $\text{Ich}\beta_{\text{S}}$, mainly originated between $\text{Ich}\alpha_{\text{B}}$ -Ile 2 to $\text{Ich}\beta_{\text{S}}$ -Obu1, Dhb2, Pro3, Dhb5, Dha8 and Trp9 together with potentially stabilizing π - π interactions of $\text{Ich}\alpha_{\text{B}}$ -Tyr19 with $\text{Ich}\beta_{\text{S}}$ -Trp9. Additionally, due to proximity the Dha8 and Thr10 were also affected (figure B.2.2.4-A.11).

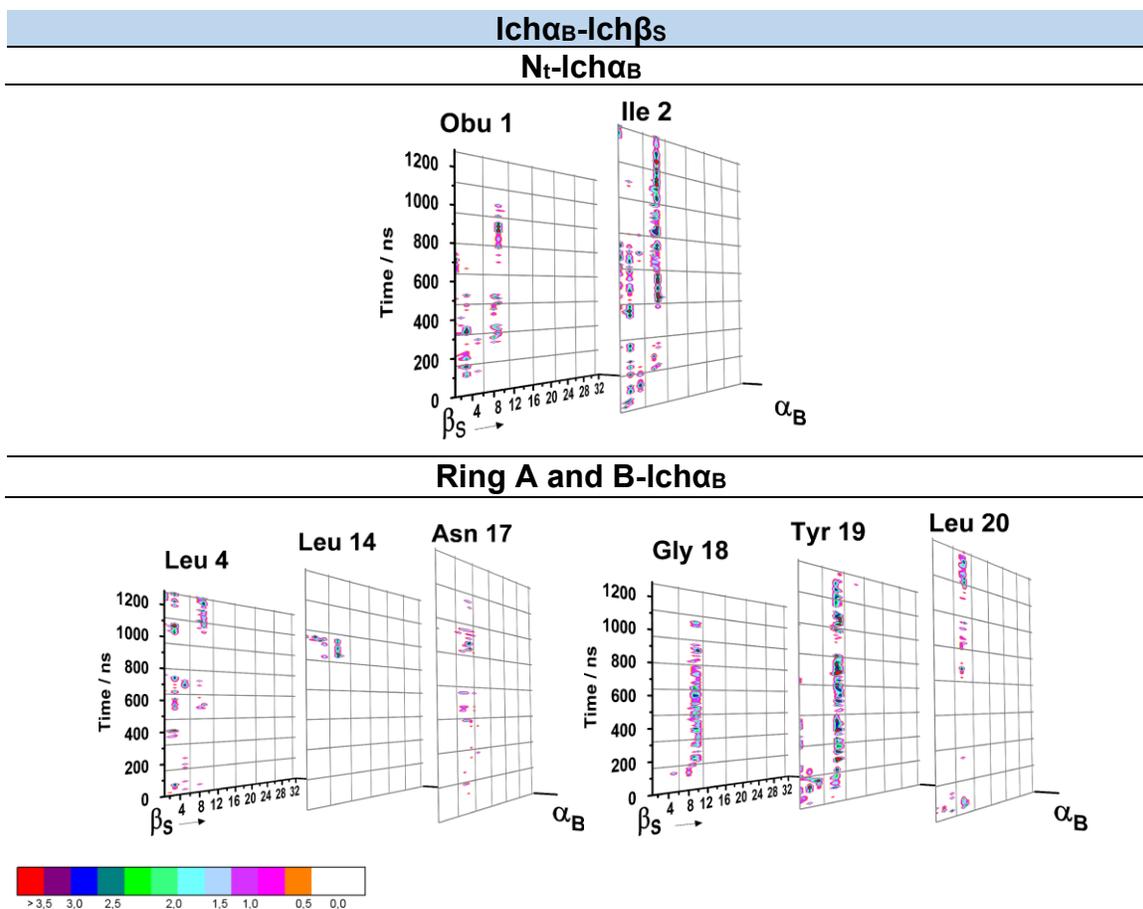


Figure B.2.2.4-A.11: TdCM of the interaction pathway followed between Ich α_B and Ich β_S peptides. Mainly produced by hydrophobic interpeptide interactions.

The hydrophobic interaction detected between Ich α_D -Ich β_S is mostly upheld from the beginning of the simulation through the following interactions i) Ich α_D -Dhb6 to Ich β_S -Val18, Lan19-23 and Ala20, ii) Ich α_D -Ile9 to Ich β_S -Ala20, Ser21, Leu22, iii) Ich α_D -Pro13 to Ich β_S -Trp9, Thr10, Dhb13, Ala14 and Val18, iv) Ich α_D -Leu20 to Ich β_S -Dhb13, Dhb17, Val18. After the approaching phase, the interaction involving Ich α_D -Asn16 to Ich β_S -Lan7-11, Trp9, Thr10, Ile12 and Dhb13 intensified the contacts mentioned before. To those major interactions, it is mandatory to remark the contribution of the proximal amino acids Ich α_D -Leu10, Lan11-21, Lys12, Leu14 and Gly15 (figure B.2.2.4-A.12).

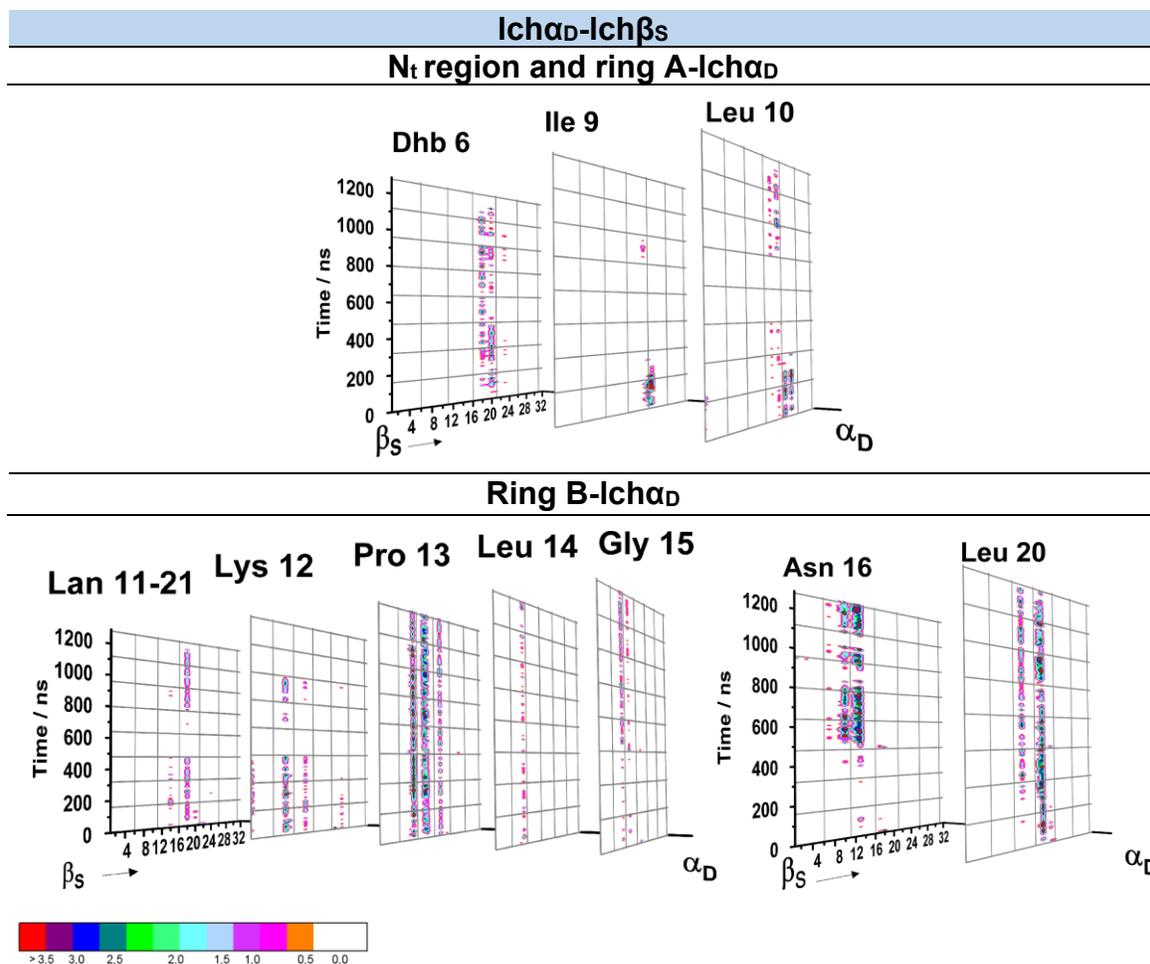


Figure B.2.2.4-A.12: TdCM of the interaction pathway followed between Ich α _D and Ich β _S peptides. Mainly produced by hydrophobic interpeptide interactions.

AG- I: Interactions Summary Ich α -Ich β

Almost all Ich α -components employ ring B as a rudder to establish hydrophobic interactions with the α -helix region and the ring B of the Ich β -peptides. Additionally, due to the different spatial arrangement of the peptides in the course of the simulation, a variety of interpeptide supportive interactions was also noticed. Notoriously, the interactions involving the Ich α -N_t-hinge region and rings C/ D to the Ich β -ring B and α -helix region, or in a minor extent, the interaction involving the Ich α -N_t or ring A.

B.2.2.4-B Interactions Inter-AG-II and AG-III Peptides

AG-II: Icha-Icha. The interactions detected between the following pairs of Icha-peptides: Icha_E-Icha_F, Icha_E-Icha_H, Icha_E-Icha_G, Icha_E-Icha_J have been classified as transitory interactions mainly produced by sporadic contacts. The pairs Icha_G-Icha_J and Icha_F-Icha_H were in a random contact at the last stage of the simulation and from the beginning until the end, respectively.

The contact detected between Icha_E-Icha_F took place during the approaching phase. The hydrophobic interaction was mainly produced between the Icha_E-N_t region and the Icha_F-Leu4, Dha5 and Dhb6. Simultaneously, Icha_E-peptide was interacting involving its ring B (residues Asn16, Asn17, Gly18 and Tyr19) with the ring C/ D of its analogue Icha_G. Additionally, during the first two stages (approaching and reorientation), Icha_E-Lys12 promoted H-bonds with the Icha_J-ring B region (residues Asn16 Asn17) generating the forthcoming to residues Leu14, Gly15 and Gly18. During the stabilization stage, some random hydrophobic contacts of Icha_E-Leu4 and Dha5 to the proximal Icha_H-N_t region and Leu4 were also detected (see figure B.2.2.4-B.1).

8. Bacteriocins

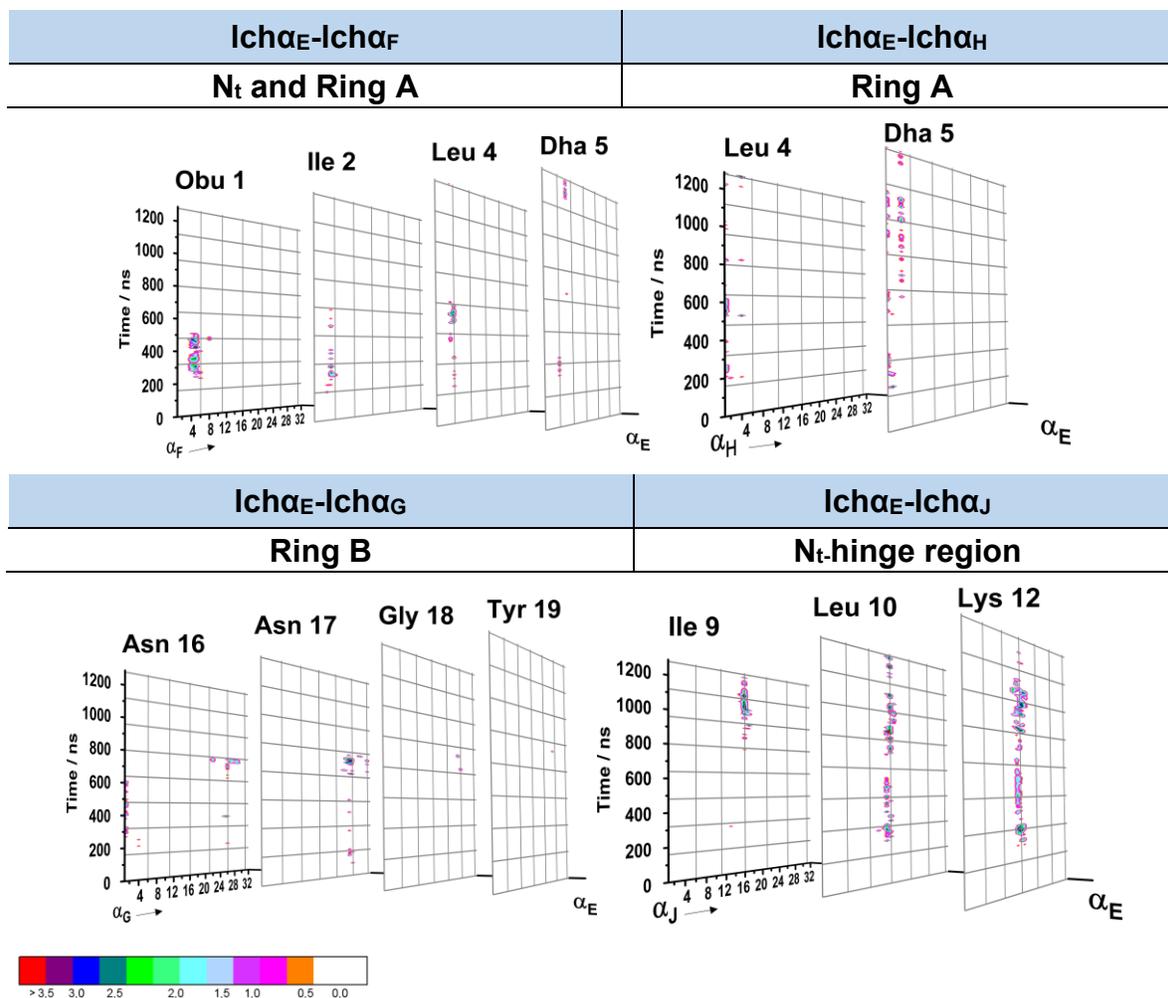


Figure B.2.2.4-B.1: TdCM of the interaction pathway followed between Icha_E with the Icha partners α_F , α_H , α_G , α_J . Mainly produced by hydrophobic interpeptide interactions.

Contrary to the observations described above, at the beginning of the stabilization stage Icha_G-Icha_J were spatially closer allowing the first hydrophobic interactions α_E between each other. The Icha_G-ring B (residues Gly18 and Tyr19) interacted with the Icha_J-ring D (residues Met28, Pro29 and Ser30) forming hydrophobic interactions and H-bonds respectively. This first approach allowed the movement of the Icha_J-N_t region towards Icha_G-ring B (residues Asn16 and Asn17) stabilized by H-bonds (see figure B.2.2.4-B.2).

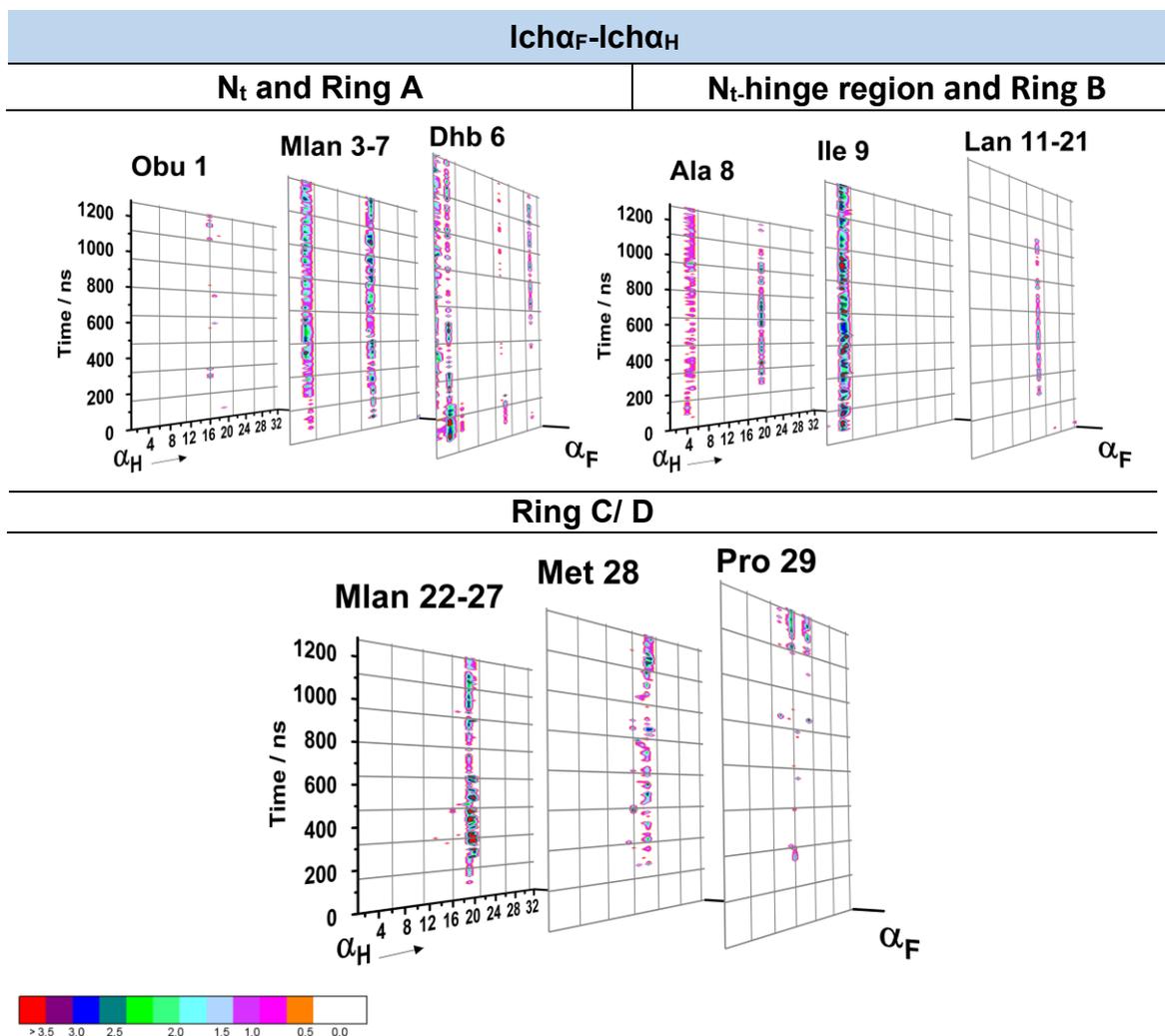


Figure B.2.2.4-B.3: TdCM of the interaction pathway followed between Icha_F and Icha_H peptides. Mainly produced by hydrophobic interpeptide interactions.

AG-II: Interactions Summary Icha-Icha

Analogous to the Icha-Icha interactions observed in AG-I, the stabilizing hydrophobic interactions between two Icha-partners in AG-II are ascribed to the ring B of one component to the rings C/ D of the other. Additionally, given an optimal arrangement, the hydrophobic interactions between N_t-hinge region and ring A also contributed to the stabilization among Icha-peptides.

AG-II: Ich β -Ich β . Due to the high number of Ich β -constituents in the aggregate (4 Ich β peptides), it was possible to study the interaction pattern involving two Ich β -analogues. For example, during the first 500 ns of simulation, the approaching of the Ich β_Q - α -helix region to the Ich β_N -N $_t$ area could be observed. This approach was mainly driven by the hydrophobic residue Val18 combined with the dehydrated residues 13 and 17. Additionally, during the reorientation stage, where the N $_t$ region together with the ring A of both peptides began to interact (figure B.2.2.4-B.4).

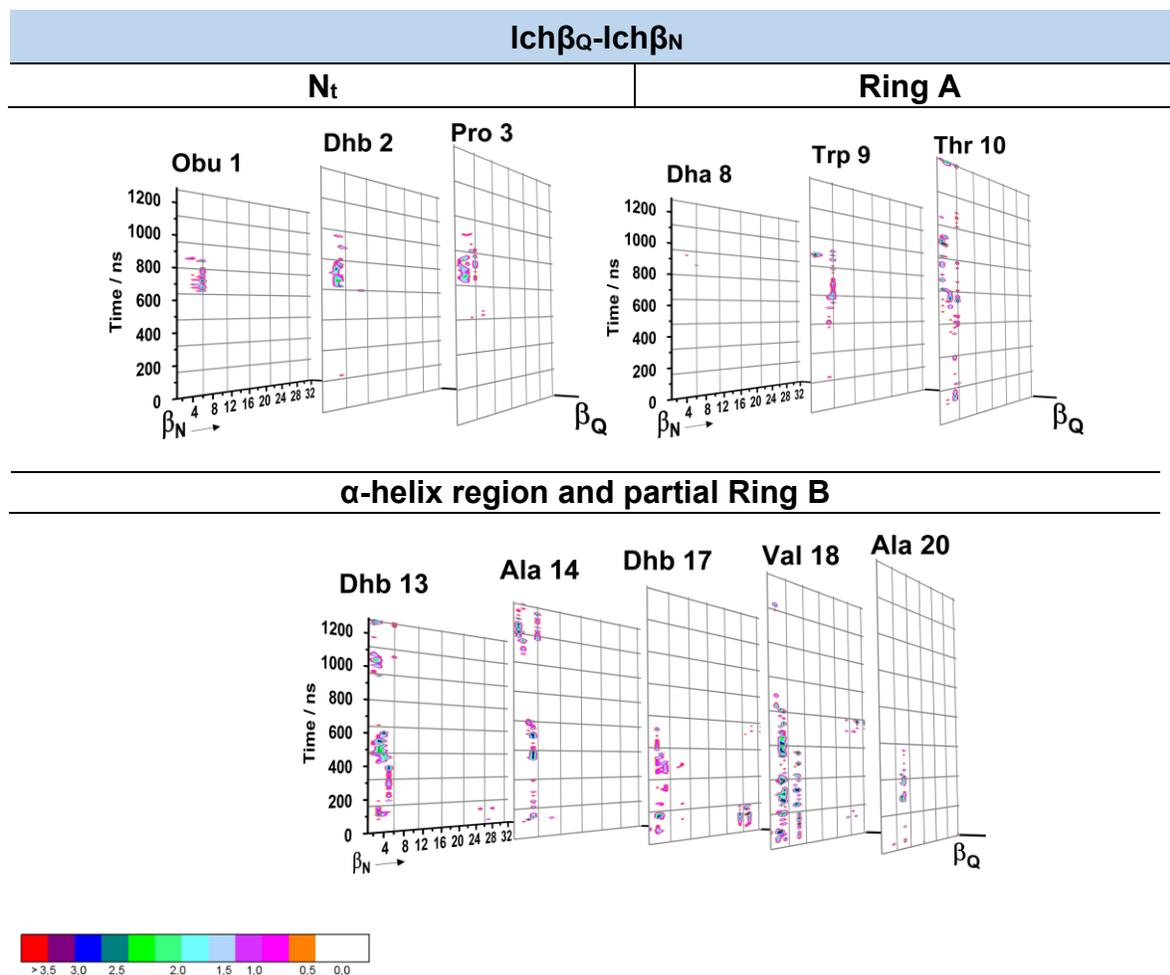


Figure B.2.2.4-B.4: TdCM of the interaction pathway followed between Ich β_Q and Ich β_N peptides. Mainly produced by hydrophobic interpeptide interactions.

Additionally, the Ich β_Q showed some relatively more stable interactions with the Ich β_U peptide. As it can be followed in the figure B.2.2.4-B.5, during the approaching stage the Ich β_Q -Ile12, Val16 and Dhb17 came near to the Ich β_U -ring C, at the same time the C $_t$ regions

8. Bacteriocins

of both peptides initiated an electrostatic interaction, which was sustained until the end of the simulation (figure B.2.2.4-B.5).

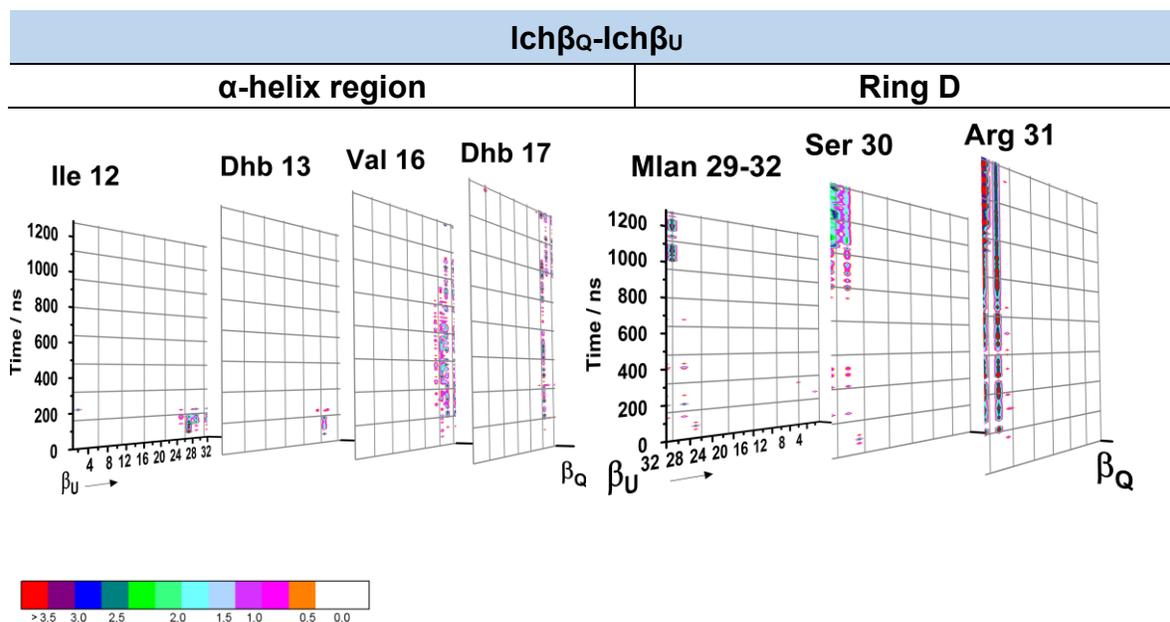


Figure B.2.2.4-B.5: TdCM of the interaction pathway followed between Ichβ_Q and Ichβ_U peptides. Mainly produced by electrostatic interpeptide interactions.

While the C_t region of Ichβ_U remained in contact with Ichβ_Q, the N_t area together with the Ring A stabilized the hydrophobic interaction with the N_t region of Ichβ_X (figure B.2.2.4-B.6).

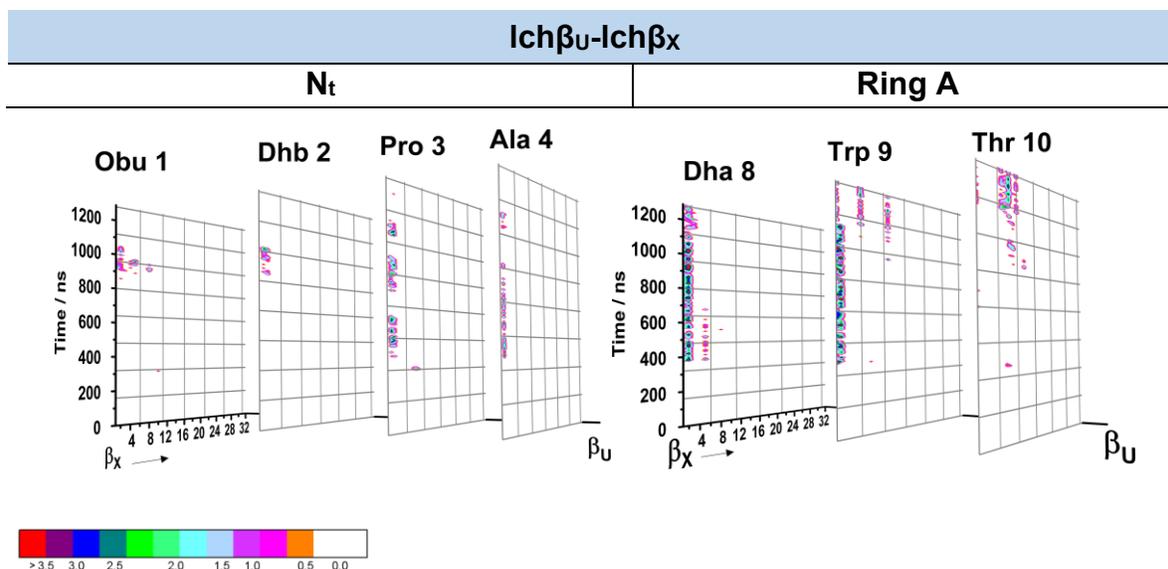


Figure B.2.2.4-B.6: TdCM of the interaction pathway followed between Ich β _U and Ich β _X. Mainly produced by hydrophobic interpeptide interactions.

Ich β _V peptide only interacted with Ich β _N. Responsible of this interaction were Ich β _V-ring A, residues Ile12, Dhb13, Dhb17 and Lys27 and Ich β _N-N_t, Val16, Dhb17 during the approaching phase. In addition, Ich β _V-Trp9, α -region, ring C (Lys27) and ring D (Arg31) to the Ich β _N-ring D caused the stabilization of the aggregate during the reorientation and stabilization stages (figure B.2.2.4-B.7).

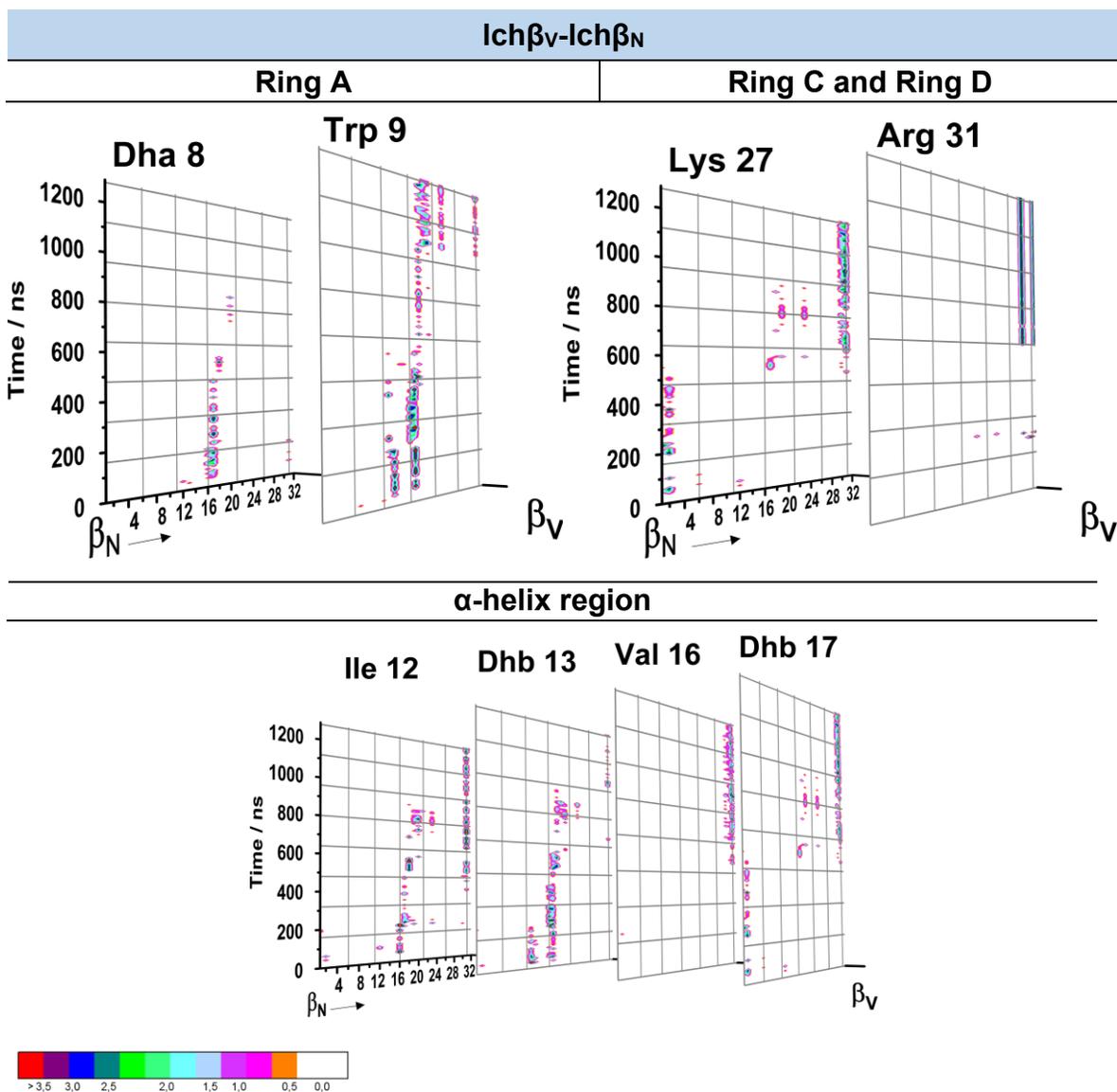


Figure B.2.2.4-B.7: TdCM of the interaction pathway followed between Ich β_V with Ich β_N . Mainly produced by hydrophobic interpeptide interactions.

Lch β_X showed sporadic interactions with Ich β_O . Interestingly, after 175 ns of simulation, Ich β_X peptide started to interact with the adjacent Ich α_J and Ich β_U constituents inducing the disconnection of the Ich β_T and Ich β_O components. Namely these last two Ich β -components in the following will be referred to as AG-III. These results prompt towards the unlikeliness of a Ich β -trimer formation (figure B.2.2.4-B.8).

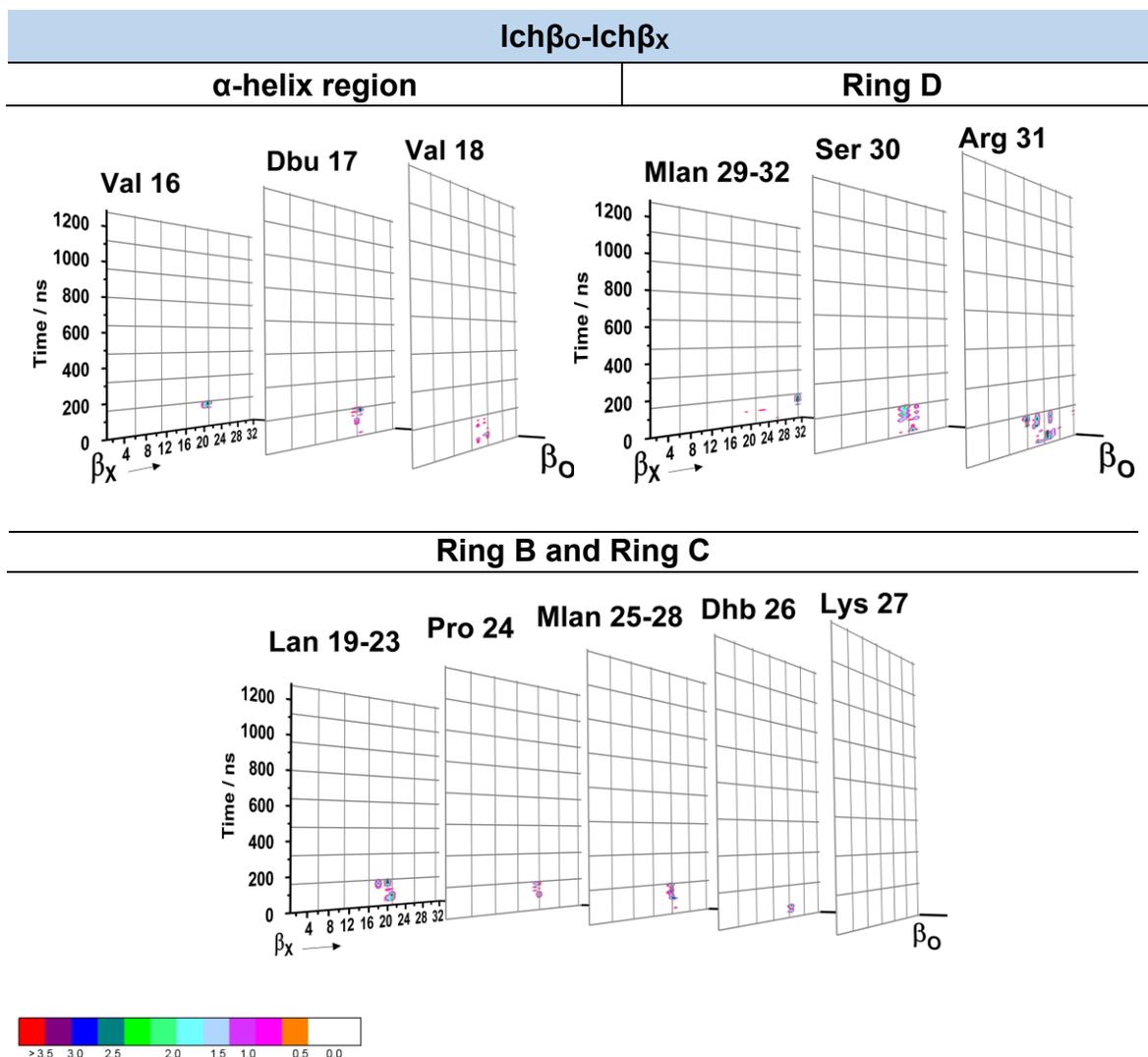


Figure B.2.2.4-B.8: TdCM of the interaction pathway followed between Ich β _x with Ich β ₀. Random H-bond interpeptide contact.

AG-III: Ich β -Ich β . The AG-III consisted of two Ich β peptides, Ich β _T and Ich β ₀. As it can be observed in the figure B.2.2.4-B.9, both peptides interacted mainly through the Ich β _T-ring C and ring D to the Ich β ₀-residues Val16, Dhb17 and Val18 with some influences of the rings B, C and D. Furthermore, the role of the hydrophobic residues for the stabilization of the dimer was unambiguous, such as the case of Ich β _T-Trp9, Ile12, Val16, Leu22 and Pro24 which over the simulation time remained in specific limited contact with the Ich β ₀-hydrophobic residues found in the first 1st-19th positions of the amino sequence.

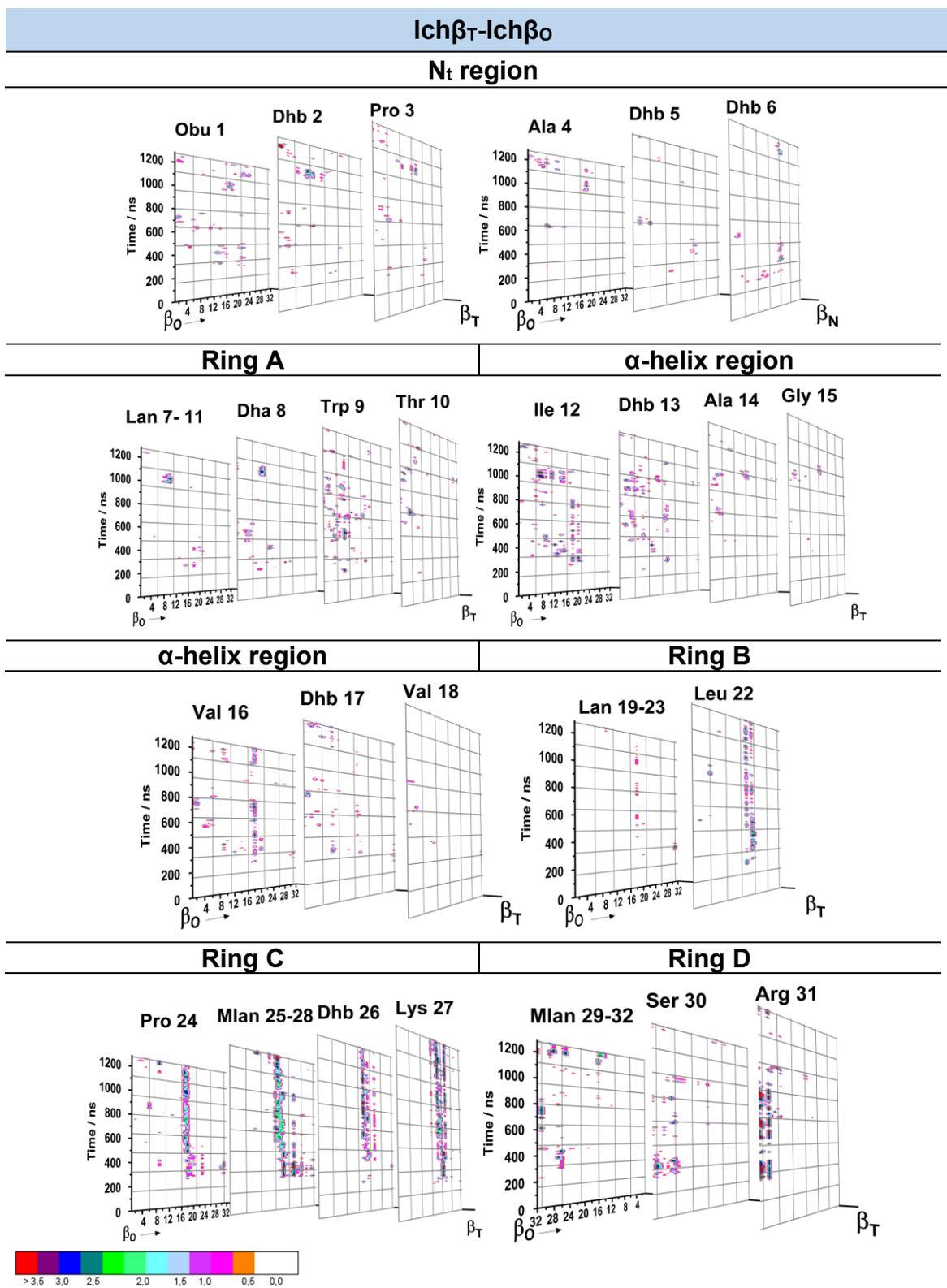


Figure B.2.2.4-B.9: TdCM of the interaction pathway followed between Ich β _O and Ich β _T. Mainly produced by hydrophobic interpeptide interactions.

Interestingly, in AG-III unfolding of the $\text{Lch}\beta$ -peptides was observed (see figure B.2.2.4-B.10). Despite that the $\text{Lch}\beta$ -peptides of the AG-III were in continuous contact almost from the beginning of the simulation, the secondary structure was not preserved until the end of the simulation. It seems that the presence of $\text{Lch}\alpha$ -component/ membrane is required to promote a correct folding of the $\text{Lch}\beta$ -component.

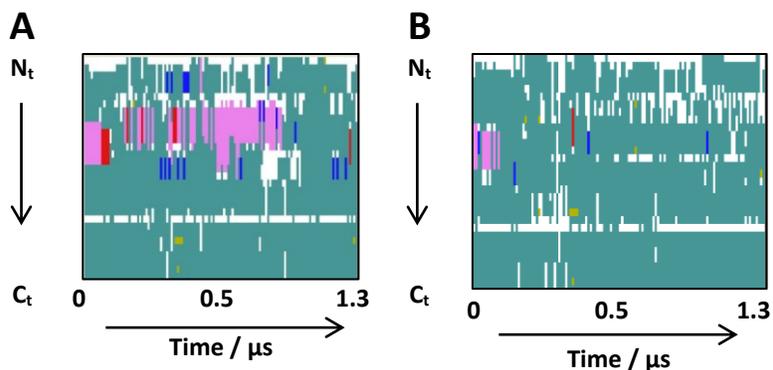


Figure B.2.2.4-B.10: Monitoring secondary structure for $\text{Lch}\beta$ -peptides in AG-III. A- $\text{Lch}\beta_\tau$ B- $\text{Lch}\beta_o$. Color code for the designation of the possible secondary structures: teal (turns), yellow (extended configurations), green (isolated bridges), pink (α -helix), blue (3_{10} helix), red (π -helix), white (coil). Unfolding detected for both peptides.

In the figure B.2.2.4-B.11, the spatial arrangement adopted by the AG-II and AG-III peptides is shown.

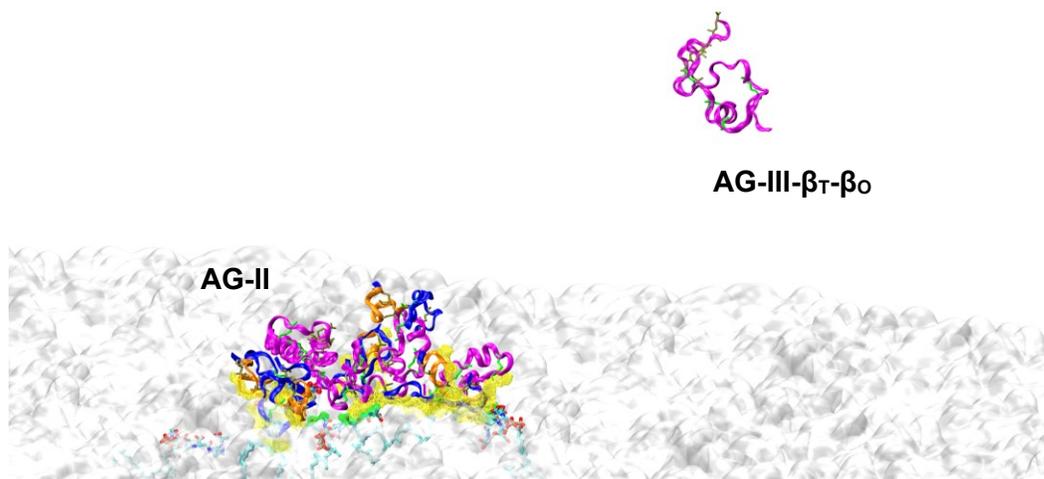


Figure B.2.2.4-B.11: Snapshot of the spatial arrangement adopted by AG-II and AG-III. Blue peptides detone the I_{ch}α-components while I_{ch}β-peptides are displayed in magenta. Ring C/ D of I_{ch}α-components is highlighted in orange. Interactions with lipid II are represented in green while the interaction with the membrane in yellow.

AG-II/ AG-III: Interactions Summary I_{ch}β-I_{ch}β

Simulations showed that dimers are favoured over trimer formation. Dimers are formed through interactions involving the full I_{ch}β-peptide chain. Indeed, the presence of the I_{ch}α-partners/ membrane seems to play a role in the correct folding of I_{ch}β-peptides.

AG-II: $I\text{ch}\alpha\text{-Ich}\beta$. The $I\text{ch}\alpha\text{E}$ peptide was the $I\text{ch}\alpha$ component in AG-II which interacted with the highest number of $I\text{ch}\beta$ -peptides. $I\text{ch}\alpha\text{E}$ -hydrophobic residues placed in the ring A and ring B (Leu4 and Leu20) together with the polar $I\text{ch}\alpha\text{E}$ (Tyr19) contacted mainly to the $I\text{ch}\beta\text{N}$ -hydrophobic N_I -region and ring D generating a combination of hydrophobic interactions and H-bonds (figure B.2.2.4-B.12).

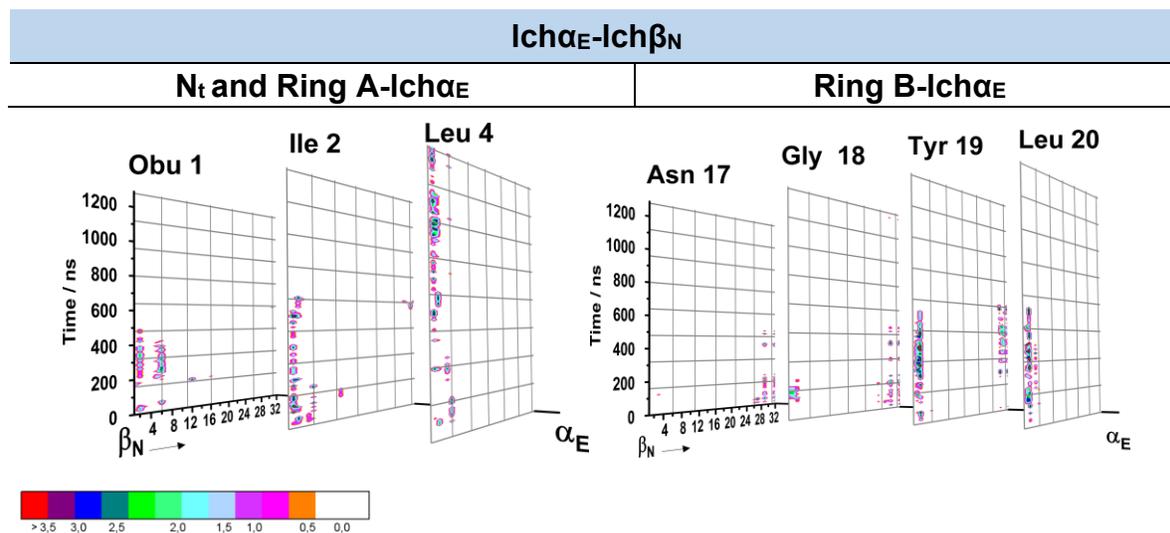


Figure B.2.2.4-B.12: TdCM of the interaction pathway followed between $I\text{ch}\alpha\text{E}$ and $I\text{ch}\beta\text{N}$. Mainly produced by hydrophobic interpeptide interactions.

A similar interaction pattern was observed with $I\text{ch}\beta\text{Q}$ peptide. During the first 500 ns mainly the $I\text{ch}\alpha\text{E}$ -hydrophobic residues placed in the ring B, were the precursors for the stabilization with the $I\text{ch}\beta\text{Q}$ -Ile12 and Val16 which was afterwards stabilized by the interactions produced between the $I\text{ch}\alpha\text{E}$ -Leu4 (Ring A) to $I\text{ch}\beta\text{Q}$ -Thr10 and Dhb13 (α -helix region), $I\text{ch}\alpha\text{E}$ -Ile9 (N_I -hinge region) to the $I\text{ch}\beta\text{Q}$ -Ring A, $I\text{ch}\alpha\text{E}$ -Leu10 (N_I -hinge region) to the $I\text{ch}\beta\text{Q}$ -Ile12 and Val16 (α -helix region) (figure B.2.2.4-B.13).

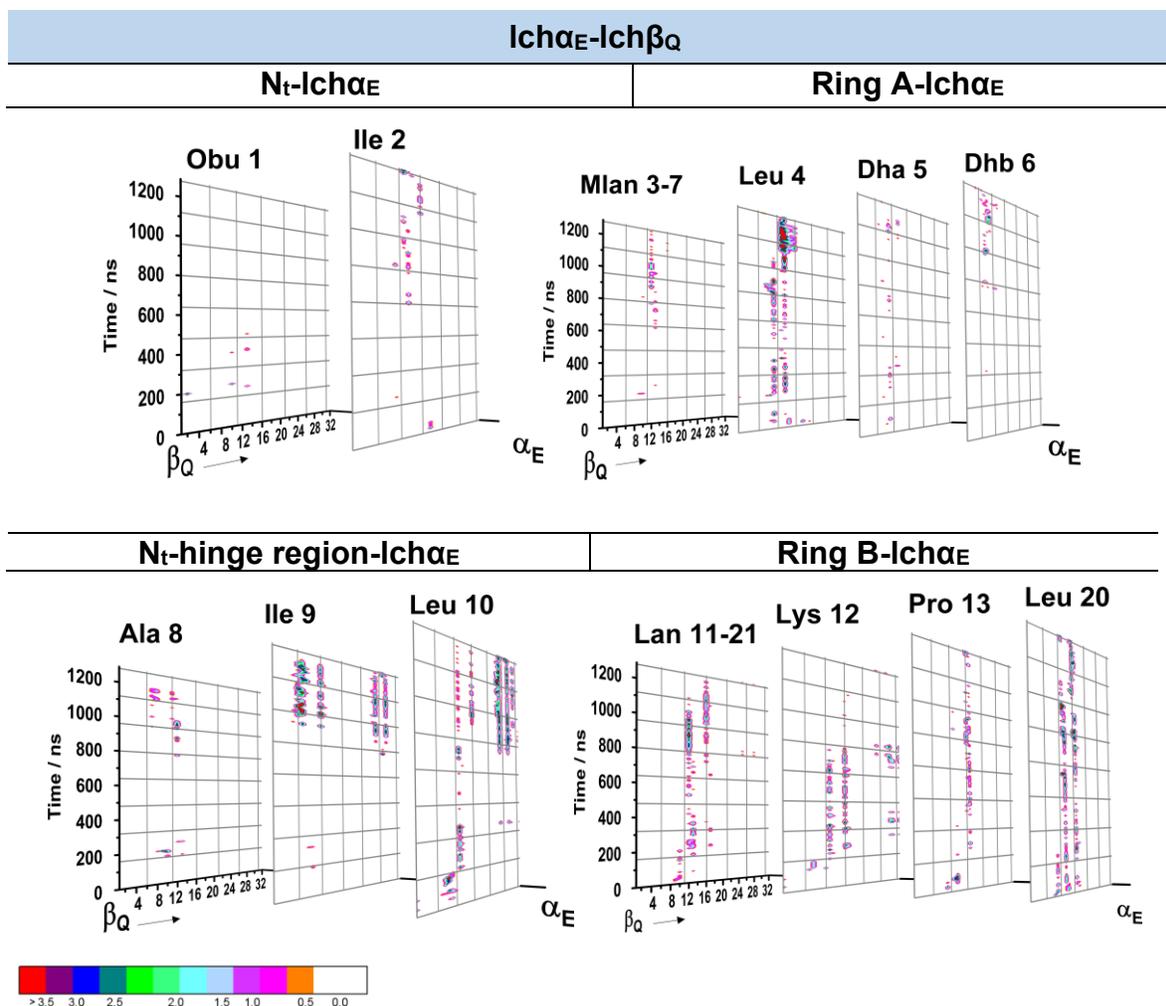


Figure B.2.2.4-B.13: TdCM of the interaction pathway followed between TdCM of Ich α_E and Ich β_Q . Mainly produced by hydrophobic interpeptide interactions.

Interestingly, at the beginning of the simulation and during the reorientation stage Ich α_E -N_t-hinge region showed a strong hydrophobic contact with the Ich β_U -N_t region (residues 1-6). This observation was detected due to the interactions produced during all simulation time through the Ich α_E -ring B to the Ich β_U -N_t region and more pronounced to the ring C and Mlan29-32 (figure B.2.2.4-B.14).

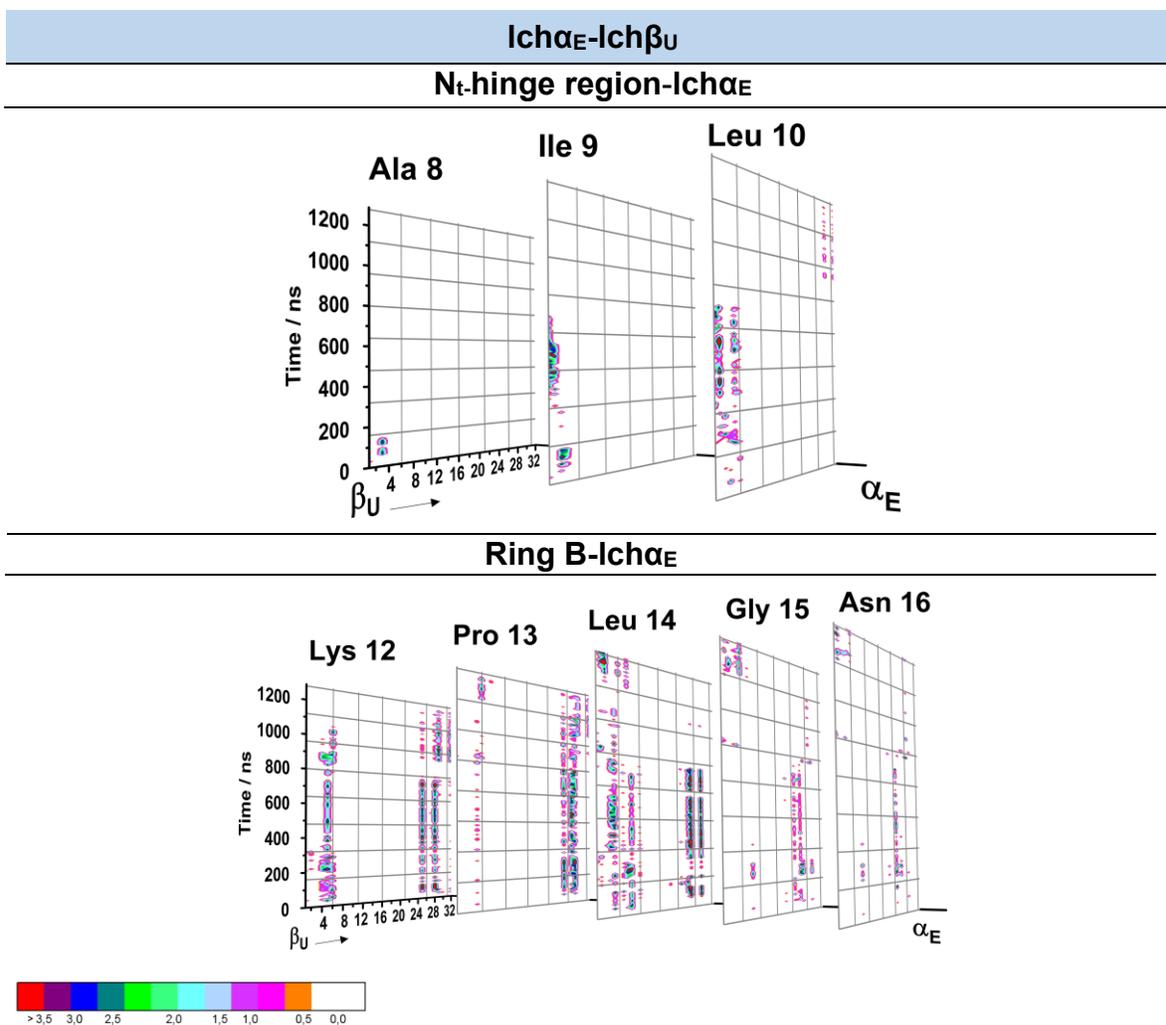


Figure B.2.2.4-B.14: TdCM of the interaction pathway followed between Ich α _E and Ich β _U. Mainly produced by hydrophobic interpeptide interactions.

The behaviour of the Ich α _E in the proximity of Ich β _V was similar. Herein, during the approaching stage weak interactions of the Ich α _E-N_t, Ring B (Tyr 19), ring C/ D (Lys25, Pro29, Asn32) to the Ich β _V- α -helix and C_t regions were detected. The strongest and most stable interactions arose during the stabilization stage involving the Ich α _E-N_t and ring A and specific contributions of the Tyr19 and Met 28 to the Ich β _V-ring B and ring C (figure B.2.2.4-B.15).

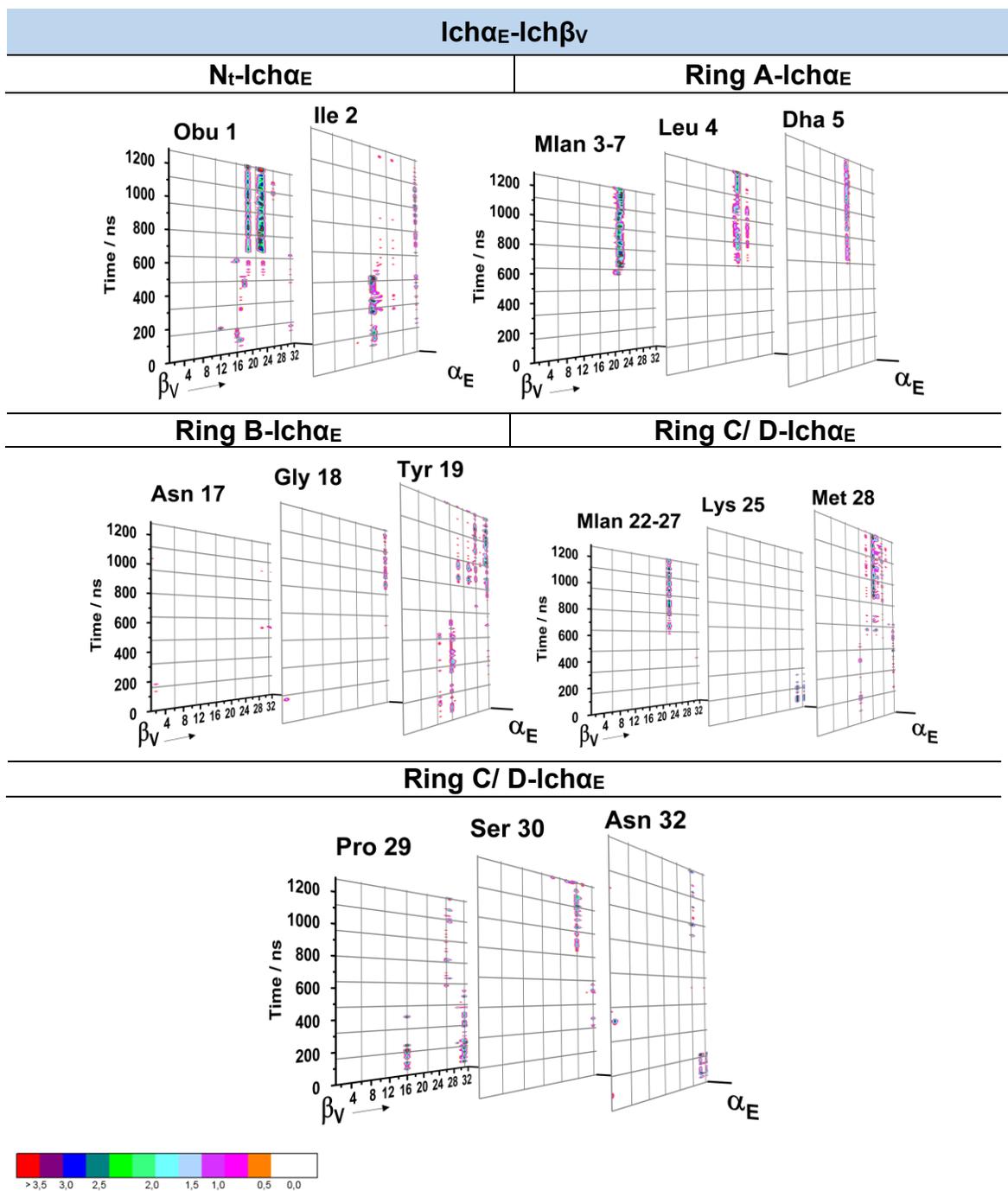


Figure B.2.2.4-B.15: TdCM of the interaction pathway followed between Ich α_E and Ich β_V . Mainly produced by hydrophobic interpeptide interactions.

The Ich α_F component just showed dual affinity for two Ich β -partners. These interactions involved the Ich α_F -N_t and rings A and B to the Ich β_N -N_t and α -helix region until the end of the simulation. Alternatively, Ich α_F was in contact with Ich β_V through hydrophobic

interactions. The $\text{Ich}\alpha_{\text{F}}\text{-N}_t$, ring A and ring B approached to the $\text{Ich}\beta_{\text{V}}\text{-ring A}$, α -helix region and ring B (figures B.2.2.4-B.16-17).

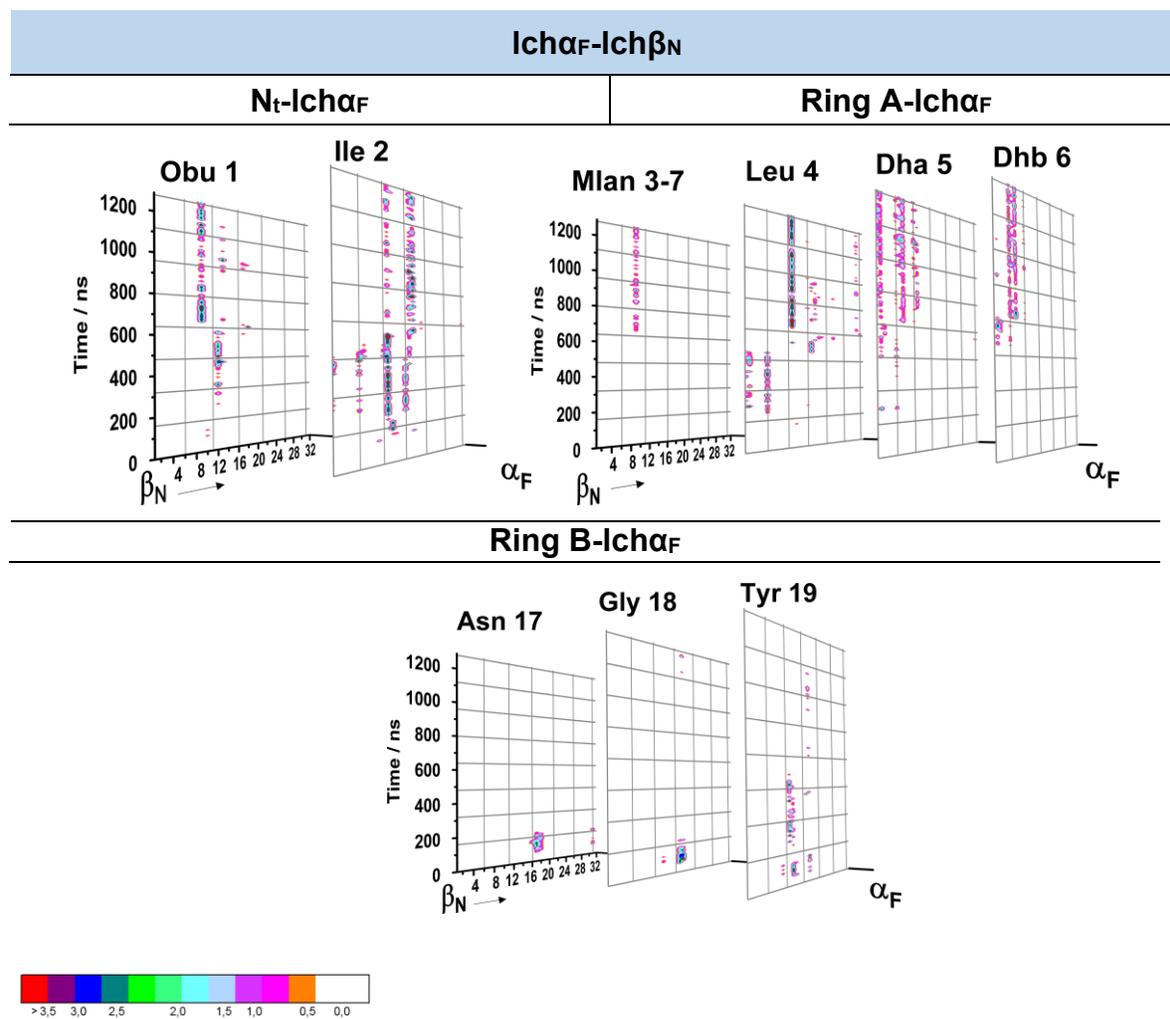


Figure B.2.2.4-B.16: TdCM of the interaction pathway followed between $\text{Ich}\alpha_{\text{F}}$ and $\text{Ich}\beta_{\text{N}}$. Mainly produced by hydrophobic interpeptide interactions.

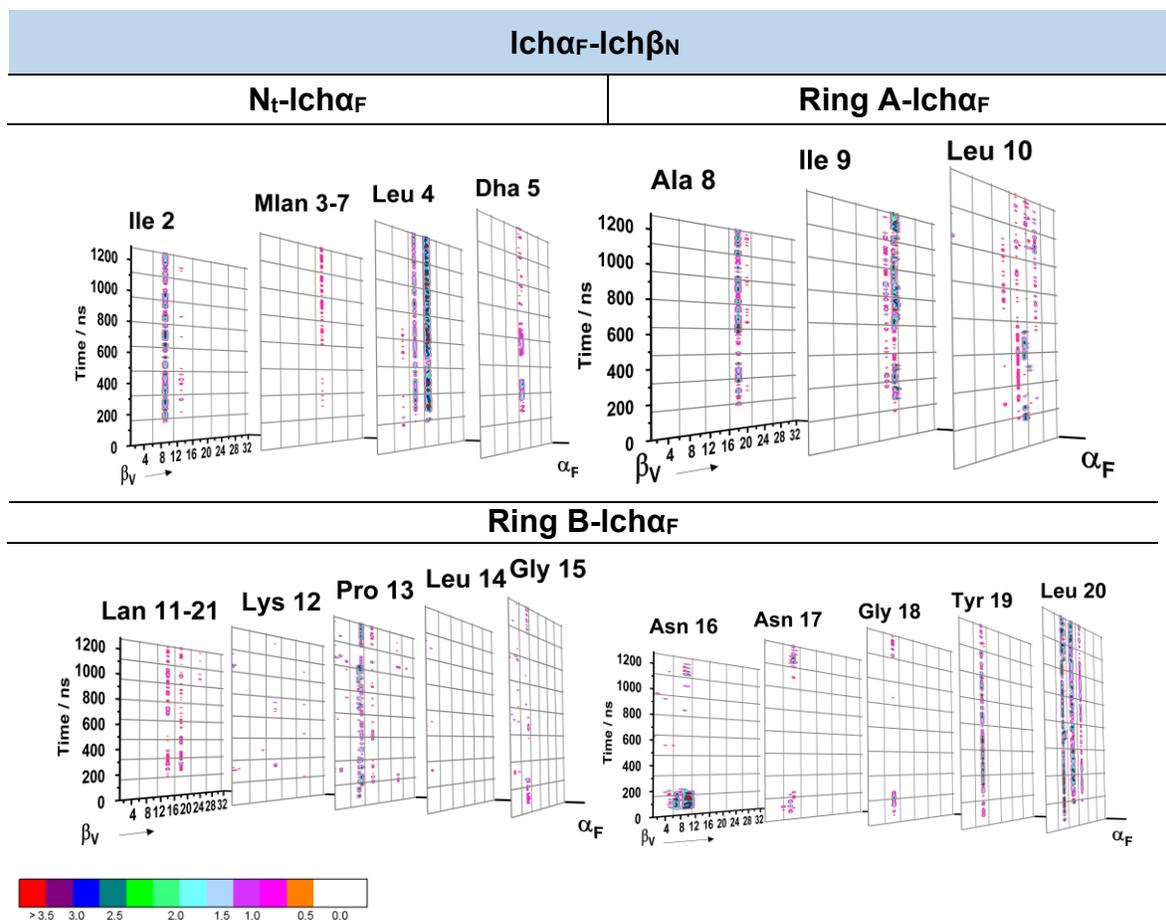


Figure B.2.2.4-B.17: TdCM of the interaction pathway followed between Icha_F and Ich_{βV}. Mainly produced by hydrophobic interpeptide interactions.

The Icha_G peptide carried out interactions with three Ich_β-partners. The Icha_G-Ich_{βN} contacts can be considered as random (figure B.2.2.4-B.18). In contrast, the scenario obtained for the pair Icha_G-Ich_{βU} was completely different. The Icha_G employed the entire sequence to keep the interaction with Ich_{βU}. Two different scenarios could be perfectly distinguished: initially, the Icha_G-N_t-hinge region, ring B and rings C/ D interacted to the Ich_{βU}-α-helix and ring B region while at the last stage of the simulation the interactions of the Icha_G- N_t and ring A to the Ich_{βU}-ring A were more pronounced (figure B.2.2.4-B.19). At the last stage of the simulation, Icha_G generated hydrophobic interactions with Ich_{βX} (figure B.2.2.4-B.20). As it can be observed in the figure B.2.2.4-B.21, these interactions were localized on the Icha_G- rings A and B to the Ich_β-residues Ile12, Dhb13, Dhb17, Val18.

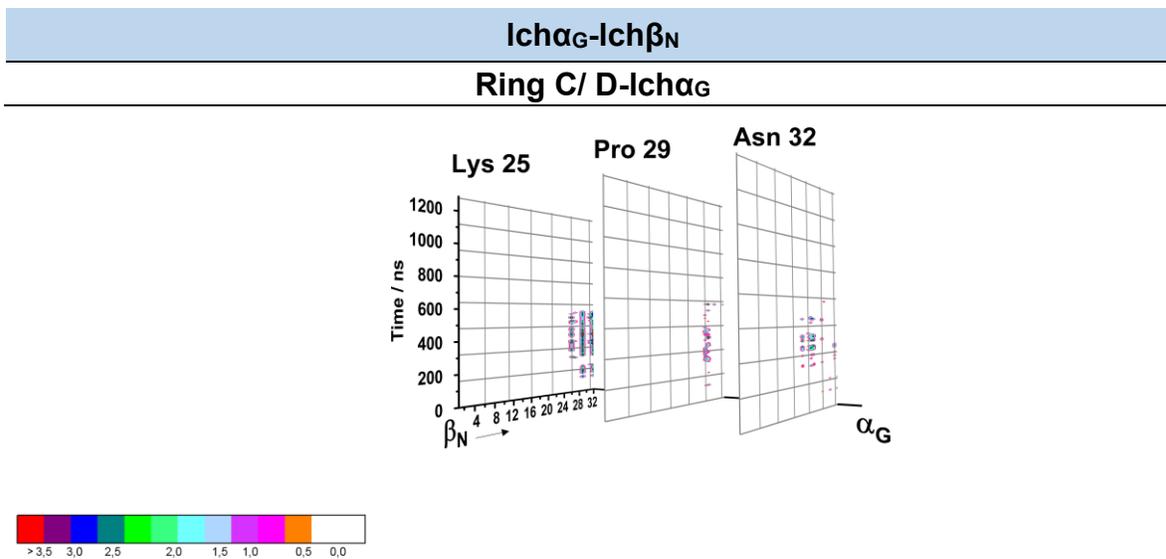


Figure B.2.2.4-B.18: TdCM of the interaction pathway followed between Ich α _G and Ich β _N. Mainly produced by electrostatic interpeptide interactions.

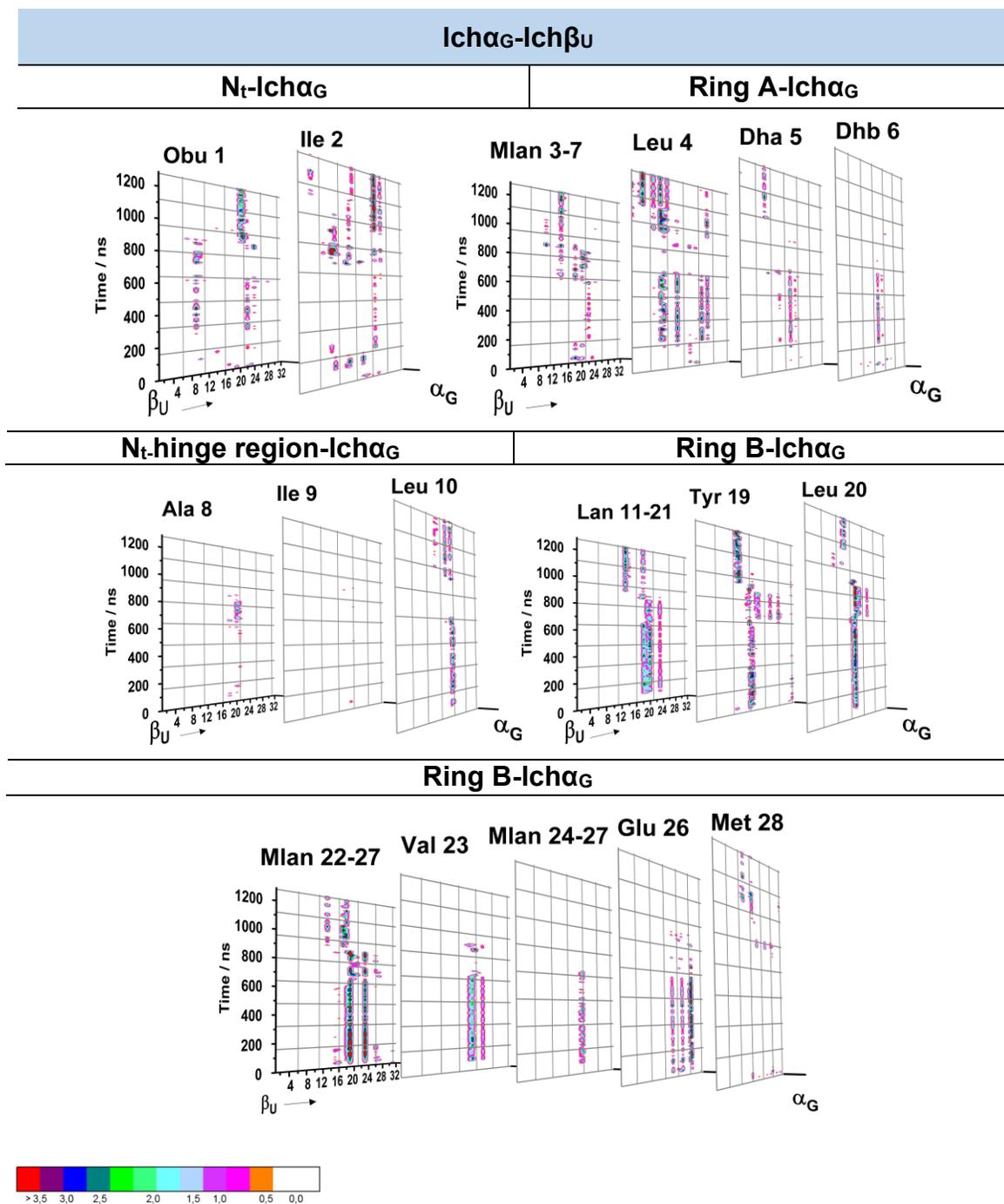


Figure B.2.2.4-B.19: TdCM of the interaction pathway followed between Ich α _G and Ich β _U. Mainly produced by hydrophobic interpeptide interactions.

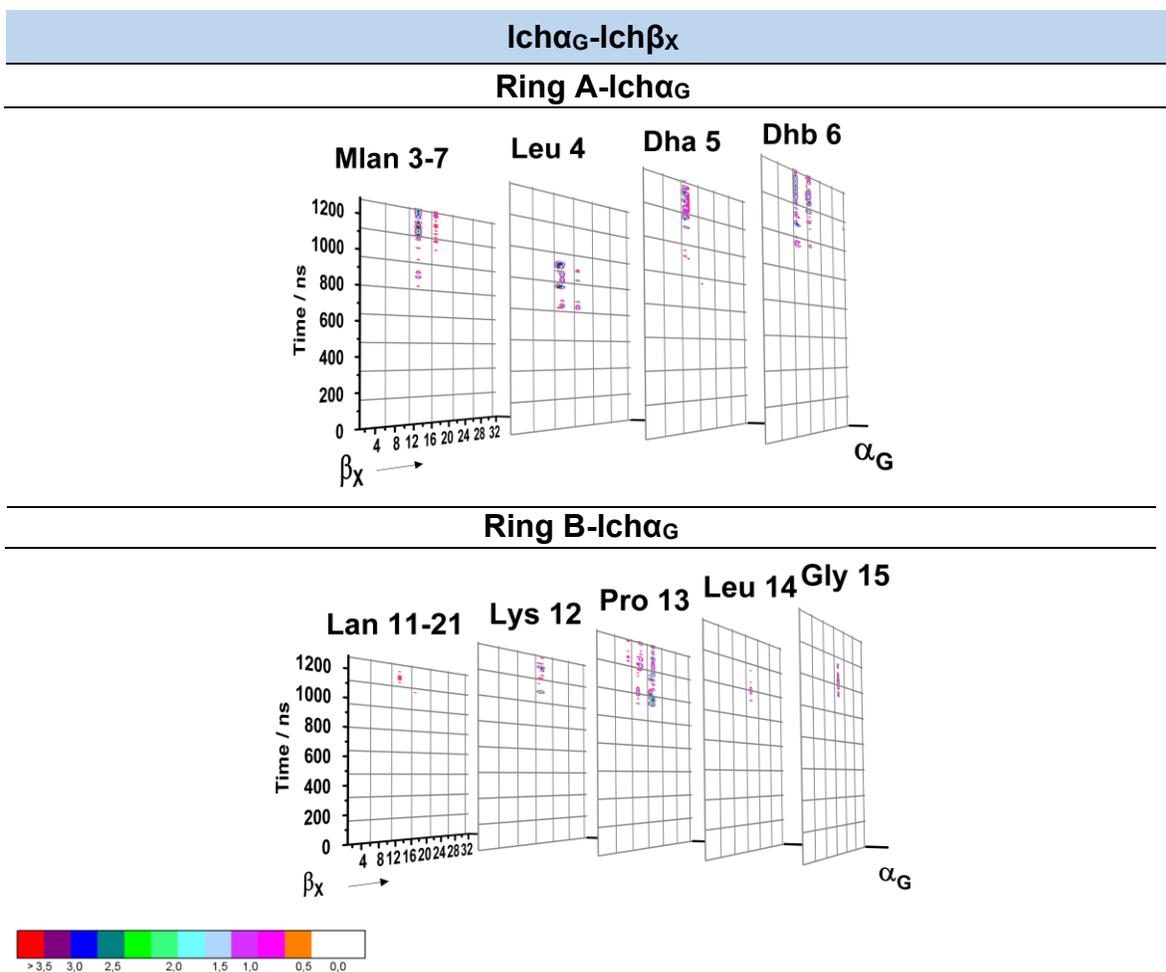


Figure B.2.2.4-B.20: TdCM of the interaction pathway followed between Ich α _G-Ich β _X. Mainly produced by hydrophobic interpeptide interactions.

The hydrophobic interactions and H-bonds for the pair Ich α _H-Ich β _N (figure B.2.2.4-B.21) peptides were detected during the reorientation stage. In this case, the Ich α _H-N_t together with some residues placed in the ring B contributed to the stabilization with the Ich β _N-ring A. Moreover, this interaction was also aided by the approximation of the Ich α _H-ring D to the Ich β _N-N_t and ring A section. Additionally, at the same stage of the simulation, this Ich α -component, generated new hydrophobic interactions due to the approximation of N_t, ring A and ring D to the Ich β _Q-N_t region (figure B.2.2.4-B.22).

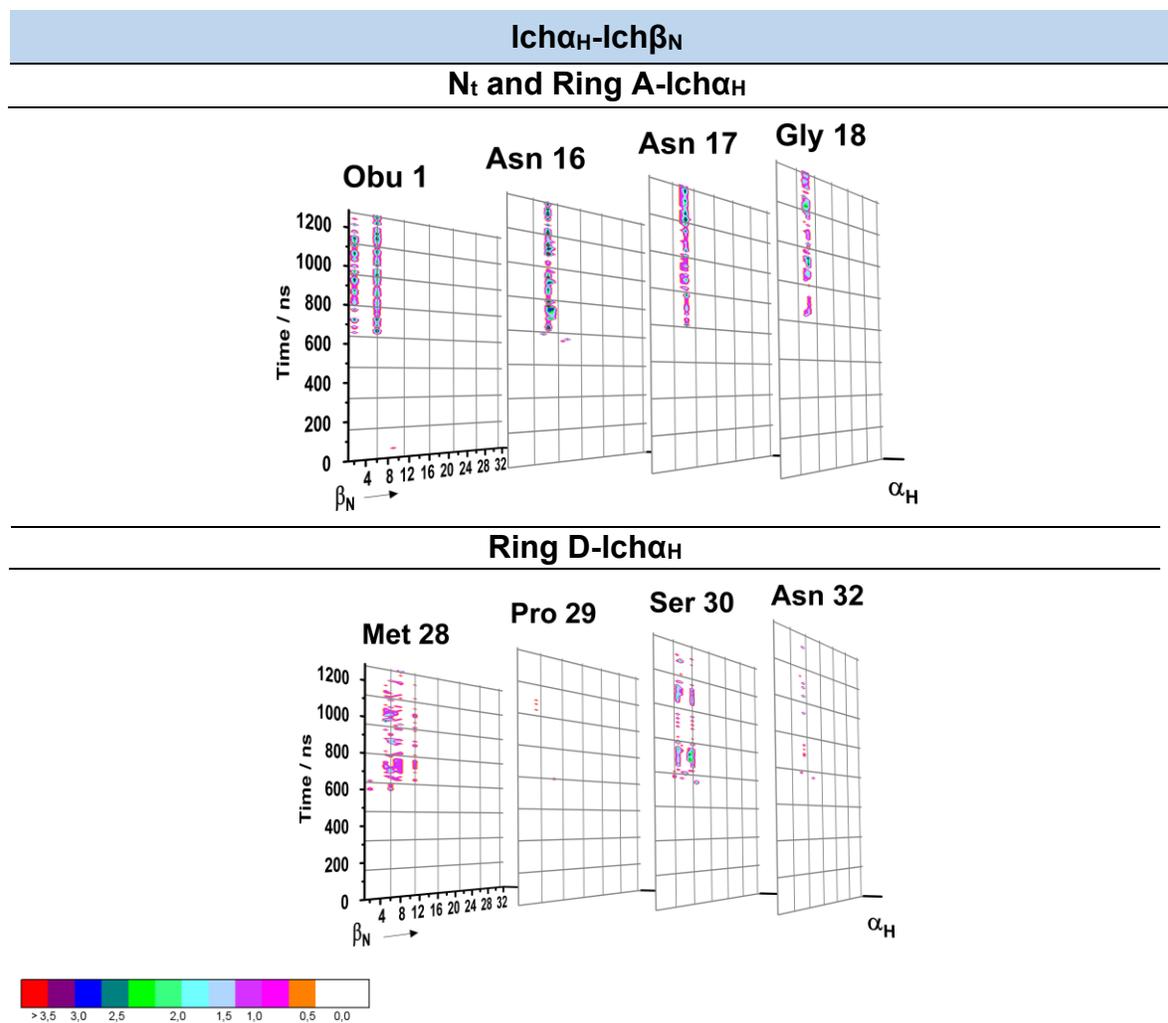


Figure B.2.2.4-B.21: TdCM of the interaction pathway followed between Ich α _H and Ich β _N. Mainly produced by hydrophobic and H-bonds interpeptide interactions.

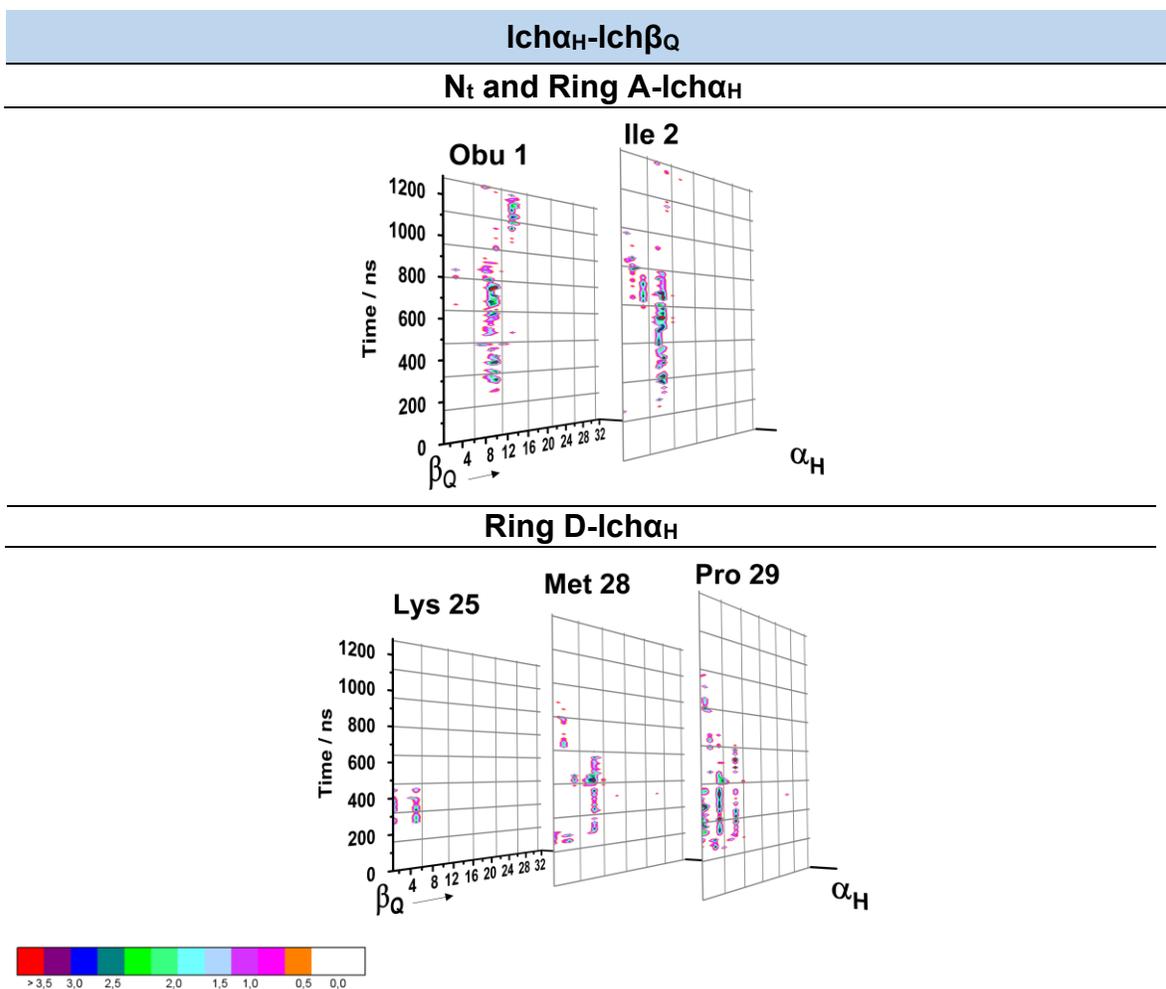


Figure B.2.2.4-B.22: TdCM of the interaction pathway followed between Ich α_H and Ich β_Q . Mainly produced by hydrophobic interpeptide interactions.

Concluding with Ich α_J , which promoted interactions with three Ich β -partners. Analogous behaviour was detected during the interaction with Ich β_Q and Ich β_U , while with Ich β_X the approach was established after few nanoseconds from the starting point. The peptides Ich α_J -Ich β_Q interacted through two different regions depending on the stage of the simulation. During the first stage, the Ich α_J -ring B region approached to the Ich β_Q -ring C and ring D which was also preserved until the end of the simulation by the H-bonds generated (figure B.2.2.4-B.23). Throughout the reorientation stage, the Ich α_J -ring B headed by the hydrophobic Pro13 and polar Tyr19 together with the rings C/ D promoted interactions with the β_U -N $_t$ region, ring A, α -helix region and ring D (figure B.2.2.4-B.24). In contrast with the previous complexes in AG-I and AG-II, the Ich α_J -Ich β_X peptides were from the beginning of the simulation in continuous contact. As it can be followed in the figure B.2.2.4-B.25, the

8. Bacteriocins

$I\alpha_J$ employed all different regions to interact primarily with the $I\beta_{\chi-N_t}$ region and ring A. Interestingly, the $I\alpha_J$ hydrophobic residues Ile2, Leu4 and Ile9 contributed to achieve a stable interaction with the $I\beta$ -rings B and C.

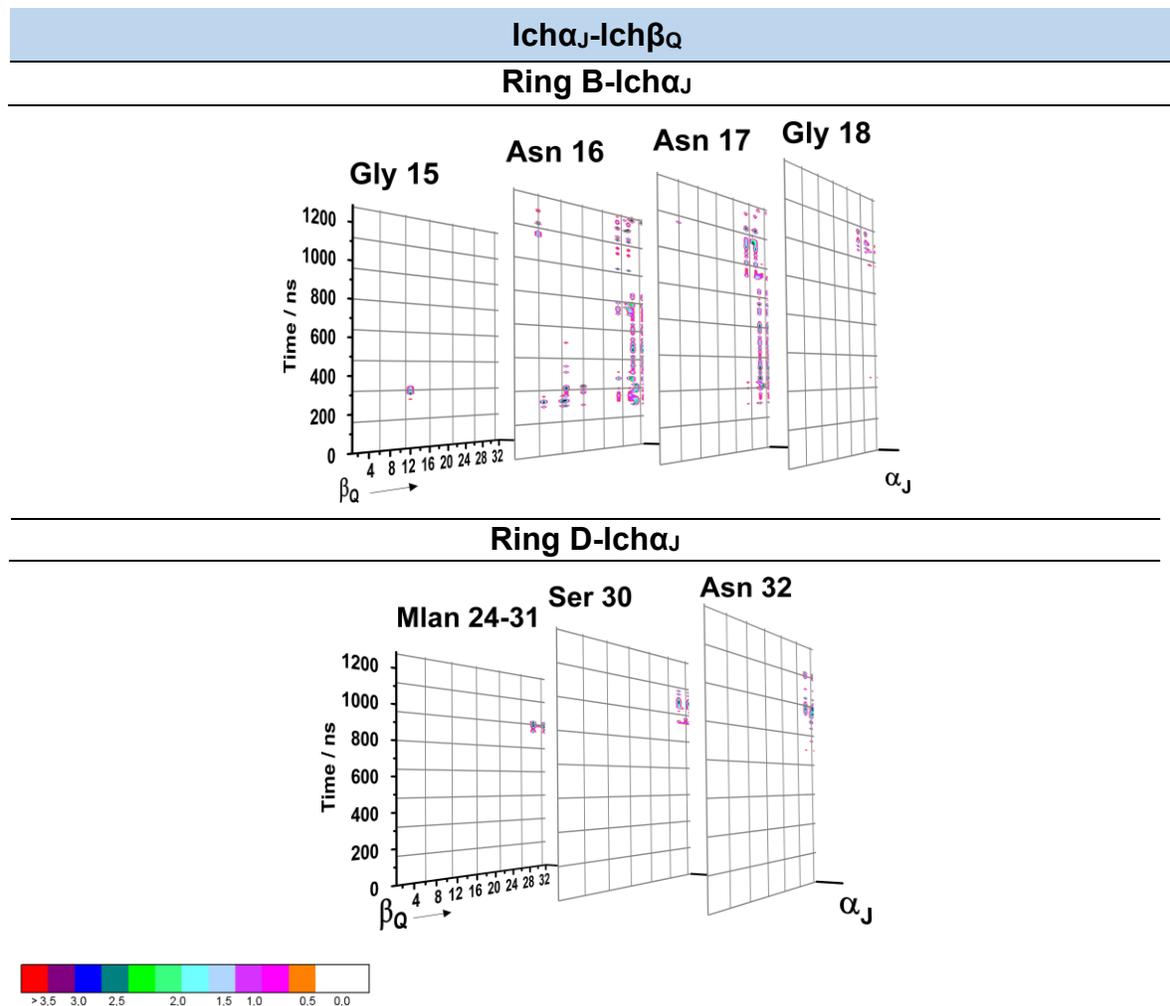


Figure B.2.2.4-B.23: TdCM of the interaction pathway followed between $I\alpha_J$ and $I\beta_Q$. Mainly produced by hydrophobic interpeptide interactions.

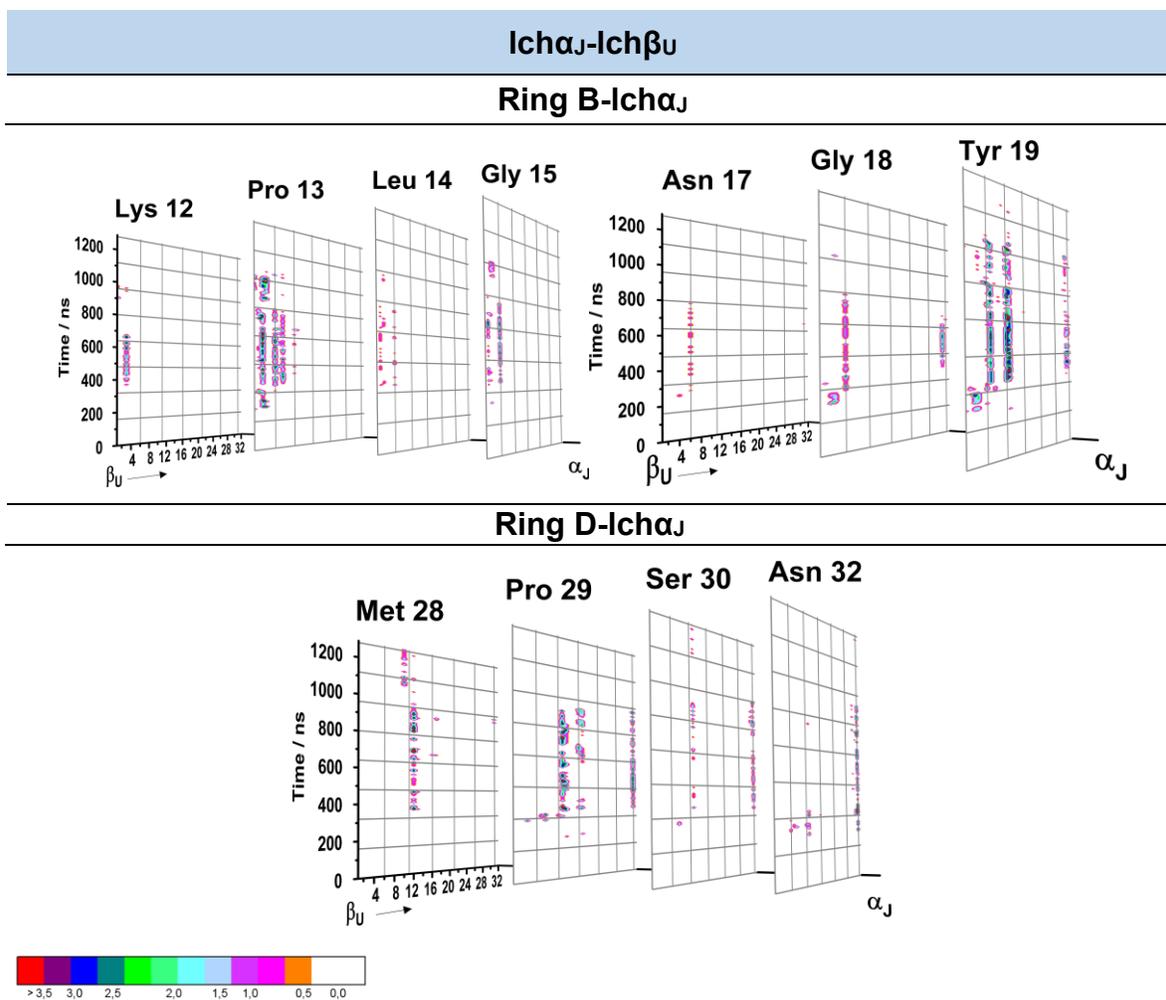


Figure B.2.2.4-B.24: TdCM of the interaction pathway followed between Ich α _J and Ich β _U. Mainly produced by hydrophobic interpeptide interactions.

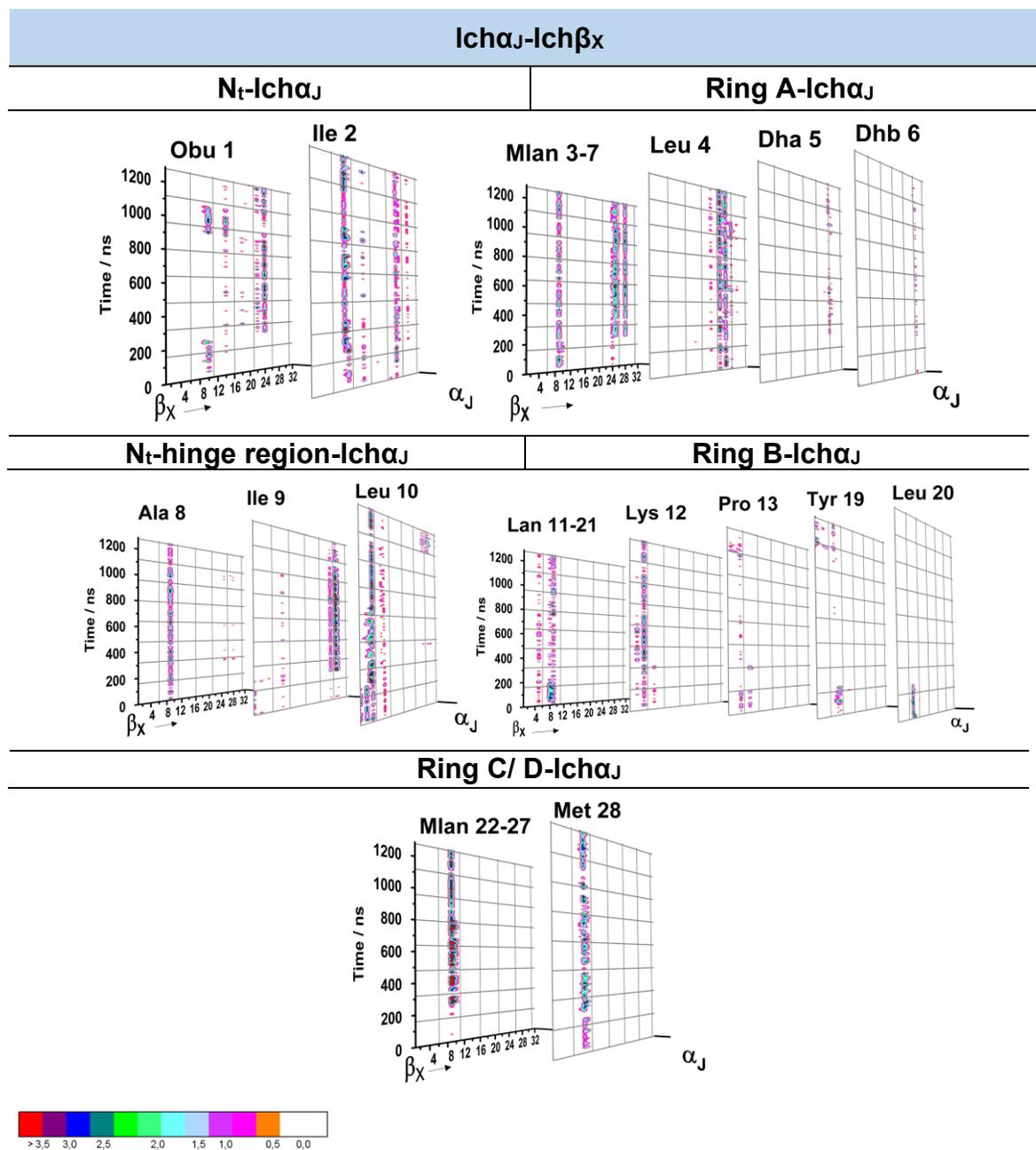


Figure B.2.2.4-B.25: TdCM of the interaction pathway followed between Ich α _H and Ich β _X. Mainly produced by hydrophobic interpeptide interactions.

AG-II: Interactions Summary $I\text{ch}\alpha$ - $I\text{ch}\beta$

Due to the high number of peptides involved in the AG-II aggregate, the interaction pathway between the two types of lichenicidin peptides slightly differ from the AG-I. In general, the $I\text{ch}\alpha$ -components approached the $I\text{ch}\beta$ -peptides through the rings A, B and the N_T -hinge region, while the $I\text{ch}\beta$ -peptides were more susceptible to interact with the α -helix and the N_T -region, in agreement with the observations detected for AG-I. It was clear, that depending on the position of the $I\text{ch}\beta$ -peptide in the aggregate, interactions of rings B, C and D were also found. Additionally, the role of the $I\text{ch}\alpha$ - N_T region was confirmed as an assistance point to promote and/or stabilize the interaction with the $I\text{ch}\beta$ -partners.

Additionally, the α - β interactions were mainly of hydrophobic nature which is why the $I\text{ch}\alpha$ -ring B and $I\text{ch}\beta$ -helix region together with the N_T were practically involved in every single α - β combinations described in this section.

B.2.2.4-C Membrane Effects for the Aggregate-Stabilization

Figure B.2.2.4-C.1 and C.2 describe the stabilization of the AG-I on the membrane surface during the course of the simulation, per peptide and per residue. The analysis was done by plotting i) the profile of the number of contacts with the membrane per residue at 500 ns and 900 ns of simulation-representing the approaching and reorientation stages (A-figure B.2.2.4-C.1 and B.2.2.4-C.2), ii) the evolution of the contacts strength per residue in the course of the simulation (B-figure B.2.2.4-C.1 and B.2.2.4-C.2), iii) total number of contacts to POPG and POPE lipids (C-figure B.2.2.4-C.1 and B.2.2.4-C.2) and iv) evolution of the number of contacts of the membrane established by the most active residues (D-figure B.2.2.4-C.1 and B.2.2.4-C.2).

In A-figures B.2.2.4-C.1 and C.2 is displayed the specific interacting profile observed at 500 ns (approaching phase) and at 900 ns (after the reorientation phase). Generally, in these two selected check points, α -peptides showed random interactions with the membrane while the β -components were perfectly anchored to it.

During the first 500 ns (figure B.2.2.4-C.1, α_A -POPG/POPE, see B), $I\text{ch}\alpha_A$ approaches the surface employing Pro13, Asn16 and Leu20, although at 500 ns just the interaction of Leu20 with the membrane surface was noticed (figure B.2.2.4-C.1, $I\text{ch}\alpha_A$ -

POPG/POPE, see A). While in the reorientation phase the Asn16 and Asn17 are the residues responsible for the assembling of the peptide in the surface. The anchoring role of Asn16 is aided by Asn17. The interaction strength with the membrane is increased (figure B.2.2.4-C.1, I α_A -POPG/POPE, see D). I α_A peptide starts the interaction with the surface of the membrane after almost 200 ns to the POPG lipids (figure B.2.2.4-C.1, I α_A -POPG/POPE, see C), some interactions were also noticed with the POPE lipids but in a minor proportion until the end of the simulation, were the contacts to lipids seem to be almost identical. In contrast to the findings found for I α_A , I α_B showed a different behaviour (figure B.2.2.4-C.1, I α_B -POPG/POPE, see B). Only some occasional contacts involving mainly Ile2, Leu4, Dha5, Dhb6 and randomly the Tyr19 and Pro29 with the surface of the membrane were detected. The contact with the membrane surface is attributable to the hydrophobic character of the Ile2, Leu4 and electrophilic nature of the Dha5 and Dhb6. Due to this alternating approach, the contact intensity is weaker (figure B.2.2.4-C.1, I α_B -POPG/POPE, see D). Furthermore, despite that the most active lipid interacting precursor is POPG (contributing in a higher proportion to the surface composition) the contacts with POPE are intensified in comparison with the I α_A (figure B.2.2.4-C.1, I α_B -POPG/POPE, see C). The I α_C peptide, which was in a close contact with LII₂, employs primarily the ring B (mainly Lys12 with contributions of the hydrophobic residues Leu14, Gly15 and polar Asn16 together with the N_T-hinge Leu10) and the ring D Mlan24-34 to be anchored to the surface of the membrane until the end of the simulation (figure B.2.2.4-C.1, I α_C -POPG/POPE, see B and D). Contrary to the behavior of observed in the previous peptides I α_A and I α_B , (figure B.2.2.4-C.1, I α_C -POPG/POPE, see C) a stable interaction with the surface of the membrane was detected. This observation confirms that the achievement of an adequate spatial arrangement of the peptide with the lipid II contributes simultaneously to the stabilization of the peptide on the membrane surface. The interactions of I α_D peptide with the surface were localised essentially on the N_T-hinge region (figure B.2.2.4-C.1, I α_D -POPG/POPE, see C) through Ala8, Ile9 and Leu10 with some influences of Dhb6 and Lys12. Interestingly, the number of contacts increase (figure B.2.2.4-C.1, I α_D -POPG/POPE, see D) during the reorientation stage (500-900 ns) were the peptide starts to interact with the POPE lipid (figure B.2.2.4-C.1, I α_D -POPG/POPE, see C).

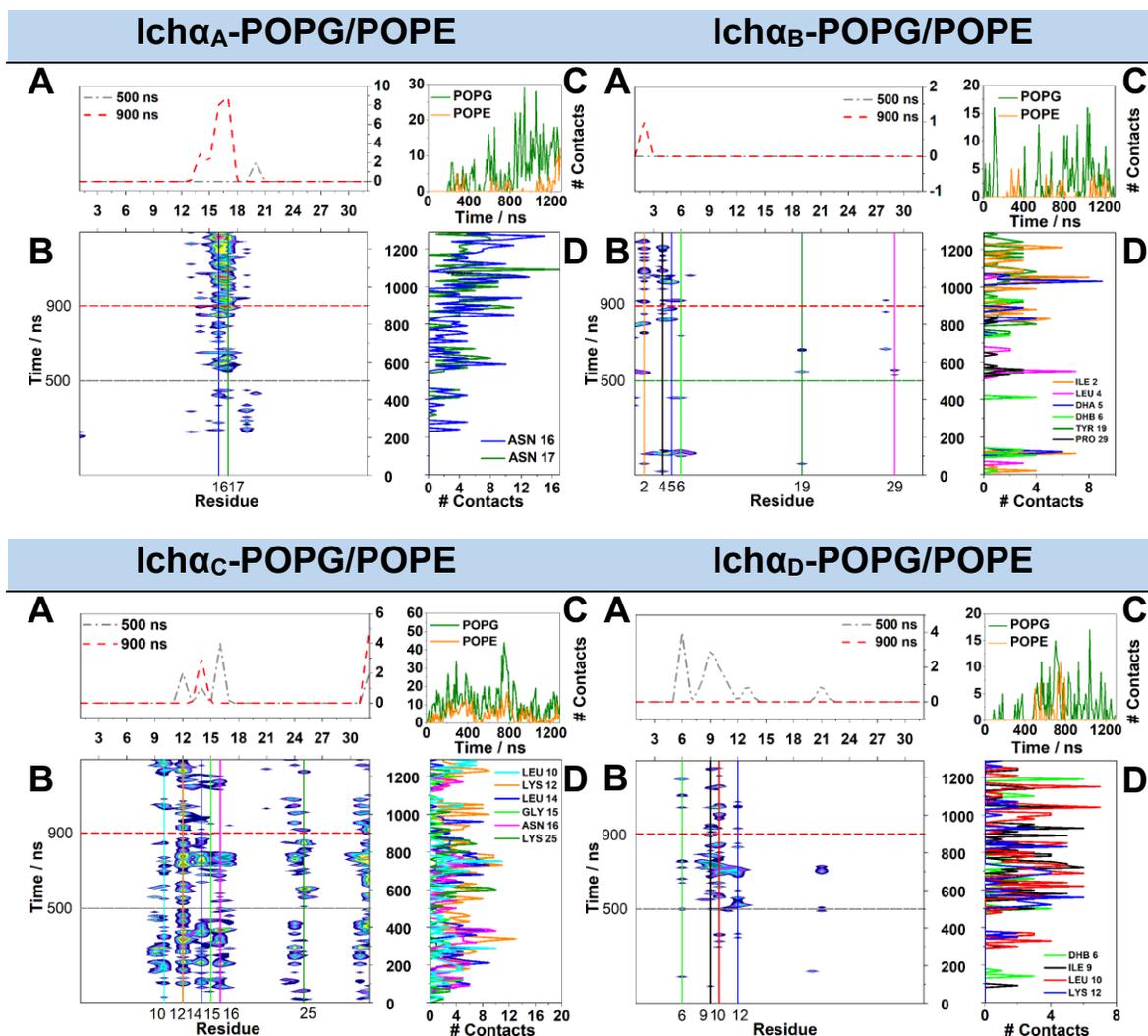


Figure B.2.2.4-C.1: Monitoring membrane- $Ich\alpha$ -peptide interaction for AG-I. **A-** Profile of the number of contacts per residue at 500 ns and 900 ns of simulation—representing the approaching and reorientation stages, **B-** Evolution of the contacts strength per residue in the course of the simulation, **C-** Total number of contacts to POPG and POPE lipids and **D-** Evolution of the number of contacts of the membrane established by the most active residues.

The β -peptides behaviour on the membrane differ significantly from the $Ich\alpha$ -components. Concisely, from Ile12 to Val18 the $Ich\beta$ -peptides are more susceptible to interact with the $Ich\alpha$ -partners (see graphics $Ich\alpha$ - $Ich\beta$). With a marked difference in comparison with the cases studied so far, the N_t region together with ring B and more strongly, rings C and D promote the anchoring to the membrane (figure B.2.2.4-C.2, $Ich\beta_R$ and $Ich\beta_S$ -POPG/ POPE, see B, D). Concluding with the fact that the total number of contacts to POPG and to POPE

8. Bacteriocins

resulted quite similar in both peptides (figure B.2.2.4-C.2, $\text{Ich}\beta_{\text{R}}$ and $\text{Ich}\beta_{\text{S}}$ -POPG/ POPE, see C).

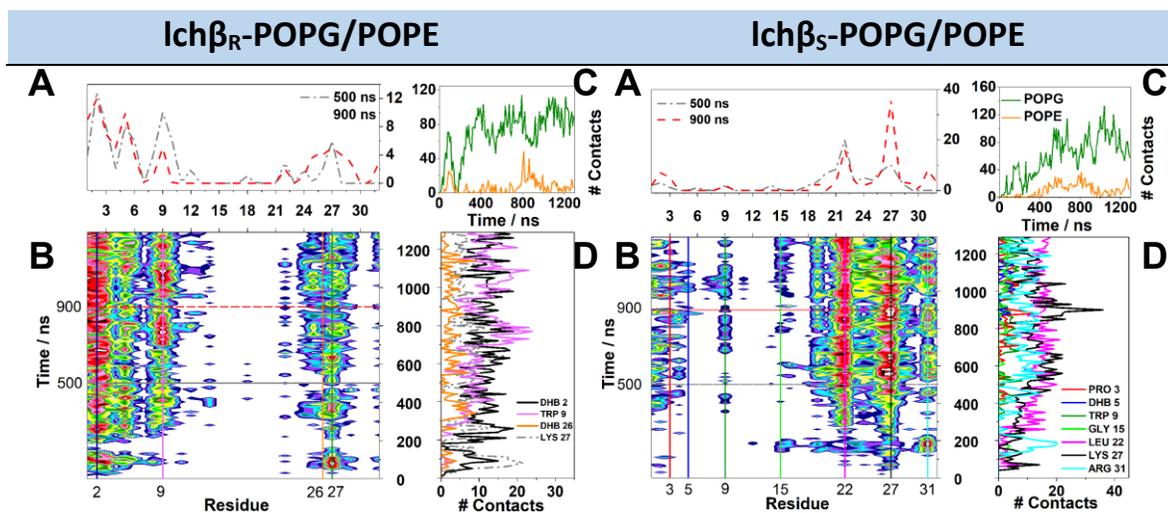


Table B.2.2.4-C.2: Monitoring membrane- $\text{Ich}\beta$ -peptide interaction for AG-I. **A-** Profile of the number of contacts per residue at 500 ns and 900 ns of simulation-representing the approaching and reorientation stages, **B-** Evolution of the contacts strength per residue in the course of the simulation, **C-** Total number of contacts to POPG and POPE lipids and **D-** Evolution of the number of contacts of the membrane established by the most active residues.

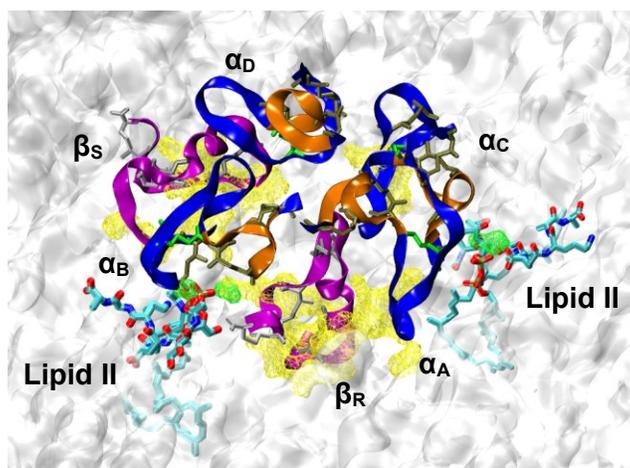


Figure B.2.2.4-C.3: Snapshot of the AG-I stabilization on the membrane surface. $\text{Lch}\alpha$ -peptides are displayed in blue with the ring C/ D highlighted in orange while $\text{Ich}\beta$ -components are displayed in magenta. Interactions with lipid II are shown in green and the yellow color denotes the interactions with the surface.

AG-I: Interactions with the Membrane Surface Summary

The $\text{Ich}\beta$ -peptides employ the N_t as well as rings C and D to strongly interact with the membrane. In contrast to the $\text{Ich}\alpha$ -units, which showed lower level of interaction through the N_t -hinge and ring B region.

The stabilization of AG-II over the membrane surface differs from that predicted for AG-I due to the increased number of peptides involved (figure B.2.2.4-C.4-5). In AG-II, concisely 4 pairs of α - β -lichenicidin (8 peptides in total) interacted with the membrane. Predominantly, the $\text{Ich}\beta$ -peptides played the role of interconnectors between the α -peptides (see previous section B.2.2.4-B) causing a decreased $\text{Ich}\beta$ -interaction with the membrane surface in comparison to AG-I. Interestingly, the $\text{Ich}\beta$ -peptides β_Q , β_U , β_V , β_X which interacted with the POPG/POPE membrane employed the same regions than their analogues in AG-I. Concisely, the N_t and rings B, C and D (figures B.2.2.4-C.4-5, see B and D). Moreover, as it can be observed in the figures B.2.2.4-C.4-5 (see A and C), the strongest interactions were found during the last simulation stage, such as in the case of $\text{Ich}\beta_Q$ or $\text{Ich}\beta_X$.

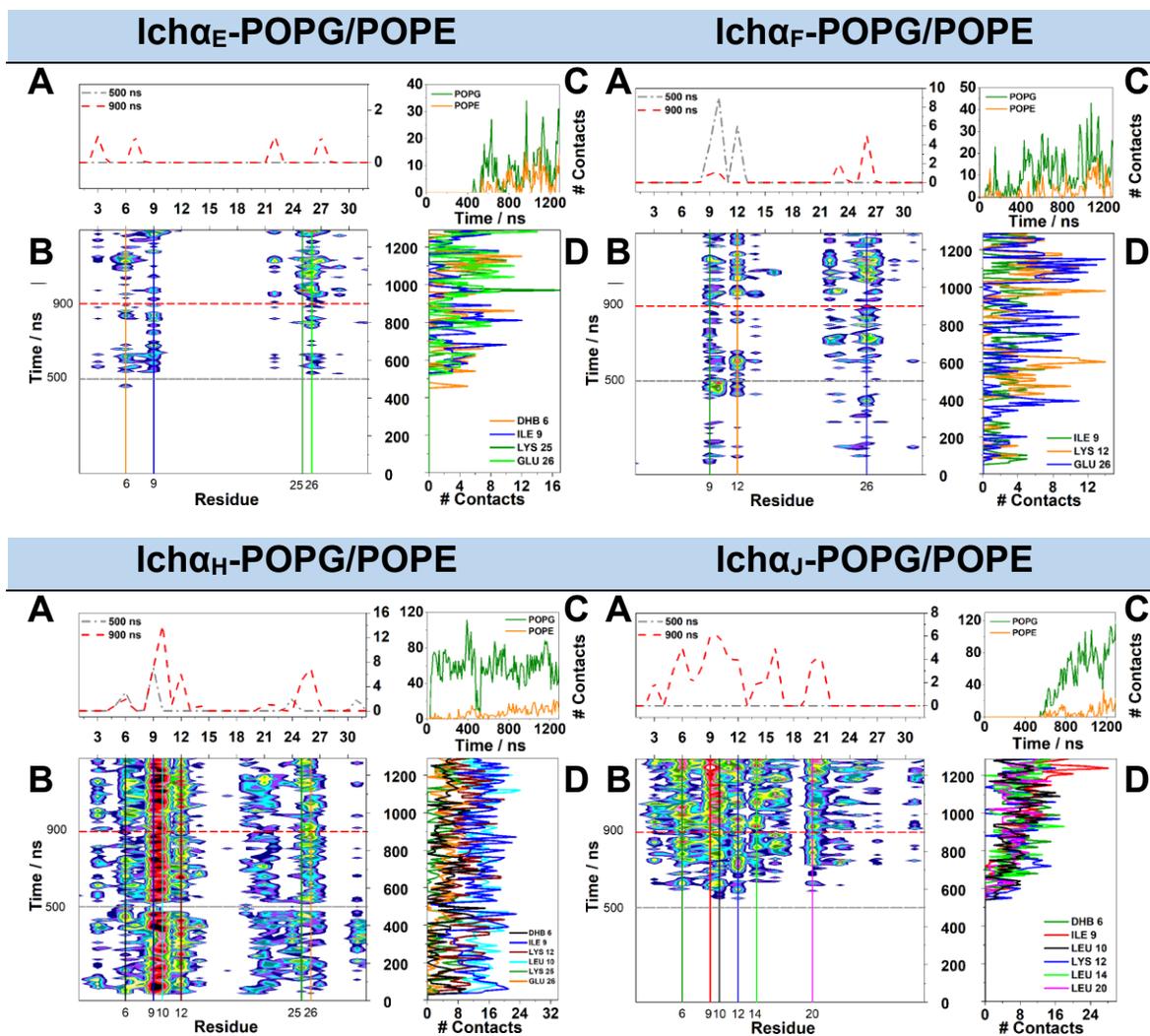


Table B.2.2.4-C.4: Monitoring membrane-Ich α -peptide interaction for AG-II. **A-** Profile of the number of contacts per residue at 500 ns and 900 ns of simulation-representing the approaching and reorientation stages, **B-** Evolution of the contacts strength per residue in the course of the simulation, **C-** Total number of contacts to POPG and POPE lipids and **D-** Evolution of the number of contacts of the membrane established by the most active residues.

Additionally, a strong interaction by the $Ich\alpha_H$ and $Ich\alpha_J$ with the membrane was identified. A noteworthy exception was the case for $Ich\alpha_E$ and $Ich\alpha_F$ peptides, which provided a similar behaviour as it was detected for $Ich\alpha_C$ and $Ich\alpha_D$ in AG-I (figures B.2.2.4-C.1-2). In AG-II, the $Ich\alpha$ -Dhb6, Ile9, Lys12, Lys25 and Glu26 were the essential residues interacting with the membrane and stabilizing the aggregate over the surface. Analogous to AG-I, in AG-II only a single $Ich\alpha$ -peptide interacted in a stable manner with a lipid II molecule promoting indirectly a strong interaction with the membrane, such as the case of $Ich\alpha_H$

(figures B.2.2.4-C.4, $\text{Ich}\alpha_{\text{H}}$ -POPG/POPE, see A). Generally, at the selected check points simulations controls, 500 ns and 900 ns figures B.2.2.4-C.4-5, see A, all peptides showed precise interactions with the membrane without any particular difference regarding $\text{Ich}\alpha$ or $\text{Ich}\beta$ peptides.

To conclude, neither $\text{Ich}\alpha_{\text{G}}$ and $\text{Ich}\beta_{\text{N}}$ showed any interaction with the membrane surface nor the AG-III composed by $\text{Ich}\beta_{\text{O}}$ and $\text{Ich}\beta_{\text{T}}$ peptides.

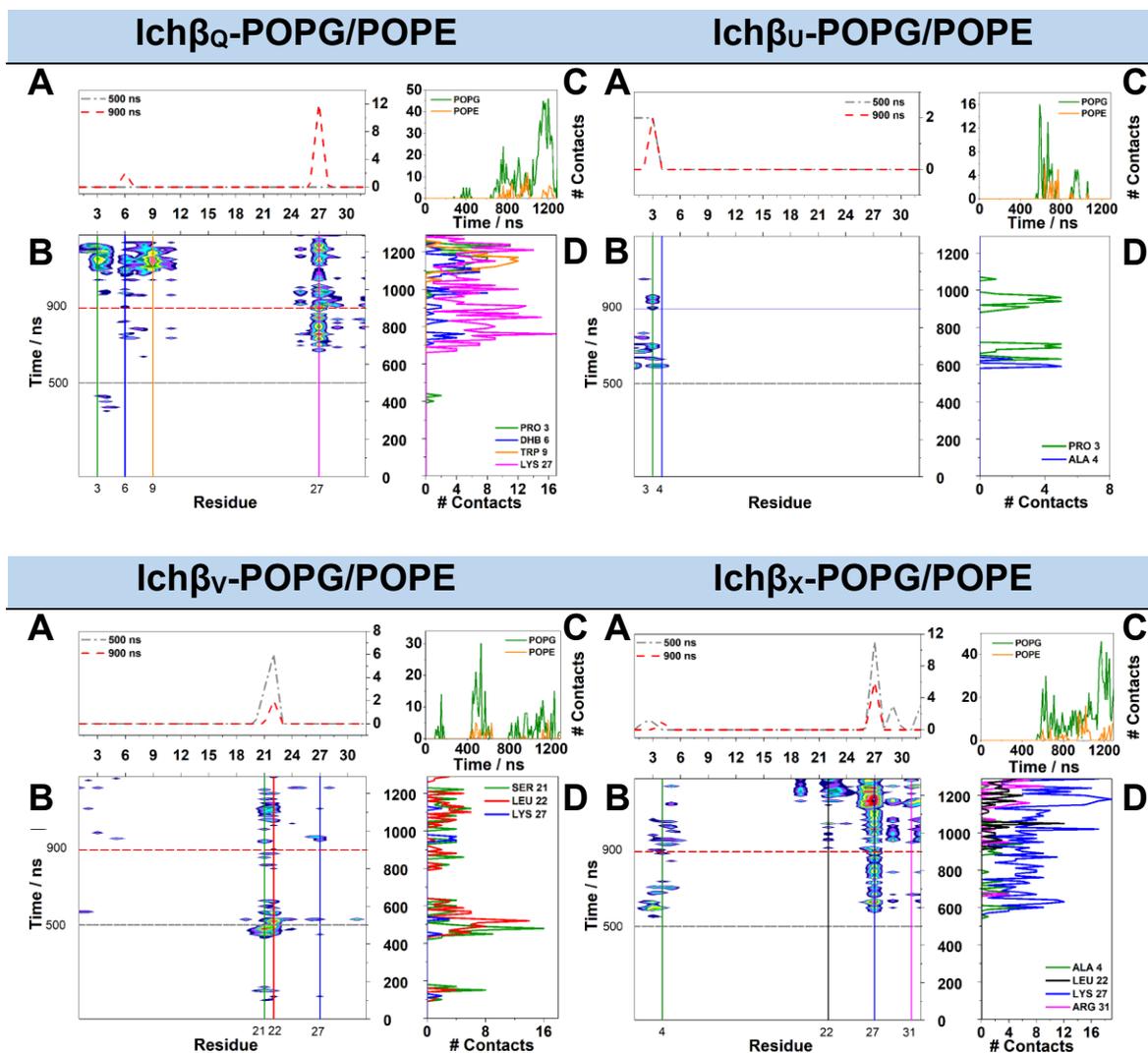


Figure B.2.2.4-C.5: Monitoring membrane- $\text{Ich}\beta$ -peptide interaction for AG-II. **A-** Profile of the number of contacts per residue at 500 ns and 900 ns of simulation-representing the approaching and reorientation stages, **B-** Evolution of the contacts strength per residue in the course of the simulation, **C-** Total number of contacts to POPG and POPE lipids and **D-** Evolution of the number of contacts of the membrane established by the most active residues.

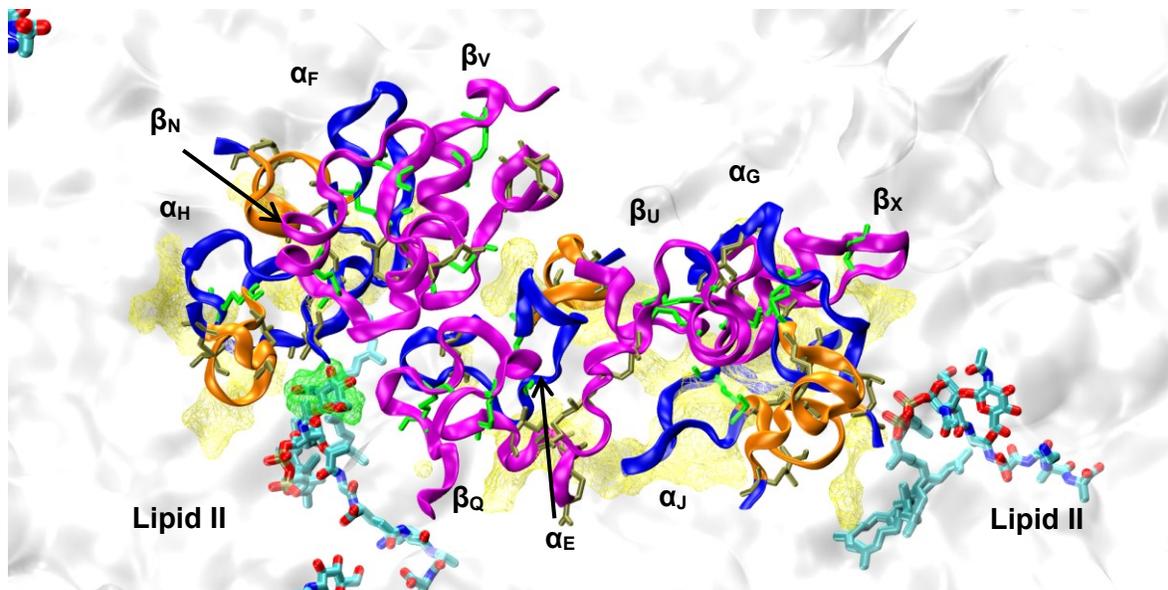


Figure B.2.2.4-C.6: Snapshot of the AG-II stabilization on the membrane surface. lch α -peptides are displayed in blue with the ring C/ D highlighted in orange while lch β -components are displayed in magenta. Interactions with lipid II are shown in green and yellow color denotes the interactions with the surface.

AG-II: Interactions with the Membrane Surface Summary

Mostly, due to the size of the aggregate, 5 lch α : 5 lch β , a different interaction pattern with the membrane surface was detected in comparison with AG-I. Herein, the lch α -peptides promote the strongest interactions with the membrane surface involving almost the entire sequence leaving the ring B region partially free. Whereas the lch β -peptides anchored to the membrane involving the N_t region and the ring C, in line with the behaviour observed in AG-I. Only a single lch α -component preserved a stable interaction with Lipid II.

B.2.2.5 Role of Lys25 and Glu26

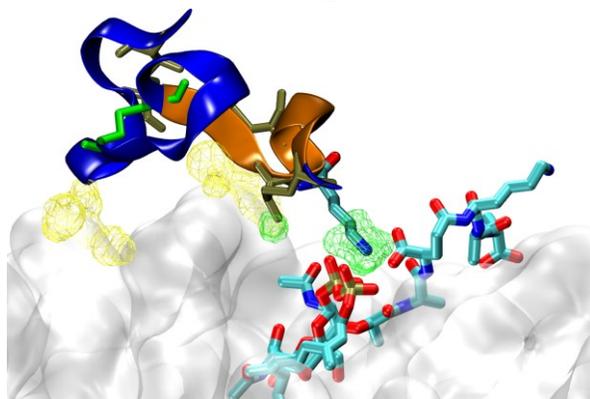
Comparatively to the up-to-date discovered lch α -peptides, (see figure 8.1.1.1 in section 8.1.1.1-B), the residue at the 25th position is variable. For example, α -Lichenicidin is characterized by the presence of a lysine at the 25th position as in Frca^[131], in contrast to Ltn α ^[127], Plw α ^[120], C55 α ^[125,126] or Hal α ^[128], Bht α ^[124], Smb α ^[123] and Thsa^[130] where a histidine (pK_R = 6.0 neutral) or a valine, are respectively found.

The scenario found in AG-I where a single α -peptide is interacting with a lipid II molecule without a direct interaction between the docking molecule and lch β -peptide, is fundamental to explain the experimental observations regarding the bacteriocin activity of most of α -components on its own, which resulted stimulated by the presence of the lch β -peptides.^[120,127] The MD simulation is consistent with this observation, since a stable α -lipid II (see figure B.2.2.1.1) interaction was found without the direct lipid II-lch β linkage.

Indeed, as it can be observed, the stable interaction of the Lys25 with the phosphate cage is also aided by the FGA of the lipid II side chain. This interesting observation opens a second scenario to be studied according to the arrangement of the peptide in the surrounding of lipid II, accompanied with the partial maintenance of this residue along the two-component lantibiotic family.

A α_C -LI₂

LYS25-phosphate cage



B α_H -LI₅

GLU 26-phosphate cage

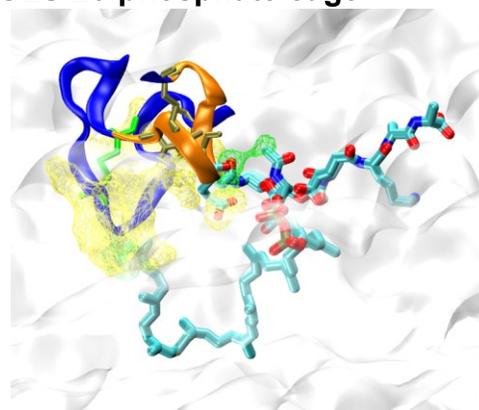


Figure B.2.2.5.1: Interaction between α and lipid II component. A- in AG-I lch α_C peptide, **B-** in AG-II lch α_H peptide. The peptides around α_C and α_H are not displayed for clarity. Green color shows the electrostatic interaction, while in yellow the interaction with the membrane surface is shown. The peptide region involved for the stabilization with the POPG/ POPE membrane is described in section B.2.2.4-C.

On the other hand, the significant overall repulsive electrostatic interaction between α_H with LI₅ was observed (section B.2.2.2). The general positive electrostatic interaction between α_H with LI₅ resulted of relevant importance in order to understand the role of the α -ring C/ D in the attachment to lipid II component. Responsible for the interaction with the phosphate cage in AG-II is Glu26, a conserved amino acid in all α -peptides of the

two-component lantibiotic family, except for Cytolysin and bicereucin which do not belong to the category of Mersacidin-like systems (see section 8.1.1.1-B, figure 8.1.1.1.1 and figure 8.1.1.1.2). Accordingly, an evolutionary cause must exist for the high degree of conservation of this residue along all the types of α -peptides (figure B.2.2.5.2). Indeed, experimental studies have reported the abolishment of antibacterial activity through the mutation of this specific residue (ring C-Glu)^[128,152] in line with previous analogous studies carried out with mersacidin^[128,153]. Based on these results, it is reasonable to suggest that the overall repulsive interaction between the Glu26 (in the case of Lichenicidin)-ring C and the Lipid II phosphate moiety perturbs the membrane surface, and consequently i) this residue is conserved in all two-component lantibiotic family and ii) as it is the case of Lichenicidin, if the previous residue is a positively charged fragment the phosphate cage attachment to lipid II can be additionally stabilized via attractive electrostatic interactions with FGA.

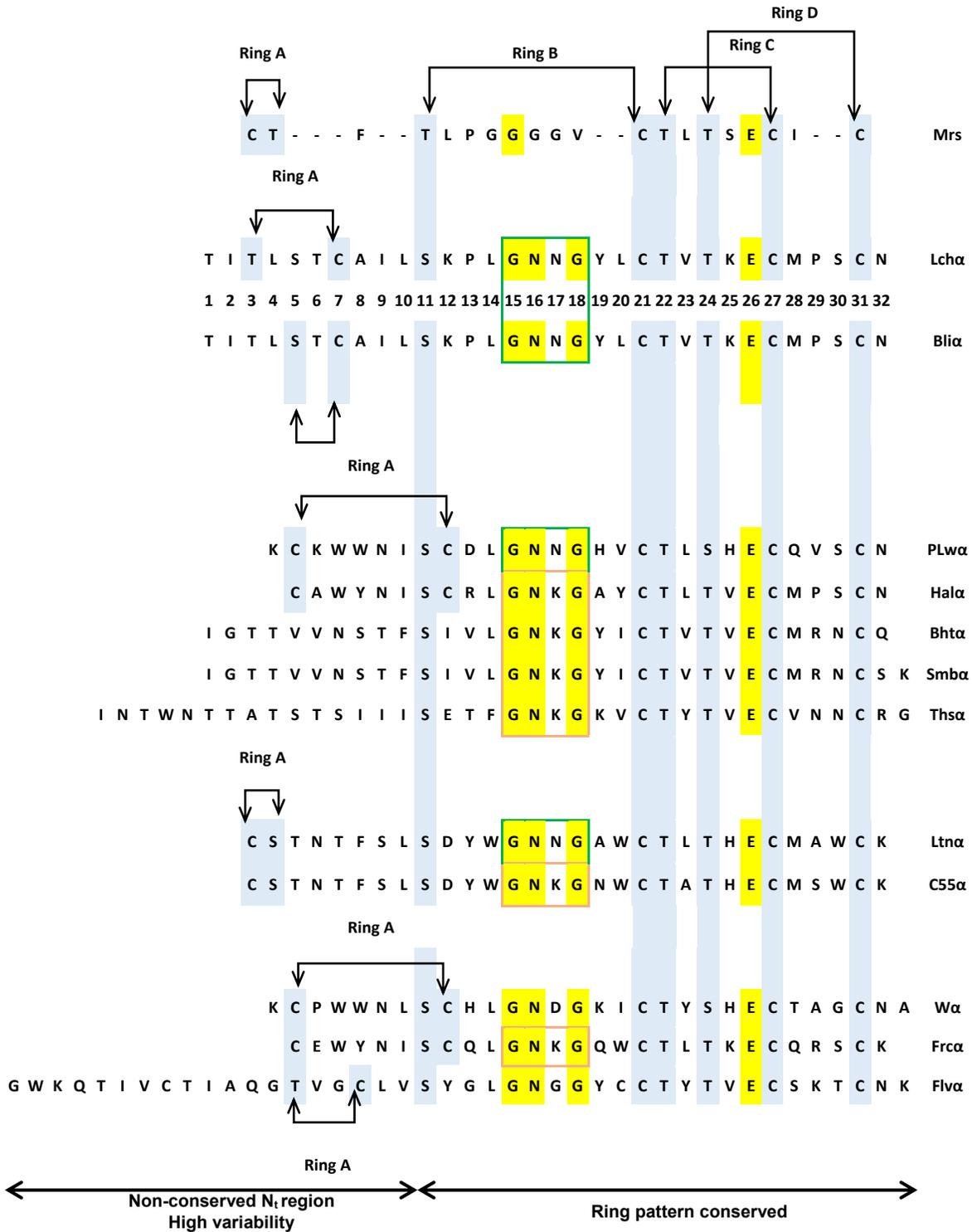


Figure B.2.2.5.2: Overview of the primary sequences of α -components discovered up-to-date. The structures after the post-translational modifications are displayed in the figure 8.1.1.1.1, section 8.1.1.1-B.

8. Bacteriocins

While for the Ich β -components (figure B.2.2.5.3), a high variable sequence and partially conserved ring pattern have been found.

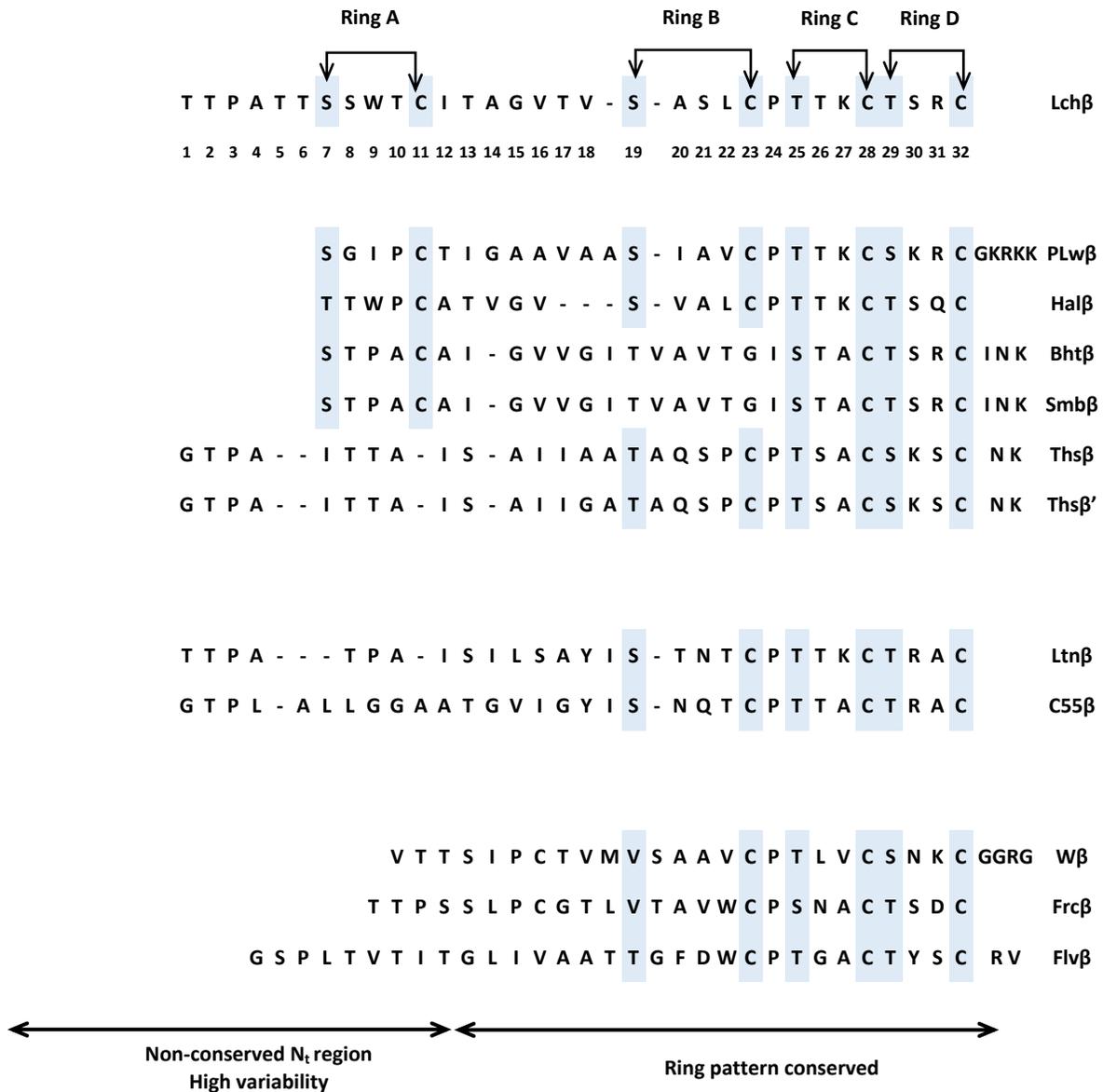


Figure B.2.2.5.3: Overview of the primary sequences of all- β -components discovered up-to-date. The structures after the post-translational modifications are displayed in the figure 8.1.1.1.1, section 8.1.1.1-B.

B.2.2.6 Ring Pattern Analysis

Despite the high variability regarding the bridging pattern at the N_t region, the intramolecular cross-link denoted as rings B, C/ D are well-conserved in all α -peptides. (see figure B.2.2.5.2). This stable and unique structural moiety has been, consequently, associated to an important functional role. Genetic engineering studies have revealed that mutations involving the bridging pattern resulted to be detrimental for the antimicrobial activity.^[120,128] Surprisingly, α -ring C (according to the nomenclature used in this thesis) was said to be unnecessary for bioactivity, which is uncommon given the high degree of conservation amongst bacteriocins.^[128] For instance, studies in which the Hal α ring B (according to the nomenclature used in this thesis) were disrupted, highlighting its importance but rendered it non essential due to the still observed (less effective) inhibition of lipid II polymerization.^[128] While ring D (according to the nomenclature used in this thesis) was suggested to be essential.^[128,131,154]

In this study, it was observed that the α -peptides achieved a stable interaction with lipid II, and approached it involving rings C/ D (see section B.2.2.1) like previously observed in the case of Mersacidin. Interestingly, the thioether ring pattern for the β -peptides of the two-component family (figure B.2.2.5.3), concisely for rings B, C and D is well-conserved. While the N_t region displays a lower degree of conservation amongst the bacteriocins. The N_t region is generally characterized by hydrophobic residues which have been suggested to play a role in the membrane insertion^[131], which was confirmed by the trajectory analysis of this thesis (see section B.2.2.4-C). Additionally, it has been proven that even in the not mersacidin-like peptides such as bicereucin, the presence of the ring is crucial for the β -activity.^[129] Interestingly, for the β -haloduracin, disruption of rings C and D promoted a decreased antimicrobial activity, while ring A was dispensable. The role of ring B could not be evaluated due to perturbations from the rest of the rings.^[128,131] These findings are in line with the results explained in section B.2.2.5-C.

The following table 2.2.6.1 summarizes the consequences in the modification of the rings observed at experimental level for Haloduracin two-component lantibiotic. Additionally, due to the similarities found in the two-component lantibiotic family some characteristics have been also added. These experimental observations are further compared with the theoretical results obtained within this thesis.

Role	Experimental observations	Computational proof after 1.3 μ s of all-atom MD for Lichenicidin
α		
Ring A	High variability Sequence and ring pattern	Responsible together with the N _i region of the ' supportive interaction ' for all possible inter-peptide interactions studied here <i>See-section B.2.2.4-A</i>
Ring B	Bridging pattern conserved Important but not essential Sequence partially conserved: GN*G with *= N or K or D or G N-Glutamine (Asn) K-Lysine (Lys) D-Aspartic acid (Asp) G-Glycine(Gly)-also case of Mersacidin <i>Figure B.2.2.5.1</i>	<ul style="list-style-type: none"> • Rudder region-sequence partially conserved because: <ol style="list-style-type: none"> 1. G-provides flexibility to the loop in order to act as rudder ring. 2. N-directing residue • β-concentration dependent functionalities: <ol style="list-style-type: none"> 1. Anchoring to the membrane aided by the presence of hydrophobic residues combined with possible H-bonds attributed to N with the phospholipid head groups 2. Redirect the β-peptides <i>See-section B.2.2.4</i>
Ring C	Bridging pattern conserved Essential Glu residue	Electrostatic interaction with the phosphate cage, in case of Lichenicidin Lys25 and GLu26
Ring D	Bridging pattern conserved Variable sequence Essential for bioactivity	Anchorage to the bacterial membrane supporting/ contributing to the interaction of α with lipid II <i>See-section B.2.2.4-C. lchα_H and lchα_C, and at the last stage lchα_J</i>
β		
Ring A	High variability Ring unessential for activity	Hydrophobic interactions (Pro2, Ala4, Trp9) combined with the 2-Obu, H-bonds/ hydrophobicity promoted the immersion in the membrane surface. <i>See-section B.2.2.4-C</i>
α -helix region		Section in continuous contact with α -partners <i>See-section B.2.2.4-B, C</i>
Ring B	Bridging pattern conserved Uncertain results after disruption	Hydrophobic interaction with the membrane surface headed mainly by Leu22, some H-bond contribution of Ser21 <i>See-section B.2.2.4-C</i>
Ring C	Bridging pattern conserved Important but not essential	Main ring for membrane-interaction Electrostatically through Lys27 <i>See-section B.2.2.4-C</i>
Ring D	Bridging pattern conserved Important but not essential	Supportive activity for ring C, through the electrostatic interaction with membrane surface by Arg31 <i>See-section B.2.2.4-C</i>

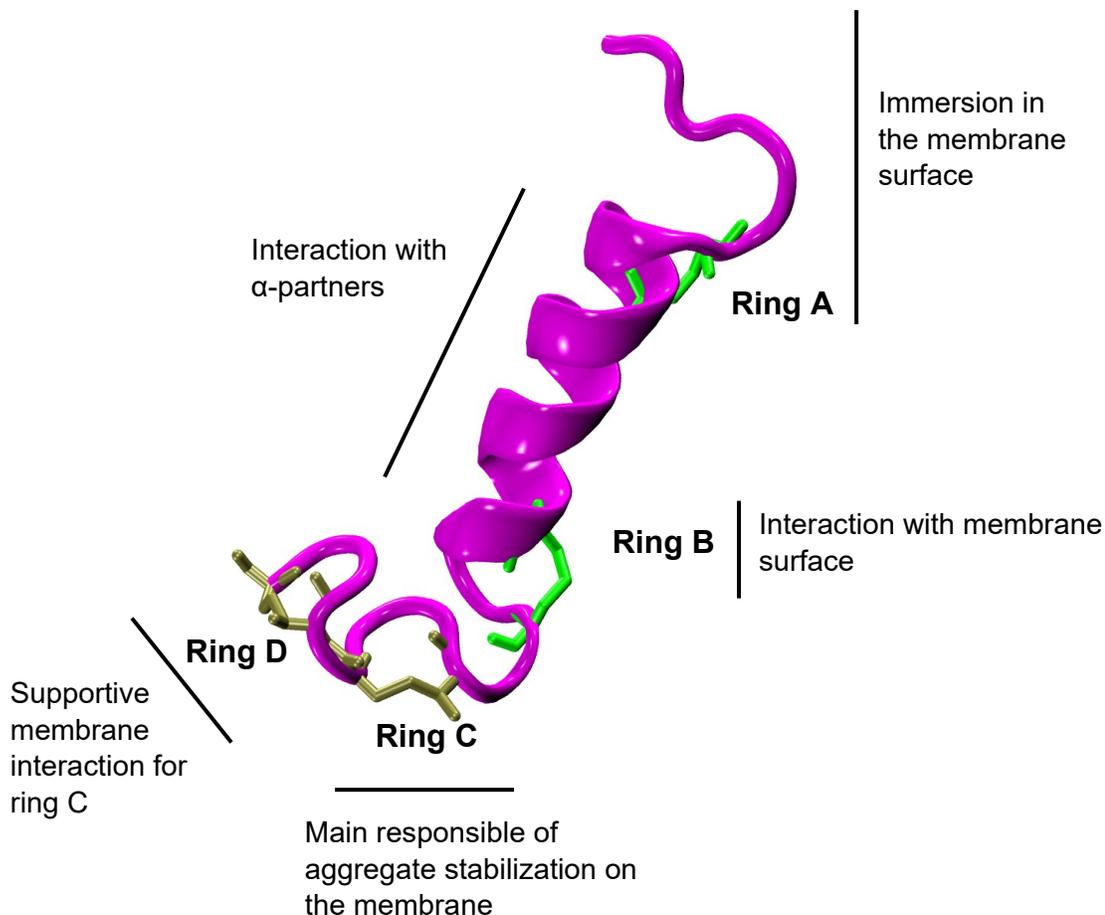
Table B.2.2.6.1: Analysis of the thioether rings in the lantibiotics by comparison of the experimental and theoretical results. *According to the nomenclature used in this thesis for the denomination of the rings (figure 8.1.1.1.1, section 8.1.1.1-B).

Indeed, despite the random interactions between different α - β -peptides observed during 1.3 μ s-long all-atom MD simulation only a stoichiometry consisting of 1 α : 1 β : 1 lipid II ratio was stable (section B.2.2.1). These findings are in good agreement with a series of experimental data showing that the optimal synergistic activity is found with a 1:1 peptide ratio for lacticin3147^[127], plantaricin W^[120], staphylococcin C55^[125,126], Smb^[123], enterocin W^[121,122], Bht^[124], thusin^[130] and formicin^[131].

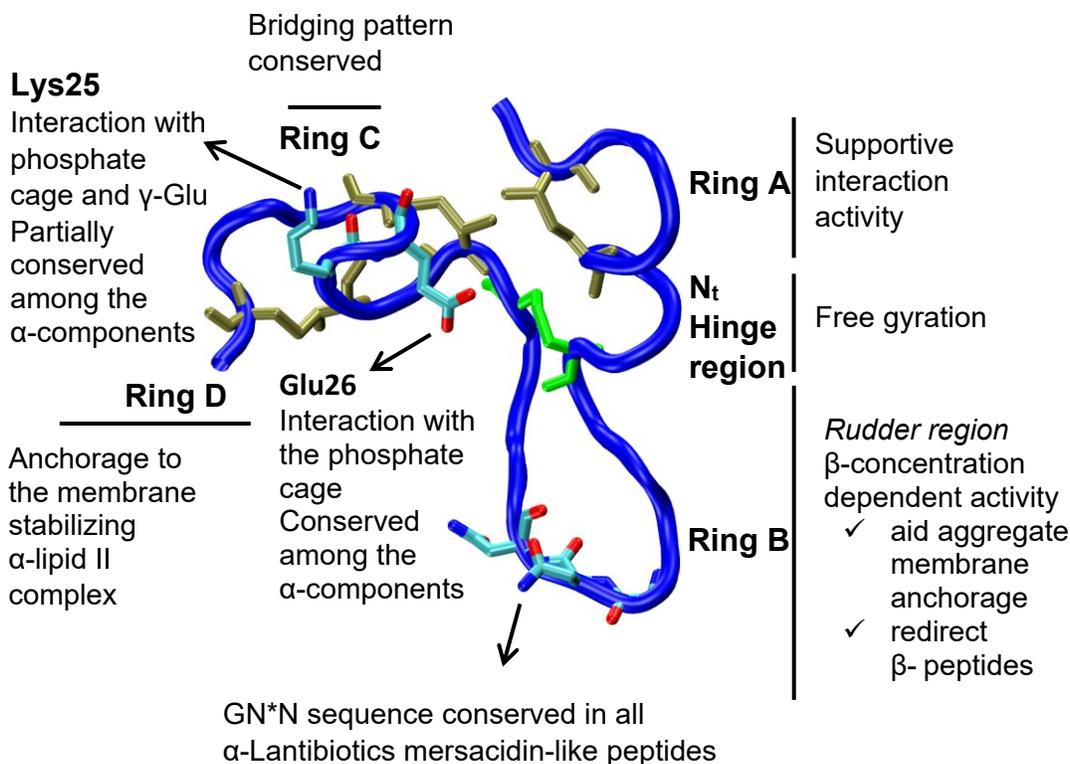
B.2.2.6 Concluding Remarks

After 1.3 μ s, from the analysis of the all-atom MD simulation, it is feasible to conclude:

- i. **Two-component Lichenicidin** (lch α and lch β peptides) **forms aggregates** of different size-stoichiometries which enclose ratios of: 4 α :2 β ; 5 α :5 β ; 2 β .
- ii. **Attachment to lipid II** component is only achieved by **α -peptide** while **β -peptide** plays the role of a **lipid II-recruiter**.
- iii. **Functional fragments** have been identified for the **β -peptide**, and bridging pattern similarities are related to those classified as mersacidin-like peptides (see section 8.1.1.1-B):
 - a. **N_t region and ring A.** Bridging pattern and sequence are variable among all β -components lantibiotic family. Hydrophobic residues promote the immersion in the membrane.
 - b. **Inter-rings α -helix region.** Interaction with α -partners. Predominantly originated from hydrophobic interactions.
 - c. **Rings B.** Bridging pattern conserved in the β -peptides of the two-lantibiotic family, used for membrane interaction.
 - d. **Ring C.** Bridging pattern conserved among all the β -component lantibiotics, main responsible of the aggregate stabilization on the membrane (Lys27).
 - e. **Ring D.** Bridging pattern conserved in the two β -component lantibiotic family, supportive membrane interaction (Arg31) for ring C.



- iv. **Functional fragments** have been identified for **α -peptide** and bridging pattern similarities are related to those classified as mersacidin-like peptides (see section 8.1.1.1-B):
- a. **N_t region and ring A supportive interaction activity.** Bridging pattern and sequence are highly variable among the α -component lantibiotics.
 - b. **Ring B-rudder region.** Bridging pattern conserved among the two component lantibiotic family. Sequence $-GN^*N-$ (section B.2.2.5) conserved among the α -peptides lantibiotics. Concentration β -dependent activity: (1) aid the aggregate anchorage to the membrane (2) redirect the β -peptides.
 - c. **Rings C and D.** Bridging pattern conserved among the two component lantibiotic family. **Ring D anchorage to the membrane surface** contributes to the **stabilization of α -lipid II complex**.



- v. **Role of Lys25. Residue partially conserved** among the α -peptides of the two-component lantibiotic. **Interacts with the phosphate moiety and with the γ -Glu of the stem chain.**
- vi. **Role of Glu26. Main responsible of the interaction with phosphate moiety.** Based on these results, it is reasonable to suggest that the **overall repulsive interaction** between the Glu26 (in the case of Lichenicidin) ring C and the Lipid II phosphate moiety perturbs the membrane surface, and consequently i) this residue is conserved in all two-component lantibiotic family

Chapter IV

9- Mammalian Cathelicidins

9. Mammalian Cathelicidins

The innate immune system constitutes the first line of defence against the attack and growth of microorganisms. It promotes the production of endogenous antibiotic peptides, in mammals, the cathelicidins among others.^[155]

The cathelicidins are a family of host defence antimicrobial peptides which have been found in a wide diversity of mammalian species. Despite the considerable heterogeneity of *inter*- and *intra*-species in the C-terminal region, which encodes the mature peptide, a high interspecies homology in the propeptide sequence denoted 'cathelin' domain has been found.^[156]

All mammalian cathelicidins have been in an early stage defined as *neutrophil-specific constituents*.^[157] Their abundance in these cells differs between and within species, being stored as granule subsets^[158,159]. From them, the cathelicidins are discharged extracellularly or in phagocytic vacuoles upon induced inflammatory or infectious response,^[156,160,161] confirming that the cathelicidins are also tissue-specifically expressed. Interestingly, cathelicidins have been also discovered to be produced in other types of cells.^[162,163]

9.1 Diversity, Structure and Classification

In mammalian species, around 30 cathelicidin family members have been discovered. Remarkably, in humans (namely LL-37^[164]), monkeys (RL-37^[165,166]), mice (mCRAMP^[162]), rats (rCRAMP^[163]), guinea pigs (CAP11^[159]), rabbits (CAP18^[167]) and dogs (K9CATH)^[168] just one cathelicidin has been isolated while other species such as horses^[169], pigs^[170–172], cattle^[173,174], buffalo^[175], deer^[176], sheep^[177–180], and goats^[177] showed a high variety of them. Additionally, cathelicidins have been also found in non-mammalian species like chicken^[181], snakes^[182], hagfish^[183] and trouts^[184] but they do not belong to the scope of this work. Typically, the cathelicidins are encoded by genes consisting of four exons. The first exon involves the sequence encoding the signal peptide, followed by the 2nd and 3rd exons which encode the cathelin domain 'part *pro*-' described by 99- 114 AAs.^[156] Concluding with the last 4th exon, which almost always encodes the mature peptide differing in sequence and length (figure 9.1.1).^[156]

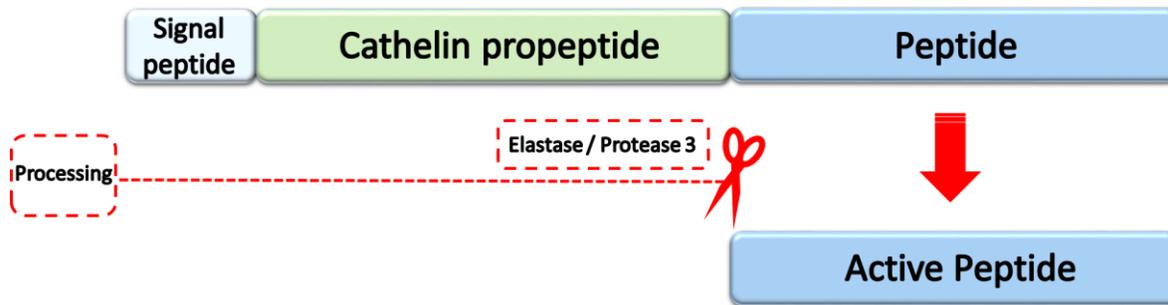


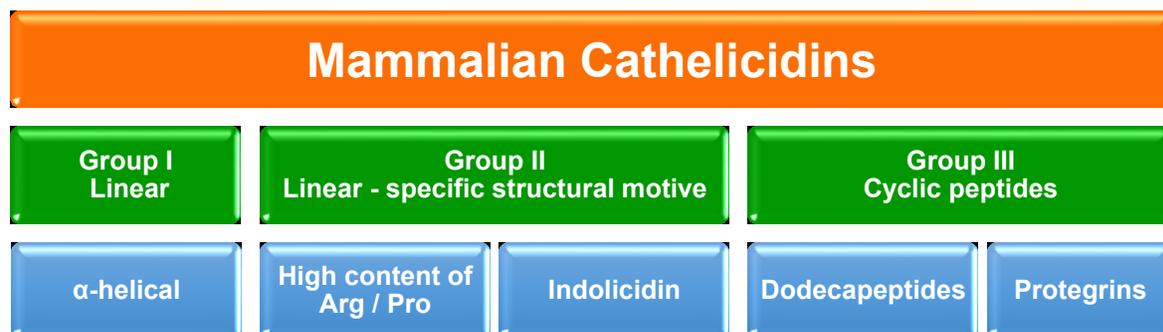
Figure 9.1.1: Simplified gene scheme representation of cathelicidins family. Active peptide is released by the serine proteases elastase or proteinase 3.

The *prosequence* is designated ‘cathelin’, due to its inhibitory activity against the cathepsin L (‘cathelin’ domain = cathepsin L inhibitor)^[185,186] which was for the first time identified from pig leukocytes. In principle, the main role of the cathelin region aids the biogenesis of the mature peptide, through the removal from the C-terminal domain, which is processed by the serine proteases elastase or proteinase 3.^[187,188] However, in humans the persistence of the uncleaved cathelin domain in infection/ inflammatory responses led to reconsider the significance of this part which should promote therefore, further studies on it.^[189]

Despite that many studies claim that the cathelin region is ‘highly conserved’ between the species, generally sharing a homology higher than 70% to cathelin^[189], some inconsistencies have been recently found. This issue has to be revised since cathelicidins from non-mammalian species just show few analogies.^[190]

As it was mentioned above, the active antimicrobial cathelicidin peptide encoded in the C-region varies in length (in a range of 12 to 100 AAs^[156,190]) sequence and function among the species. Most commonly, the cathelicidins peptides are amphipathic due to the predominant presence of basic residues and cationic overall charge at neutral pH. These peptides generally show coil structures in solution, while in lipid environments they tend to adopt α -helical conformations.^[157]

Despite the mentioned diversity, an initial four-group classification according to its structural features was carried out by R. Gennaro and M. Zanetti^[191] in 2000. Based on this prompt division, simplifications promoted by R. Bals and J. M. Wilson in 2003^[192], and revisited by Zannetti in 2005 followed (scheme 9.1.1):



Scheme 9.1.1: Latest Cathelicidin classification proposed.

- i. Group I: The most frequent, involve peptides with α -helical conformation, such as LL-37/ hCAP-18.^[193]
- ii. Group II: Linear peptides predominantly characterized by a single or a mixture of two AAs. Most commonly, arginine and/ or proline. For instance, Bac5, Bac7^[194] the first to be discovered, PR-39^[195] and the prophenins.^[196] In this group, the 13-residues peptide indolicidin^[197] is also found, characterized by a specific structural motive consisting of a high content of tryptophan (5 residues) and a regular spacer proline (3 residues).
- iii. Group III: Cyclic peptides involving the dodecapeptides^[198] and the protegrins.^[199] This group includes all those with disulphide bridges.

Evidently, according to this classification, different spatial configurations are found, contributing thus to their wide diversity. The combination of two factors i) the high conservation of the cathelin domain (in mammalian species) and ii) high sequence diversity, regarding the active peptide along the species, have caused several cross-examinations with still open questions. Mainly, the efforts have been principally focused on the justification of the evolutionary selection pressure to conserve the N-terminal section. Furthermore, the existence of at least one α -helical cathelicidin in all studied mammalian species suggested

that a peptide of this specific group was the promoter for the expansion of the family, generating a huge variety of peptides through several rounds of gene duplication.^[189]

9.2 Human Cathelicidin LL-37

The cathelicidin hCAP-18/ LL-37 was discovered by three different research groups at the same time in 1995, from myeloid lineage. Concisely, from myeloid bone marrow cDNA^[200–202] and from neutrophils^[201]. LL-37 is the lone cathelicidin found in humans and consequently, it has been extensively studied.

The name hCAP-18 denotes the estimated mass of the full-length polypeptide (~18 kDa) combined with the clarification of the cationic character of the C-terminal region (Cationic Antimicrobial Peptide). This nomenclature provides also information regarding the similarities found with the rabbit cathelicidin CAP18.^[189]

Alternatively, at the first stage the human cathelicidin was denoted as FALL-39, associated to the initially suggested four N-terminal residues (phenylalanine (F)-alanine (A)-Leucine (L)-Leucine (L)) of the mature peptide. Subsequently, after the isolation and clarification concerning the structure and sequence, the name was modified to LL-37. It referred to the active peptide of 37-AAs long consisting of two leucine residues in the N-terminal region. Currently, hCAP-18 is used for the designation of the propeptide while LL-37 represents the putative antimicrobial domain^[195] which is cleaved by proteinase 3 after exocytosis from neutrophils.^[161]

The hCAP-18/ LL-37 is expressed in several different regions of the body. A summary is displayed in table 9.2.1.

Organism	References
Neutrophils	[203]
Cells involved in inflammatory and immune responses	NK cells, B cells, $\gamma\delta$ T cells, monocytes [204] mast cells [205] hematopoietic cell and leukemia cells [206]
Skin and epithelial cells of the airways, mouth, tongue, oesophagus, intestine, cervix and vagina	[207–209]
Glands	Sweat and salivary [210]
Epididymis and testis	[211]
Body fluids	Wound fluid [212] Airway fluid [208] Seminal plasma [213]

Table 9.2.1: Plethora of body regions where the peptide has been detected to be expressed.

It has been also reported that hCAP18/LL-37 possess a broad-spectrum of pleiotropic properties such as antimicrobial activity against bacteria, viruses, fungi and parasites. Furthermore, it has also shown cytotoxic effects against eukaryotic cells usually at higher concentrations compared with the bacterial activity.^[214]

9.2.1 Antimicrobial Activity of Cathelicidin LL-37

The antimicrobial peptide LL-37 is obtained after the cleavage of the propeptide hCAP-18. This peptide is highly charged (16 charged residues) and it is characterized by a +6 charge at physiological pH showing an amphipathic α -helical configuration. Furthermore, it is active in a micro- and sub-micromolar concentration against a wide spectrum from Gram-negative to Gram-positive bacteria.^[215] Interestingly, it has been also reported that the combination *in vitro* of LL-37 with α - β - defensins provide an enhanced antimicrobial efficiency. These studies contribute for the understanding of the work-flow employed for our immune system against microbial invasion.^[189]

Additionally, it has been also proven its functional stability under different physiological conditions, highlighting its activity even under high salt media concentrations

(up to 150 mM NaCl)^[189,215] Indeed, LL-37 is able to inhibit the formation of bacterial biofilm even at low peptide-concentrations.^[216,217] Conversely, its antimicrobial efficiency results altered in patients with cystic fibrosis.^[218]

9.2.1.1 Mechanisms of Antimicrobial Activity of Cathelicidin LL-37

For the α -helical AMPs, a general transition pattern from a coil structure in solution to a α -helical conformation upon interaction with the lipid membranes has been reported.^[189] Interestingly, LL-37 is characterized by a disordered structure in aqueous solution while at physiological buffer or upon interaction with the membrane surfaces, it shows a consistent α -helical conformation.^[214,219] This steady conformation was explained by Oren *et al.* 1999^[220] as consequence of the presence of a bipolar character in the peptide sequence (hydrophobic head combined with charged residues along the peptide) which promote i) intermolecular salt bridges and ii) intrapeptide interactions allowing the aggregation. Additionally, the presence of the surrounding anions, such as chloride, at specific positions in the structure can act as quenchers for the stabilization of the α -conformation.^[219]

Moreover, a carpet-like mechanism has been proposed to be used by LL-37^[221] when aggregating over the lipid surfaces. Promoting, afterwards the leakage through the formation of suggested toroidal pores.^[222]

9.2.2 Pleiotropic Effects of LL-37

Known for its antimicrobial activity, additional biological activities have been reported for LL-37. The peptide is involved in a widespread variety of intricate mechanisms, highlighting for instance the specific activation of cell receptors or intracellular targets. Certainly, the property of inducing the activation of a diverse set of cells is directly linked to its physicochemical properties, which include the overall positive charge and amphipathicity.^[223]

The additional processes where the LL-37 has shown an important role are briefly summarized in table 9.2.2.1. These fascinating properties do not belong to the scope of this work and consequently literature where to refer is provided.

Properties	References
Stimulation of angiogenesis	[224,225]
Mobilization of Mesenchymal stromal cells, MSCs	Mast cells recruitment/ activation at inflammatory sites ^[207–209]
Epithelial wound healing	[225]
Immunomodulation	[226,227]
Tissue-specificity in cancer: Antitumor activity Apoptosis induction in cancer cells Protumorigenic	Antitumor ^[228–230] Protumor ^[218,231,232]

Table 9.2.2.1: Summary of the pleiotropic properties of LL-37

9.3 Cathelicidin LL-37 and its Fragments LL-32 and LL-20 by Means of MD Simulations

In humans, three main classes of AMPs can be identified: the defensins, the histatins and the cathelicidins^[190,233]. While several defensins and histatins have been discovered, only a single cathelicidin, the LL-37 has been identified so far. [Further details previous section 9.2].

The main focus of this section was to investigate the antibacterial activity of peptide LL-37 and its two fragments, namely LL-32 and LL-20. An interdisciplinary investigation was developed in order to study this unique cathelicidin from three different expertise: synthesis and in vitro experiments (AG Gutschmann, Forschungszentrum Borstel-Leibniz-Zentrum für Medizin and Biowissenschaften), SEIRA spectroscopy (AG Hildebrandt from Technische Universität Berlin) and in silico experiments (AG Mroginski).

In the following, the theoretical approach will be explained in detail and an overview of the spectroscopic measurements will be also provided. The in vitro experiments have not been concluded yet, which is why these results cannot be reported for this thesis. Nevertheless, the consistence of the spectroscopic measurements and the theoretical trajectories are remarkable.

With the aim of studying its antimicrobial mechanism of action, classical MD simulations of the cathelicidin LL-37 were carried out. Additionally, two fragments, known as

LL-32 and LL-20, which are respectively cutted-off at the residues 32 and 20 in comparison with the mother peptide, were also studied employing the analogous technique. Interestingly, the region involving the residues 17 to 29 in LL-37 is denoted as core peptide because it has been identified as the shortest antimicrobial and anticancer peptide to date.^[234] (see figure 9.3.1) Consequently, this region is being used for the design of novel peptides with improved properties such as the case of OP-145^[235] among others^[164,236]. According to the aforementioned facts, the combined analysis of the three peptides provides insights about how the execution of its antibacterial activity takes place, which will undoubtedly contribute to the development of new drugs.

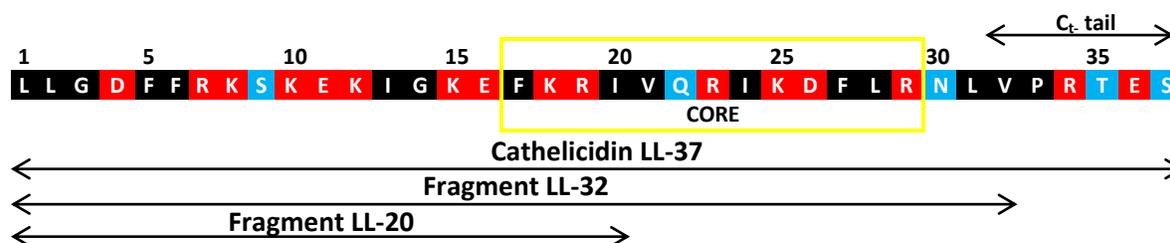


Figure 9.3.1: Amino acid sequence of cathelicidin LL-37 and fragments LL-32 and LL-20. Highlighted the core of the peptide (from residue 17 to 29). Color code used for the designation of the AAs: non-polar (black), polar non-charged (blue), polar charged (red) side chain residues.

Li *et al.*^[234] (2006), Porcelli *et al.*^[237] (2008) and Wang^[238] (2008) reported the 3D-structures of the LL-37 peptide complexed to SDS micelles, DPC micelles and to D8PG-mimicking a bacterial membrane-, respectively, by means of two- and three- dimensional NMR spectroscopy. All these studies agreed that an α -helical conformation is adopted by LL-37. Furthermore, depending on the nature of the interacting model membrane, the length of the helix varies: involving the residues 2 to 31 in SDS and D8PG, while in DPC compromises the residues 4 to 33. In the case of LL-37 complexed with DPC micelles, breakage of the α -helix was found at residue K12. However, the structure and dynamics of the isolated LL-37 peptide and its fragments in solution have not been investigated to date.

9.3.1 Conformational Dynamics of LL-37 and its Fragments LL-32 and LL-20 in Solution

9.3.1.1 Model Building and Set-up of the All-Atom MD Simulation of LL-37 and its Fragments LL-32 and LL-20 in Solution

The starting geometry for LL-37 was extracted from the Protein Data Bank (PDB-entry: 2K6O).^[238] Consequently, employing this structure as template, the initial configurations for the fragments were developed.

The three peptides case of study were modeled with a neutral N_t and an amidated C_t yielding an overall charge of +6, +6, and +4 for LL-37, LL-32 and LL-20, respectively. Furthermore, in order to reproduce the experimental conditions used by the collaboration partners, the protonation states of all ionizable groups were set according to the physiological pH 7.

Afterwards, a two-step protocol was performed in order to solvate and ionize the systems using the SOLVATE and IONIZED plugins of the VMD software^[26]: i) the peptides were individually solvated in cuboid boxes of TIP3P water molecules^[109] (of around $80 \times 50 \times 50 \text{ \AA}^3$ per system) and ii) in order to reproduce the experimental conditions, sufficient of ions were added to the solvated system mimicking an ionic strength of 100 mM of NaCl buffer. Subsequently, the systems underwent 20000 steps of energy minimization followed by a thermal equilibration at 300 K. Heavy-atoms were initially constrained, followed by unconstrained-MD simulations. The thermally equilibrated peptides in aqueous solution were simulated for 100 ns at 300 K in an NPT ensemble under constant atmospheric pressure and temperature using the Langevin Piston method.^[19] With the purpose of obtaining the equilibrium structure of the three peptides, the production run was done under periodic boundary conditions with extended electrostatics using Particle Mesh Ewald Summation^[12] and a cutoff distance of 12 Å for the Van der Waals interactions. SHAKE algorithm^[113] was used to constrain all bond lengths between heavy and hydrogen atoms assuring a time step of 2 fs.

9.3.1.2 Analysis of the Trajectories for LL-37 and its Fragments

The trajectories for the three peptides were analyzed by means of RMSD, RMSF, RG, dipolar moment together with the monitoring of the secondary structure. In general, the native behavior of the three peptides in buffer solution remains unaltered. The analysis of all aforementioned properties confirm the adoption of an α -helical conformation (see figure 9.3.1.2.1) which is consistent with experimental evidences.^[219]

- for the LL-37 and LL-32 peptides, the α -helical region comprises the residues 1 to 31 in perfect agreement with NMR observation, while the remaining residues in LL-37 form part of an unstructured VPRTES-tail (see Figure 9.3.1 and A-figure 9.3.1.2.1 for LL-37 and B-figure 9.3.1.2.1 for LL-32). As it is described in the section 3.7.3.1, the pink color is associated to an α -helical structure while the teal and white denote turns and coil structures, respectively.
- for LL-20, the α -helical motif contains the full peptide chain (C-figure 9.3.1.2.1)

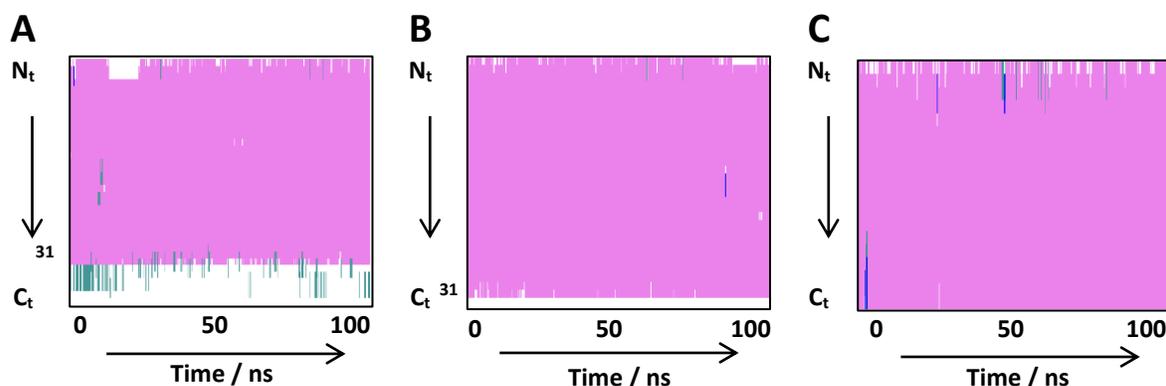


Figure 9.3.1.2.1: Monitoring of the secondary structure of the peptides in solution. A-LL-37. B-LL-32 and C-LL-20. Color code for the designation of the possible secondary structures: teal (turns), yellow (extended configurations), green (isolated bridges), pink (α -helix), blue (3_{10} helix), red (π -helix), white (coil).

A steady RMSD average value of the backbone heavy atoms (ca. 3 Å) was obtained for the LL-37 and for LL-32 while for LL-20 was lower (ca. 1.5 Å) (A-figure 9.3.1.2.2). Deviations from the stable behavior of the RMSD plots are mainly attributed to sudden bends of the α -helical spine. A similar trend between the peptides can be detected from the RMSF-plot, the core region (residues 17 to 29) showed higher C_{α} fluctuations. This fact was more pronounced in the mother peptide together with the longest fragment, while in LL-20, it is less pronounced. However, in the latest fragment, the C_{α} -fluctuations are shifted to the center of the sequence. (B-figure 9.3.1.2.2). Furthermore, the sequence length effects in the three variants had been also reflected in the global shape of the peptides, as it can be observed in the RG (C-figure 9.3.1.2.2). Comparatively, the peptides with a lack of the unstructured VPRTES-tail showed an overall regular compactness (C-figure 9.3.1.2.2: LL-32 orange line and LL-20 green line) – while the presence of the tail promote more degrees of freedom and consequently, higher RG values combined with more variability. (C-figure 9.3.1.2.2: LL-37 mauve color).

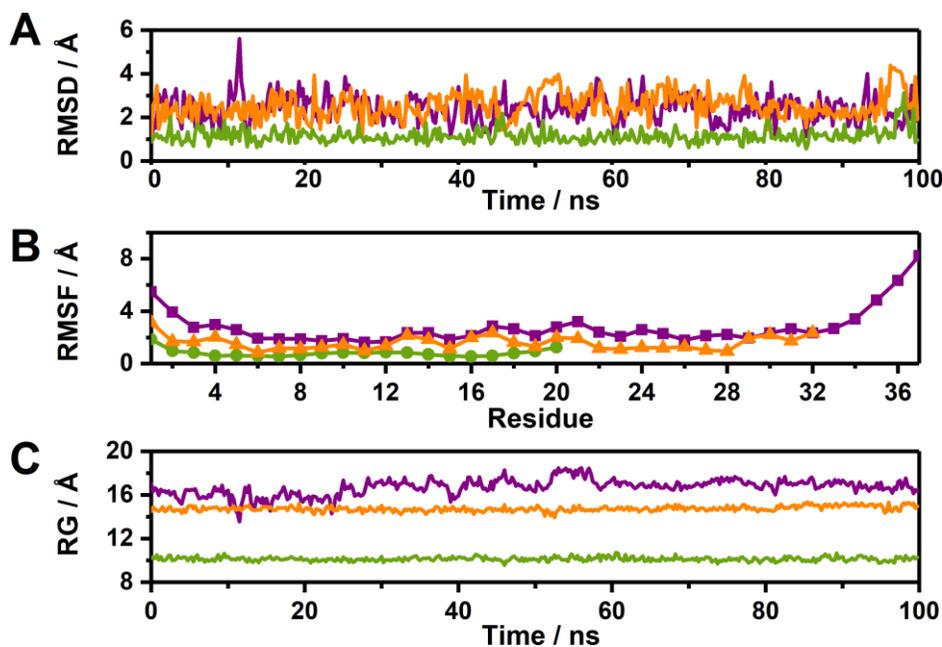


Figure 9.3.1.2.2: Analysis of the stabilization in water for the peptides by means of A-RMSD, B-RMSF, C-RG. LL-37 is represented in mauve while LL-32 and LL-20 in orange and green, respectively.

For LL-37 and LL-32, breakage of the α -helix is predicted at positions K8-K12, (figure 9.3.1.2.2: structures B, C, D, E, and G) in agreement with experimental results^[237] as well as at positions R19-R23 which are involved in the core (figure 9.3.1.2.2: A, B, F and H). Interestingly, these two hinge points are characterized by the presence of hydrophilic amino acids, S9 and Q22, which are immersed in positively charged sections K8-K12 and R19-R23, respectively (see figure 9.3.1). Furthermore, the rate of the α -helix bending is significantly lower in the LL-32 fragment which lacks the unstructured C_T-tail. In the case of the LL-20 fragment, breakage of the α -helix is only predicted at position K8-K12 (figure 9.3.1.2.2: structures I, J, K, and L).

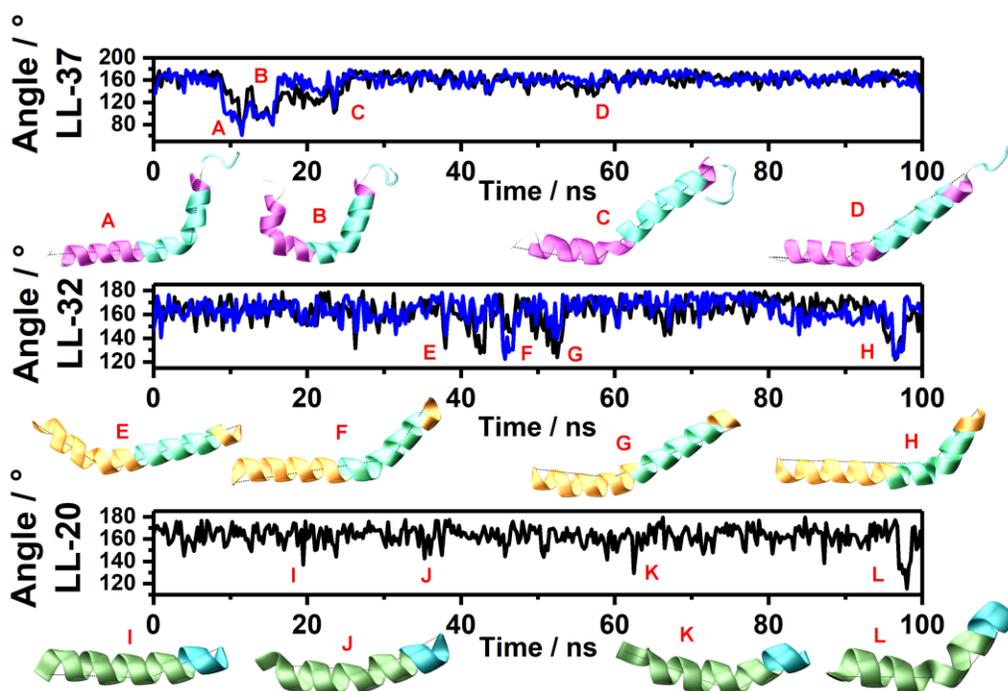


Figure 9.3.1.2.3: Graphical examples of the peptides bendings. Around L1-K12-V32 (black trace) and L1-R23-V32 (blue trace) in case of LL-37 and LL-32 and L1-K12-I20 for LL-20 (black trace).

Furthermore, despite that LL-37 and LL-32 show the same total charge (+6), their average dipolar moments are different: 256 Debye for LL-37 and 135 Debye for LL-32. This reflects a different orientation of the polar side chains. As it was expected, the LL-20 fragment showed the lowest dipolar moment of 50 Debye due to the decreased number of positively charged residues.

The formation of inter- and intra-molecular bridges have been suggested to be of important relevance for the stabilization of the secondary structure in physiological buffers, together with the fact that the presence of ions acting as quenchers play also an essential role.^[214,219] According to this, the dynamics of the salt bridge formation was also evaluated in the course of 100 ns of all-atom simulation. For this study, in addition to the mother peptide and the aforementioned fragments, a LL-20 mutant, namely S9V-LL-20, having serine 9 substituted by a valine, was also considered in order to enrich the evaluation of the dynamic role regarding the first twenty amino acids. This approach enabled the comparison with the experimental evidences of Wang. *et al.*^[239] with their LL-23V9. In the figure 9.3.1.2.4 the probability detected of the salt bridge formation during the simulation time between the side chain residues for each peptide is displayed. A-figure 9.3.1.2.4 shows a direct comparison regarding LL-37 and LL-32, in both peptides the salt bridges are built involving the side chains of residues D4-R7, D4-K8, E11-K8, E11-K15, E16-R19. Additionally, in LL-32 a salt bridge formation is detected involving the side chains of D26-R29, while for LL-37 two additional bridges are obtained between the side chains of R34-E36 and E36-K8. Interestingly, in both peptides no salt bridges are built in the core region. On the other hand, the salt bridge formation for the short LL-20 and mutant S9V-LL-20 remains very similar to the previous cases, observing the salt bridge formation between the side chain residues of D4-R7, D4-K8, E11-R7, E11-K8, E11-K15, E16-K12, E16-R19 (see B- for LL-20 and S9V-LL20). As it can be easily detected in the figure 9.3.1.2.4, the formation of the salt-bridges involve mainly the residues 1 to 19 for the four peptides with the contribution of the residues D26, R29, R34 and E36 for LL-37 and LL-32. Furthermore, by comparison the LL-20 and its mutant showed very similar behavior and no significant dynamic differences were found (see-Appendix – Cathelicidins). Very interestingly, through the dynamics of these peptides, it is feasible to detect the lack of salt bridges in the core region, which suggest the presence of unique properties in this region.

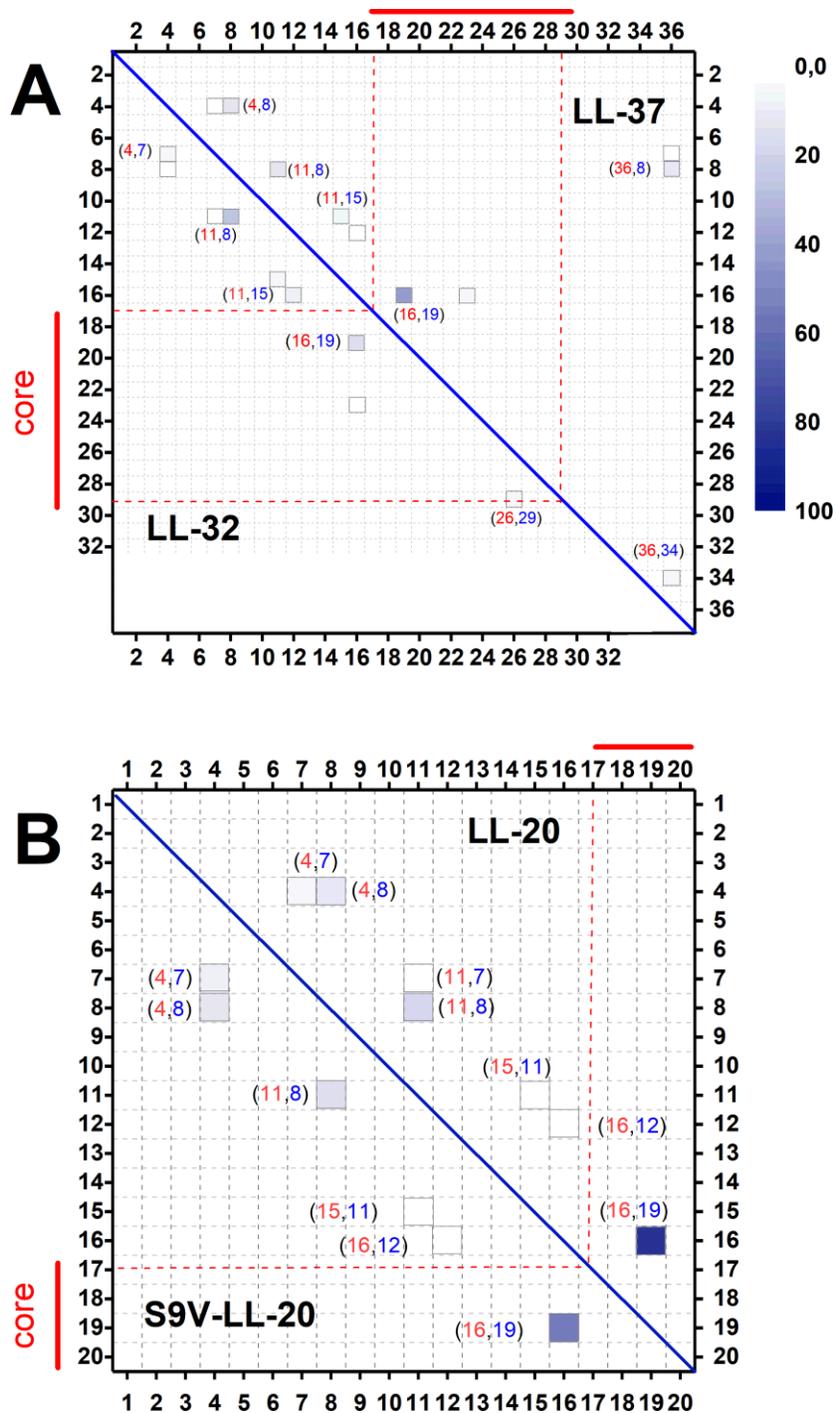


Figure 9.3.1.2.4: Evaluation of salt-bridge formation for mother peptide and derivatives. **A-** Probability of the side chain salt bridges formation regarding LL-37 and LL-32 in the course of the simulation. **B-** Probability of the side chains salt bridge formation for LL-20 and S9V-LL20. Most of the salt brigdes are found involving the first 19 AAs and the C_t region.

9.3.2 Theoretical Studies of LL-37 and its Fragments in Biologically Relevant Environments

9.3.2.1 Model Building and Set-up of the All-Atom MD Simulation of LL-37 and its Fragments LL-32 and LL-20 with Nearby Lipid Membranes

As it was previously mentioned, despite the availability of the 3D-dimensional structures of the LL-37 complexed to various lipid bilayers^{[234][237][238]}, the mechanism of action and dynamics of peptide attachment to the membrane surface has not been explored at atomic resolution. In order to do so, the individually equilibrated peptides were used for investigating the antibacterial activity in a more complex simulation system including a lipid bilayer. This section is subdivided in two parallel studies i) approaching of monomer peptides to the membrane surface and ii) aggregation evaluation in the lipid environment.

In agreement with the experimental setup, a biological membrane was modeled by a mixture of 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2oleoyl-sn-glycero-3-phosphoglycerol (POPG) molecules in a 90:10 proportion. The POPC/POPG bilayer with a surface area of 80 x 80 Å² was constructed with CHARMM-GUI.^[114]

- i. With the purpose of studying the individual peptide-membrane interactions, three systems were developed. The three peptides were initially placed 10 Å above the lipid bilayer (see A- figures 9.3.2.2-A.3-4). The membrane-peptide systems were solvated in 80 x 80 x 400 Å³ TIP3P box and ionized following the same guideline explained in section 9.3.1.

In parallel,

- ii. Two additional peptide-membrane models were constructed with either two LL-37 peptides or two LL-32 peptides placed 10 Å above the bilayer surface and 20 Å from each other in a similar manner as in the previous systems (see figure 9.3.2.2-B.1 and 9.3.2.2-B.1-3). The subsequent preparation steps were in agreement with the previous protocols.

For both i) and ii) studies, the production run was extended until 300 ns in the NPT ensemble under constant atmospheric pressure and temperature using the Langevin Piston method and 300 K, together with a constant cell size along the surface area.^[19] The strength of the peptide-membrane interaction was evaluated by computing the number of contacts. A contact between two atoms is defined by an interatomic distance below 3 Å.

9.3.2.2 Analysis of the Trajectories for LL-37 and its Fragments LL-32 and LL-20 with Nearby Lipid Membranes

9.3.2.2-A Monomers in a Lipid Environment

The length of the peptides promotes distinct dynamic behavior and adsorption properties. The figure 9.3.2.2-A.1 shows the monitoring of the contacts interaction of the three peptides with the membrane surface. As it can be observed in the figure 9.3.2.2-A.1, the fragment LL-20 showed a clear sign of desorption (green trace) which is consistent with three main factors:

- Conservation of just four residues of the peptide core (see figure 9.3.1).
- Involvement of the charged residues in salt bridges (see figure 9.3.1.2.4).
- For entropic reasons, the binding to any surface decreases, when decreasing the peptide length.^[240]

In contrast, the increase of contacts of LL-37 in the course of the simulation time was significantly notorious (figure 9.3.2.2-A.1- mauve trace). In parallel, it was also determined that once fragment LL-32 interacted with the surface, a regular steady behavior was maintained until the end of the simulation (figure 9.3.2.2-A.1- orange trace).

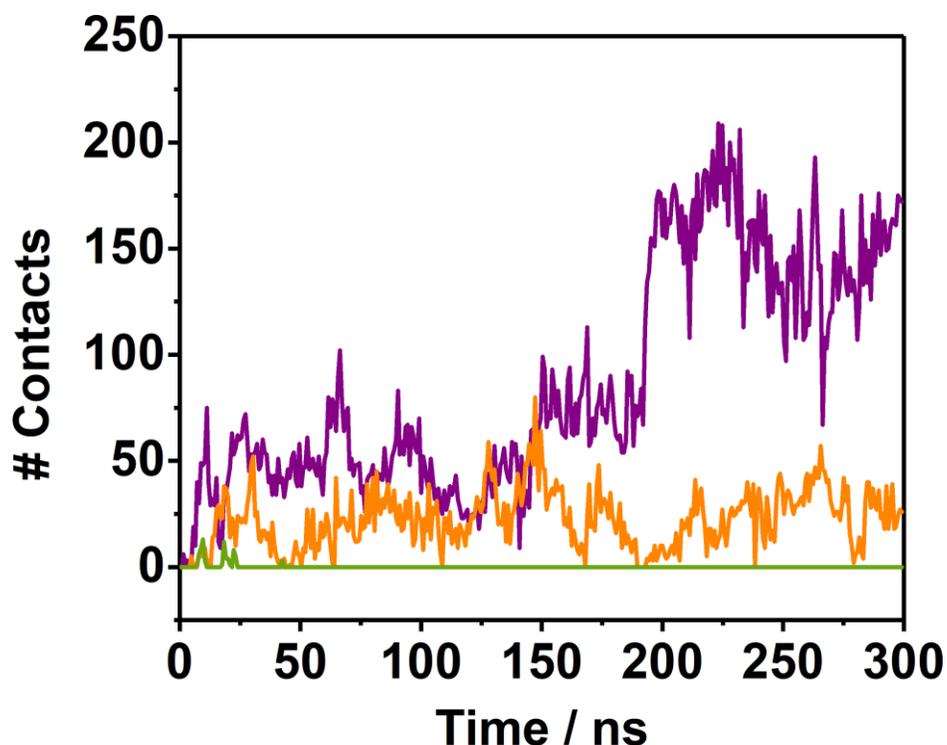


Figure 9.3.2.2-A.1: Monitoring of the monomers-membrane interaction during 300 ns of all-atom MD simulation. LL-37(mauve line), LL-32 (orange line) and LL-20 (green line). LL-37 and LL-32 showed different interactions pattern with the membrane surface while LL-20 provided clear desorption.

Peptide-membrane interaction pattern was evaluated by means of the COM-distances. The figure 9.3.2.2-A.2 shows the monitoring of the COM distance between peptides and membrane during 300 ns of all-atom MD simulation. A different behavior in the surrounding lipid environment is adopted by LL-37 and LL-32. While the LL-32 remains stable over the surface (orange trace) interacting with it through a modest number of contacts (Table 9.3.2.2-A.1), LL-37 immerses itself into the membrane. For LL-37, the distance between COM's $_{\text{pep-mem}}$ is $\sim 5 \text{ \AA}$ shorter than the average distance from the center to the surface of the membrane (20 \AA). Furthermore, the number of contacts between peptide and lipids is five-fold larger than the computed for the LL-32 fragment. Interestingly, in this simulation time the adsorption of LL-32 on the surface is mainly produced by electrostatic interactions, while for the LL-37 the stronger interactions with the surface are a consequence of the combined electrostatic and hydrophobic forces (Table 9.3.2.2-A.1). This fact is in agreement with the experimental findings reported by Henzler-Wildman *et al.*^[241]

who determined that both electrostatic and hydrophobic forces may affect the LL-37 depth insertion.

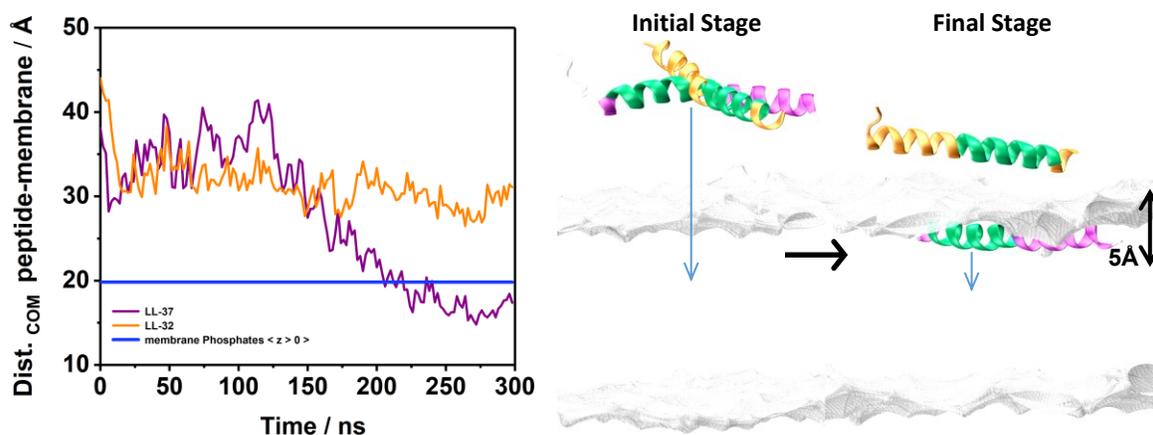


Figure 9.3.2.2-A.2: Monitoring of the distance between COM's of peptide and membrane in the course of 300 ns of all-atom MD simulation. Immersion of LL-37 (mauve trace) in the POPC/POPG while LL-32 remains over the surface (orange trace), blue line denotes the z- coordinates average of the phosphate heads (upper layer) of the membrane as reference point for immersion. Side chains of both peptides are not displayed for clear visualization.

Monomer-Surface	Hydrophobic Int.	Electrostatic Int.
LL-32	5	32
LL-37	120	90

Table 9.3.2.2-A.1: Monomers-membrane interaction type.

As it can be followed in figure 9.3.2.2-A.3, the insertion of the LL-37 into the membrane takes place stepwise. First, the core section (F17-R29) approaches the anionic membrane using the R23 as anchor (B-figure 9.3.2.2-A.3). Secondly, the R34 promotes the transient attachment of the C_T-VPRTES-tail (C- D- E figure 9.3.2.2-A.3) which destabilizes the initial arrangement and induces the immersion of the N_T-hydrophobic head of LL-37 in the membrane (D-figure 9.3.2.2-A.3). The process exemplified in C- E figure 9.3.2.2-A.3 is subsequently repeated leading to a parallel arrangement of the peptide to the membrane surface (as it has been suggested by Sood *et al.*^[221]) After 300 ns of MD simulation, the peptide is completely immersed in the membrane with its hydrophilic side and its dangling C_T-tail (F-figure 9.3.2.2-A.3) pointing towards the outside. The membrane immersion

detected for LL-37 are in line with the fact that the peptide is able to disrupt the acyl chain membrane packing.^[221]

These atomistic evidences contribute to the understanding of the open questions regarding the mechanism used by the peptide in order to exert its bactericidal activity. Several experimental studies have suggested that the peptide follows a carpet mode to exert its antimicrobial activity, while others have reported experimental observations which can be ascribed to a toroidal pore mechanism.^[220,241–244] These two proposed mechanisms adopted by LL-37 have been also associated to the nature of the membrane surface, the conditions of growth etc. According to G. Wang *et al.*^[244] both pathways could represent different stages of the peptide behaviour. The trajectories (and the experimental measurements) obtained in this work are consistent with the carpet model mechanism. It is also important to mention, that the perpendicular arrangement of LL-37 with respect to the surface observed during the simulation time (D-figure 9.3.2.2-A.3) provides insight for the understanding of the steps required for pore generation, if sufficient quantity of peptides are involved, promoting the transition from a parallel to a transversal reorientation.

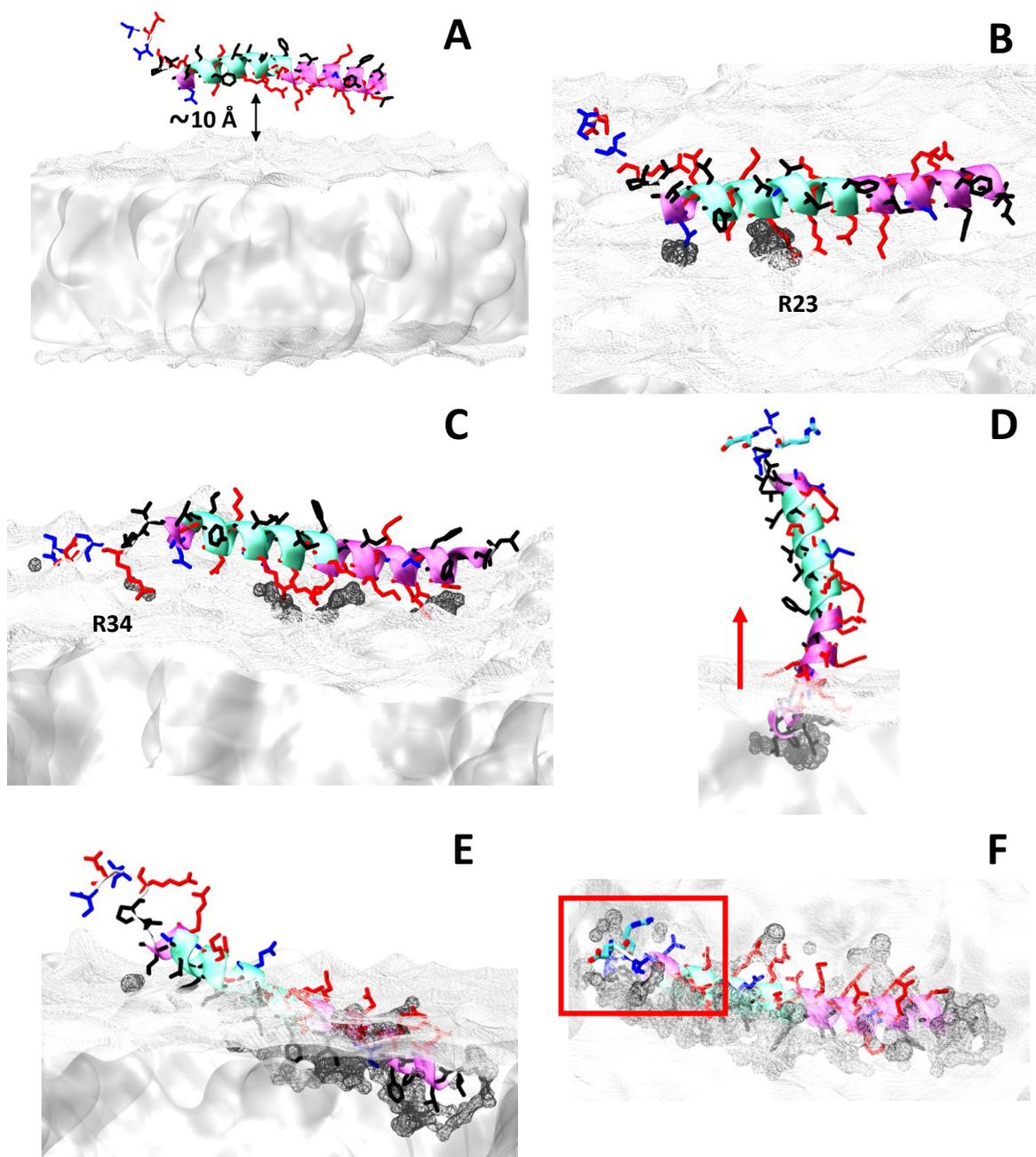


Figure 9.3.2.2-A.3: Snapshot of the monomeric LL-37-membrane interaction.

In the case of the LL-32, the MD simulations predict the adsorption of the peptide on the surface of the membrane but not its insertion (figure 9.3.2.2-A.4). Analogous to LL-37, the core residues of LL-32 (F17-R29) steer the interactions with the surface using the R23 as anchor (B-figure 9.3.2.2-A.4). However, after accommodating itself on the top of the lipid bilayer (30 Å away from the center of mass of the membrane) (C-figure 9.3.2.2-A.4), the peptide remains stable generating a mean of 37 contacts with the surface of POPC and

POPG lipid (table 9.3.2.2-A.1). Unlike the LL-37, the insertion of the LL-32 into the membrane is not predicted.

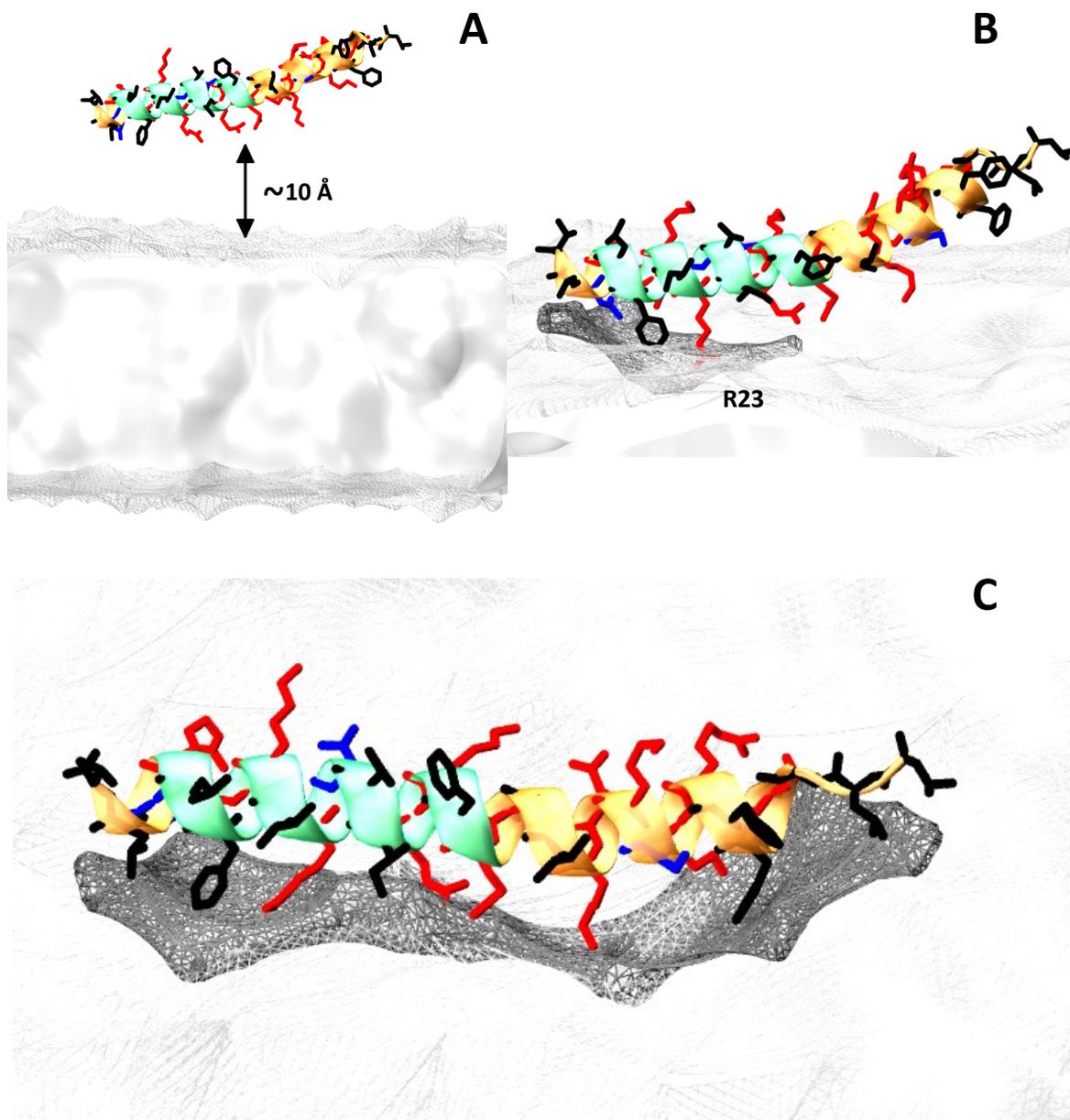


Figure 9.3.2.2-A.4: Snapshot of the monomeric LL-32-membrane interaction. Approaching to the lipid membrane and stabilization on the membrane surface.

9.3.2.2-B Dimerization Studies for LL-37 and Fragment LL-32

According to experimental studies, LL-37 self-assembles and forms oligomers^[219,241] although the order of oligomerization decrease when in contact with the membrane^[219]. To investigate this issue in greater detail a dimerization course was studied by means of MD simulations.

Atomistic details of peptide oligomerization were achieved through the analysis of the MD trajectories. From the beginning of the simulation the two LL-32 peptides try to interact via hydrophobic interactions (table 9.3.2.2-B.1) and remained in an antiparallel arrangement until the end of the simulation (see B-figure 9.3.2.2-B.1). In order to highlight the approach between the peptides, the interactions with the membrane surface are not depicted. The presence of the tail promotes a differentiated dynamical behavior when comparing LL-32 and LL-37. (figures 9.3.2.2-B.1 and 9.3.2.2-B.3).

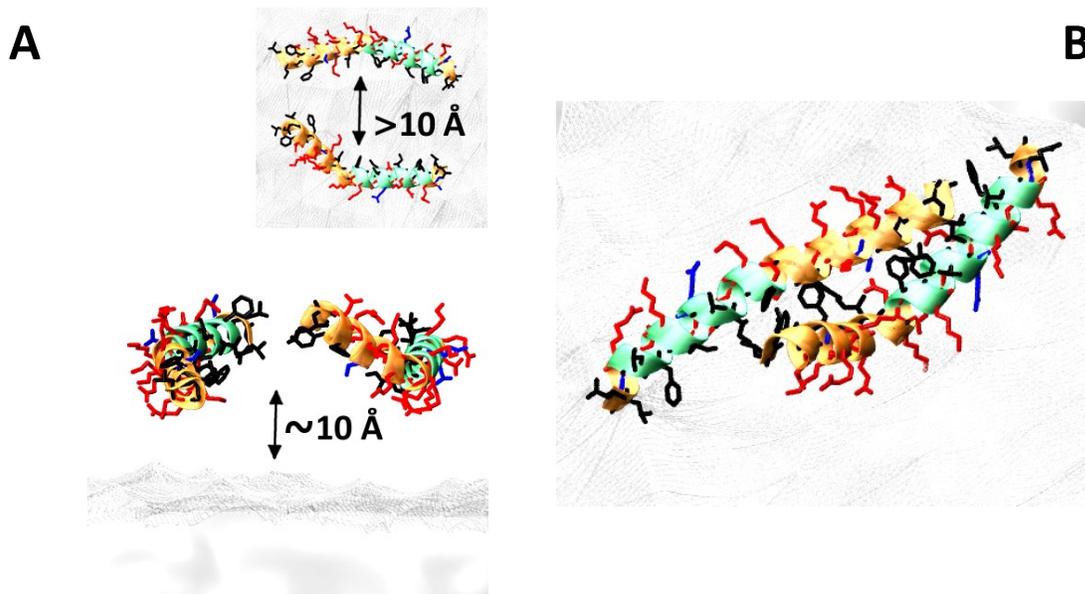


Figure 9.3.2.2-B.1: LL-32-dimerization process fragment over the membrane surface. Interactions with the membrane surface are not displayed for clarification. Side chain peptide color code detailed in figure 9.3.1

As observed in the figure 9.3.2.2-B.2, most of the contacts detected between these two fragments take place through the hydrophobic residues (black trace) in agreement with

the Hofmeister effect or ‘salt-out’, shielding the hydrophobic face.^[214] The dimerization process is relatively fast because it is produced during the first 100 ns of simulation.

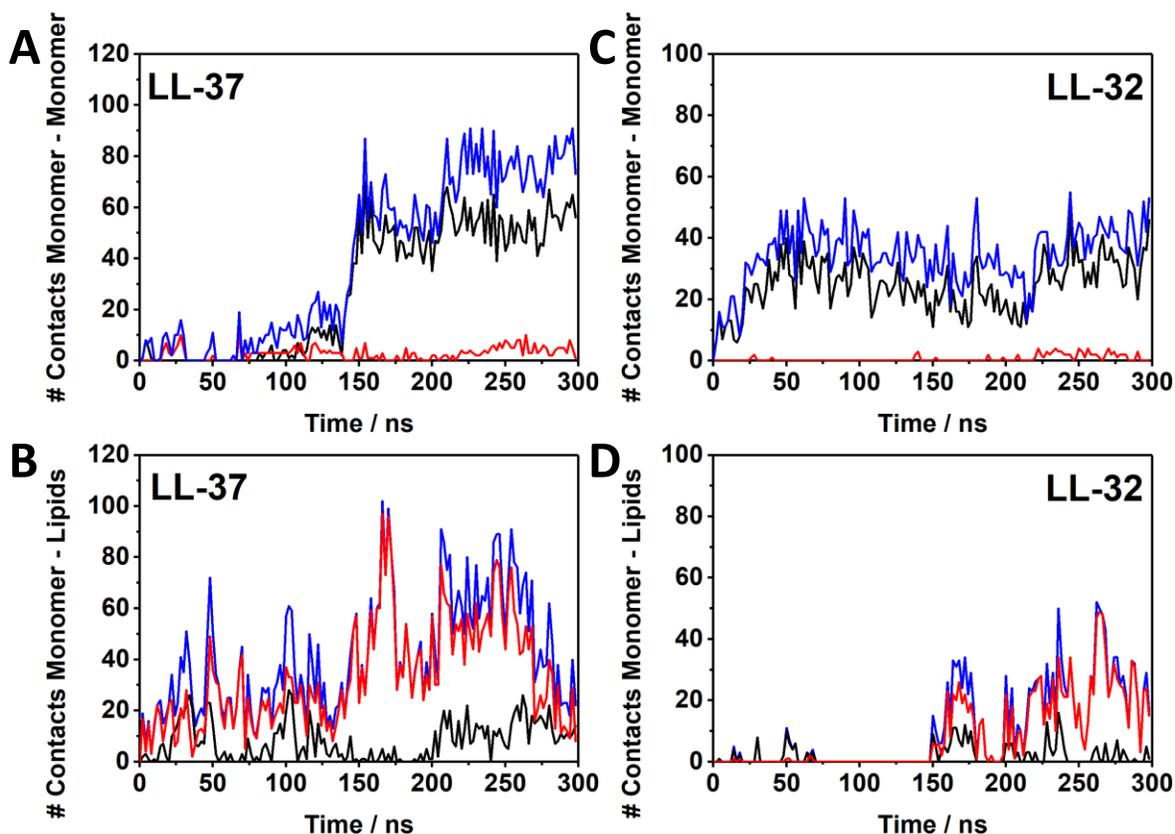


Figure 9.3.2.2-B.2: Monitoring the number of contacts between the monomers and the peptides with the membrane surface for LL-37 and LL-32. Total number of contacts (blue traces), involving only hydrophobic residues (black traces), only hydrophilic residues (red traces). **A-** LL-37-Monomer-monomer contacts. **B-** LL-37 monomer-lipid membrane contacts. **C-** LL-32-Monomer-monomer contacts. **D-** LL-32 monomer-lipid membrane contacts.

Parallel interactions produced with the surface of the membrane, which simultaneously allocates the peptides, and the lack of the C_T-VPRTES-tail, promote an overall steady hydrophobic interaction between the peptides (see figure 9.3.2.2-B.2). However, the interactions produced between the monomers differ from those detected for LL-37 dimer formation (next section). Table 9.3.2.2-B.1 summarizes the average number of contacts between both monomers over the last 100 ns of MD simulation. Noteworthy, from the total number of contacts (40), 28 are of hydrophobic nature. Interestingly, the interaction with the surface fades once the dimer is formed (see figure 9.3.2.2-B.4), as reflected by the

reduced number of contacts to 12. After 300 ns of MD simulation, the fragment LL-32 builds an antiparallel dimer via hydrophobic inter-monomer interactions while the electrostatic interactions keep it in contact with the lipid surface.

Dimer	Monomer-Monomer			Monomer Unit-Surface	
	Hydroph. Int.	Electrost. Int.	Mixed	Hydroph. Int.	Electrost. Int.
LL-32	28	2	10	1	11
LL-37	54	4	23	6	22

Table 9.3.2.2-B.1: Average number of contacts between both monomers over the last 100 ns of MD simulation

Similarly, to LL-32, LL-37 dimerizes (see figure 9.3.2.2-B.3). Dimerization represents the first step for the suggested up-to-hexamer aggregate built by LL-37.^[219] In this process, the role of the C_T-VPRTES-tail can be also determined, which is used to reach the reaction partner without losing the contact with the membrane surface. This process is achieved through i) the anchoring-role of R34-monomer *i* to the surface of the membrane while ii) in parallel, E36-monomer *i* built a salt bridge with R34-monomer *j* (see B-figure 9.3.2.2-B.3). This important stage is predominantly of electrostatic nature and it is detected within the first 150 ns of simulation (table 9.3.2.2-B.1 and figure 9.3.2.2-B.4). The transient salt bridge between the VPRTES-tails of the two LL-37 monomers does not only recruit peptides, but it holds them together until energetically favorable hydrophobic interactions between monomers are produced, in line with the Hofmeister effect reported experimentally^[214] and in contrast to the results reported by Shahmiri *et al.*^[245] who suggested that salt-bridges stabilize the aggregate formation. The latter reported that the salt bridge formation involved the amino acids in the N- and C-termini and around the residue D36 in LL-37. Interestingly, LL-37 is characterized by the presence of a glutamic acid at the 36th position (E36) instead of an aspartic acid (D).

From the MD simulations, it can be demonstrated that once that the dimer is stabilized, the R34-E36 salt bridge is broken, enabling thereby the R34 to interact with further partners such as a third peptide. In this system, both monomers acquire a parallel orientation to each other and the intermonomer interaction is strengthened, as it is reflected by the interaction energy plots of LL-37 compared to the LL-32 dimer. Comparatively, the interaction nature remains identical.

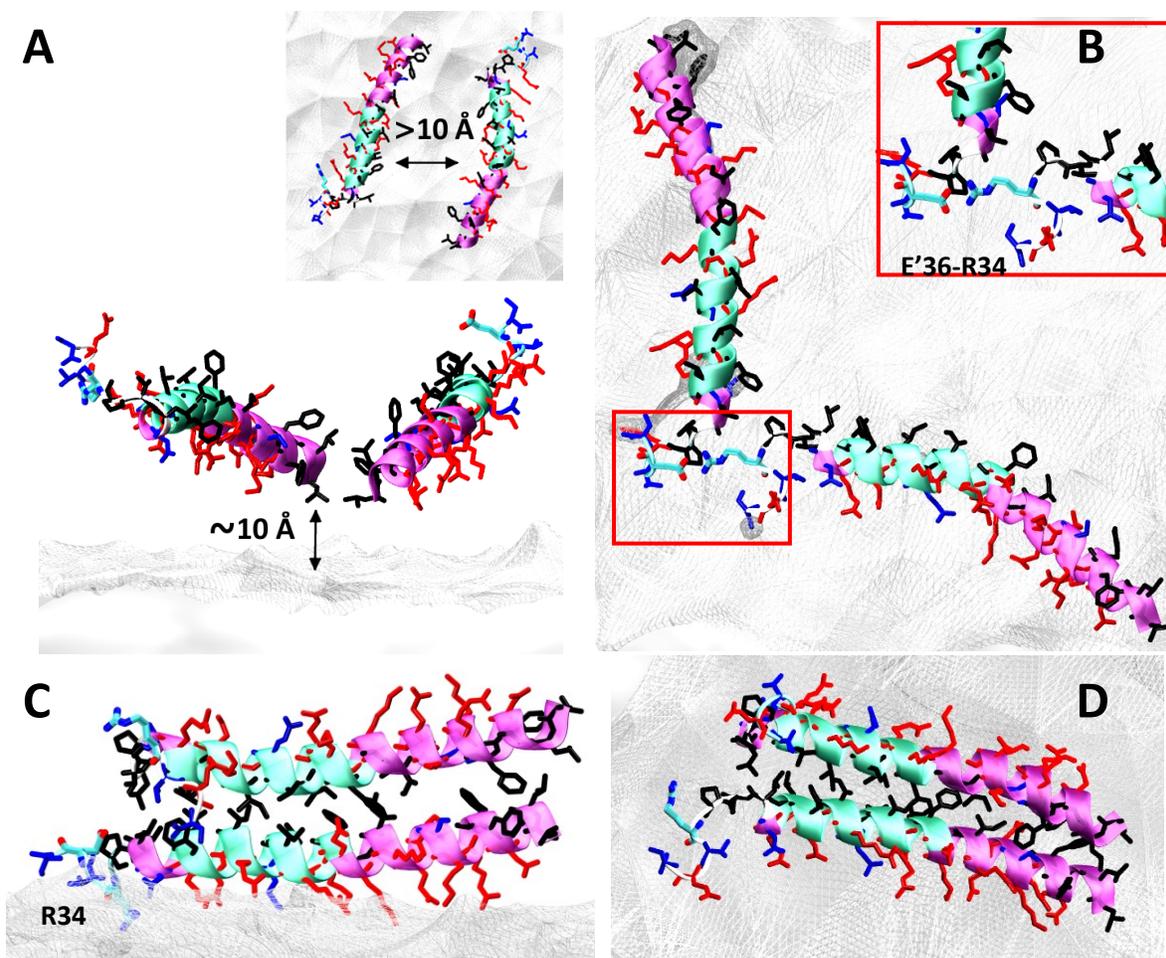


Figure 9.3.2.2-B.3: LL-37-dimerization process over the membrane surface. Interactions with the membrane surface are not displayed for clarification. Side chain peptide color code detailed in figure 9.3.1

In conclusion, the driving force for the self-association of the peptide is based on the Hofmeister effect, consisting on the generation of a hydrophobic core shield of the amphipathic helix from the aqueous environment. This process is aided by the VPRTES-tail of LL-37, which promotes an efficient and ordered oligomerization of cathelicidin on the surface of the membrane.

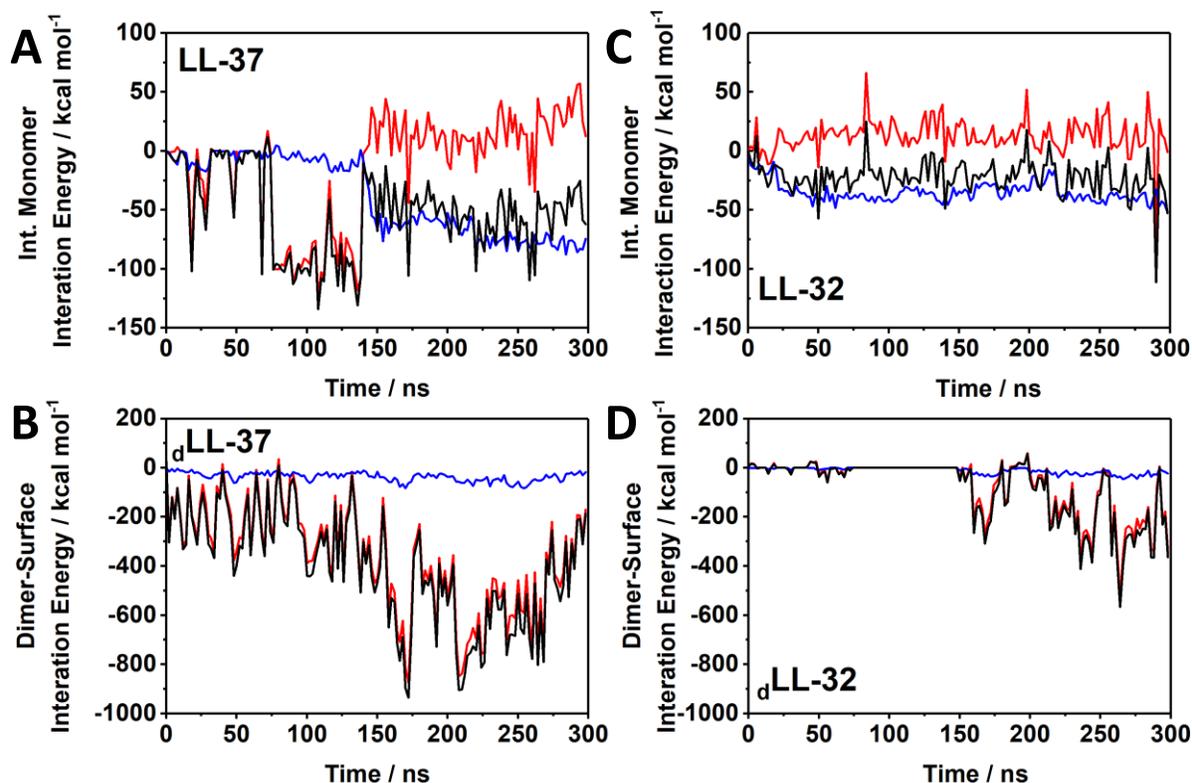


Figure 9.3.2.2-B.4: Evolution IE for LL-37 and LL-32. **A-** IE between two LL-37 monomers. **B-** IE between LL-37-dimer-membrane. **C-** IE between two LL-32 monomers. **D-** IE between LL-32-dimer-membrane. Black traces represent total energy, blue traces represent Van der Waals contribution and red traces the electrostatic contribution to the corresponding total energies.

9.3.3 Agreement with the Experimental Measurements

LL-37 and its fragments were studied by a combined spectroscopic approach. The second method, carried out by Msc. Enrico Forbrig, was based on two types of infrared spectroscopic experiments:

- Fourier transform infrared (FTIR) transmission experiments of the three peptides in buffer, to be used as reference.
- Incorporation of the peptides into the tBLM with a negatively charged POPC:POPG (90:10) membrane. This step was monitored to track the spectral changes associated with the peptide-membrane interaction.

9.3.3.1 Transmission IR Experiments

The figure 9.3.3.1.1 shows the transmission IR spectra of the amide region of LL-37, LL-32, and LL-20 in buffer. Interestingly, LL-37 and LL-32 show very similar spectra, in which the amide I and amide II regions can be modeled efficiently with four and two Lorentzian functions, respectively. The most pronounced bands are located at ca. 1652 cm^{-1} and 1550 cm^{-1} . These bands can be ascribed to the amide I and II vibrations, respectively, which are in line with an α -helical secondary structure of both peptides (A-B-figure 9.3.3.1.1 for LL-37 and LL-32, respectively).^[246] The additional bands at 1682, 1666, 1637-1634 cm^{-1} and at 1531 cm^{-1} can be related to vibrations of amino acid side chains (Appendix Cathelicidins). Concisely, the peptides contain a high percentage of ARG and LYS AAs together with the presence of polar charged and non-charged residues of ASP and ASN, GLN respectively. Consequently, the following explanation can be extrapolated for all systems:

- The guanidinium group of ARG gives rise to very intense asymmetric and symmetric vibrations at 1695 – 1652 cm^{-1} and 1663 – 1614 cm^{-1}
- LYS shows an asymmetric and symmetric ammonium vibration at ca. 1630 cm^{-1} and 1530 cm^{-1} .
- The amide bonds of ASN and GLN absorb at ca. 1680 cm^{-1} and in the range around 1610 cm^{-1} , as well as the asymmetric COO^- mode of GLU and ASP that can be encountered almost in the entire amide region, depending on its specific interaction.

Also, the amide I/amide II ratios are similar with values of 2.1 ± 0.2 and 2.0 ± 0.1 for LL-37 and LL-32, respectively. The spectrum of LL-20 is of much worse signal-to-noise ratio due to the reduced number of charged side chains residues and the associated decreased solubility (C-figure 9.3.3.1.1). However, similar vibrations can be identified:

- Amide I and II at 1649 cm^{-1} and 1556 cm^{-1} .
- Only the amino acids ARG, LYS, GLU, and ASP can contribute to the spectrum: depicted by the amino acid residue vibrations at ca. 1681, 1666, 1634, and 1531 cm^{-1} .

Herein, the amide I and II are less pronounced due to the shorter sequence of the peptide and thus the amide I/amide II ratio is slightly decreased to 1.8 ± 0.1 .

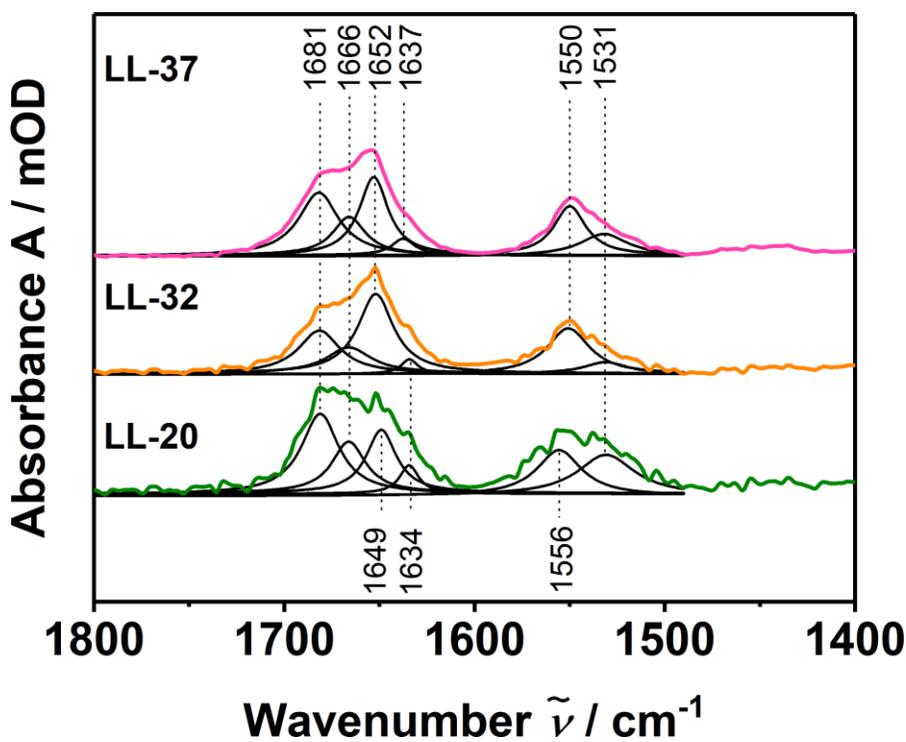


Figure 9.3.3.1.1: Transmission IR spectra of the amide region of LL-37, LL-32, and LL-20 in buffer.

9.3.3.2 Surface-Enhanced IR Spectroscopy (SEIRA) Experiments Using Tethered Bilayer Lipid Membranes (tBLMs)

With the aim of studying the peptide-membrane interactions of LL-37 and its fragments LL-32 and LL-20 in a negatively charged POPC:POPG (90:10) membrane, a tethered bilayer membrane system (tBLM)^[37] was constructed as described in the experimental section. Secondly, POPC:POPG (90:10) liposomes were spread on top of the SAM layer leading to a lipid layer of alternating lipid bilayer and lipid monolayer areas.

The SEIRA difference spectrum of the POPC:POPG (90:10) membrane taking the spectrum of the mixed SAM as reference is shown in the figure 9.3.3.2.1. The successful

tBLM formation can be verified by the vibrational modes with positive intensity, which are characteristic for lipids that are deposited onto the SAM, i.e. the $\nu(\text{C}=\text{O})$ stretching vibration at 1738 cm^{-1} and the symmetric (s) and antisymmetric (as) vibrational CH_n bands ($\nu_s(\text{CH}_2)$, $\nu_{as}(\text{CH}_2)$, and $\nu_{as}(\text{CH}_3)$) at 2854 cm^{-1} , 2925 cm^{-1} and 2957 cm^{-1} .^[247] In contrast, the negative bands at ca. 3400 cm^{-1} and 1647 cm^{-1} can be assigned to water molecules that are removed from the SAM interface at the same time. Furthermore, this process was also monitored using electrochemical impedance spectroscopy and also proves a successful assembly of the tBLM.^[38]

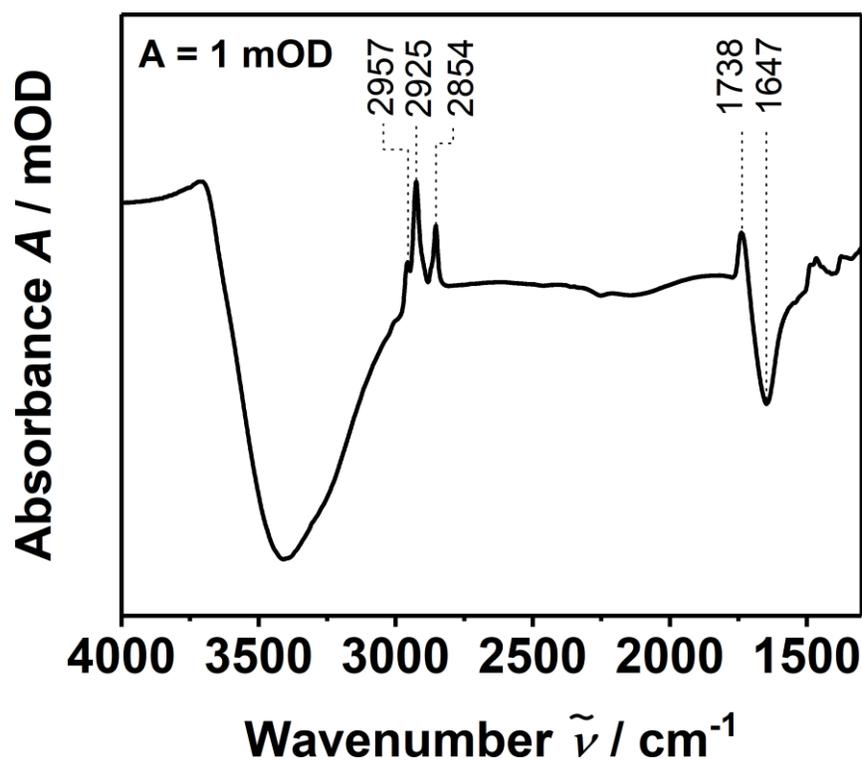


Figure 9.3.3.2.1: SEIRA difference spectrum of the POPC:POPG (90:10) membrane taking the spectrum of the mixed SAM as reference.

The incubation of LL-37, LL-32, and LL-20 on tBLMs formed of POPC:POPG (90:10) was monitored using SEIRA spectroscopy resulting in the spectra plotted in figure 9.3.3.2.2. Clear differences in adsorption of the peptides to negatively charged membranes can be determined from the incubation on the tBLM. In accordance with the previous evidences, LL-37 and LL-32 showed similar spectra with sharp and pronounced amide I and II bands at ca. 1655 and 1550 cm^{-1} (figure 9.3.3.2.2). Further spectral contributions observed in the transmission IR experiment can be detected at the same range of 1681 cm^{-1} , but with a

noticeable relative decrease of intensity which can be explained by the surface selection rule in SEIRA, that allows detecting solely normal modes with a transition dipole moment perpendicular to the surface.

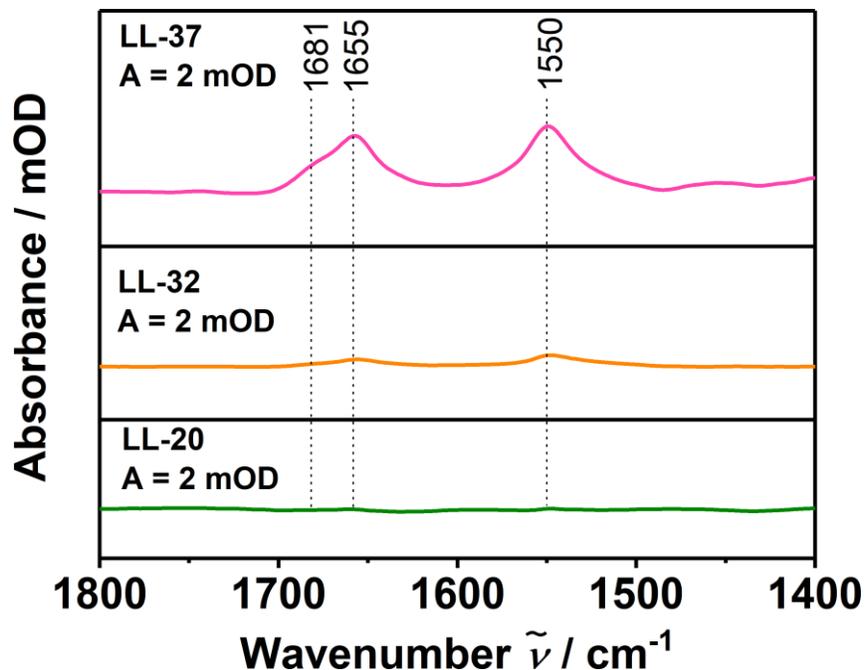


Figure 9.3.3.2.2: Incubation of LL-37, LL-32, and LL-20 on tBLMs formed of POPC:POPG (90:10) monitored using SEIRA spectroscopy.

Likewise, due to this effect, the amide I/amide II ratio is modified in the SEIRA spectra and determined to be 0.6 ± 0.1 for LL-37 and LL-32, which reflects uniform horizontal orientation of both α -helical peptides when bound to the membrane. However, while LL-37 shows a SEIRA spectrum with ca. 2 and 4 mOD (for amide I and amide II) (figure 9.3.3.2.3), the intensities of LL-32 are smaller by a factor of ca. 10 (0.3 and 0.6 mOD for amide I and amide II) (figure 9.3.3.2.3) suggesting an enhanced binding of LL-37 or a deeper incorporation into the membrane core (based on the distance dependence of SEIRA with d^{-6}). Additionally, as it is shown in figure 9.3.3.2.3, the amide I/amide II ratio of LL-37 changes during the membrane binding process from ca. 0.4 to 0.6. In contrast to the value for LL-32, which stays constant at a value of ca. 0.6. This evidence suggests that LL-32 experiences rather a simple one-step process of direct binding to the membrane surface while the LL-37 binding involves a more sophisticated process including more steps, which may explain the 10-times higher intensity.

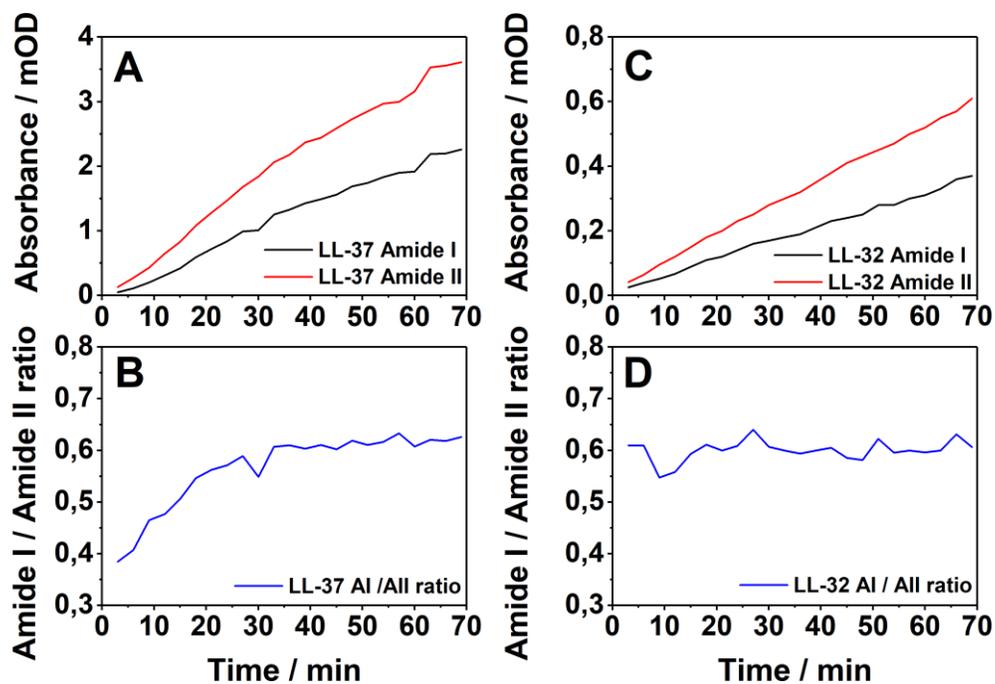
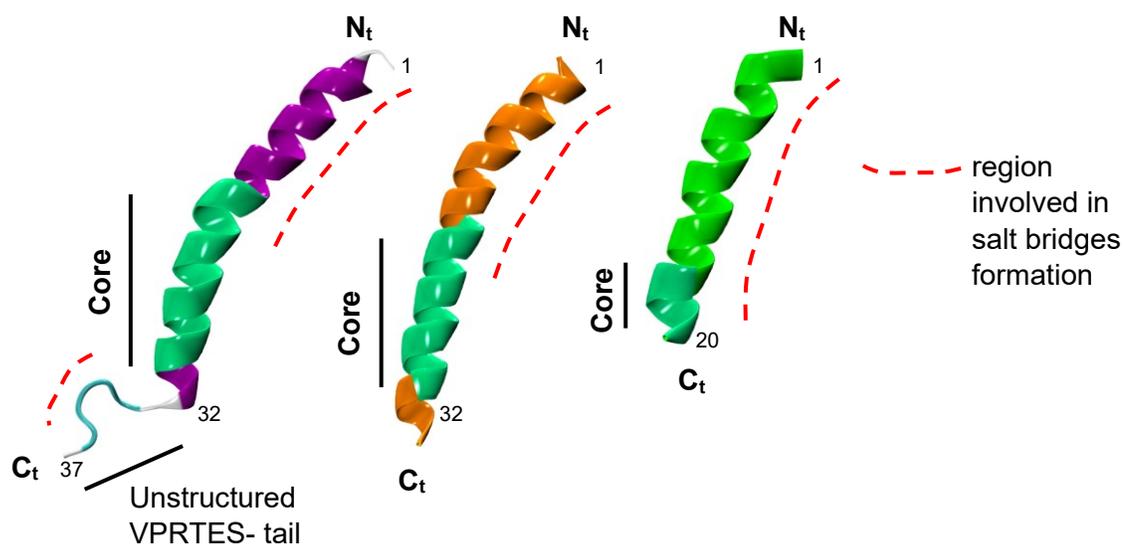


Figure 9.3.3.2.3: Evaluation of the amide I/amide II ratio for LL-37 and LL-32.

9.3.4 Concluding Remarks Combining Theoretical and Experimental Measurements

The aim of this work has been shedding light to the several open questions regarding the mechanism of action of LL-37. According to the combination of the theoretical and experimental results for LL-37 and its fragments (LL-32, LL-20) it is feasible to conclude that:

- i. **An α -helical structure is adopted by the three peptides in physiological buffer solution.**



- a. MD simulations of LL-37 under the same conditions used at the experimental level predicted an α -helical fold of the peptide and an unstructured VPRTES-tail. From the FTIR transmission experiments it is detected that the amide I band of LL-37 and LL-32 appears at the same position, confirming that the VPRTES-tail at the C-terminus (residues 32-37) does not contribute to the overall α -helix structure.^[219]
 - b. LL-20 amide I position of LL-20 is downshifted by ca. 2 cm^{-1}
- ii. **Peptide–membrane activities differ:**
 - a. **LL-20 fragment no membrane binding:** the MD simulations predicted basically no or very weak interaction of the LL-20 with the anionic surface while SEIRA spectra showed barely any peptide signals and thus, no binding to the membrane surface or into the membrane.

b. LL-37 and LL-32 membrane binding:**b.1 Analogies in the membrane binding**

- **Salt bridges formation detected for LL-37 and LL-32 in analogy with those found for LL-20 and variant S9V-LL-20.** Most charged sidechains of the remaining residues (20-37) despite their (i+4) spacing are not involved in salt bridges formation. Only one exception is found in the LL-37 involving the interaction between R34 and E36. These evidences suggest that the peptide core (residues 17-29) (see figure 9.3.1) is determined by the dynamics of salt bridge formation. The lack of salt bridges in the cathelicidin core leads to flexible cationic side chains which are free to interact with an anionic surface and are responsible for a five-fold increase of the total dipole moment of the peptide.
- **MD simulations predicted that LL-37 and LL-32 approach to the membrane surface employing the core (R23-as membrane anchor AA) and adsorb onto the membrane without loss of the secondary structure.** At this stage, different behavior was detected despite of displaying analogous nature of the chemical interactions.

b.2 Differences in the membrane binding

- **For LL-32**, from the trajectory a **one-step process** is predicted where the peptide directly adsorbs onto the anionic surface via primarily electrostatic interactions.
 - **For LL-37**, after the analysis of the MD simulation, a **two-step process** was observed, initiated by the electrostatic attachment of the peptide on the surface followed by the immersion of the LL-37 into the membrane aided by the N_t-hydrophobic head. In this last stage, the residue R34 of the VPRTES-tail, only present in LL-37, plays the essential role of destabilizing the initial electrostatic adsorption and promoting the interaction between the N_t-hydrophobic head and the membrane.
- c. **Experiments were consistent with the MD simulations.** LL-37 attachment was higher extended, i.e. with ca. 10 times higher intensity of the amide

bands. This observation can be explained either by a 10-fold higher occupancy of adsorbed peptides or based on the distance dependent surface enhancement effect in SEIRA, which is directly related on the fact that LL-37 is located closer to the Au surface and thus more deeply immersed within the membrane. Furthermore, during the incubation of LL-32 the amide I/amide II ratio stays constant at a value of ca. 0.6 while the amide I/amide II ratio for LL-37 evolves from ca. 0.4 to 0.6. These distinct amide I/amide II ratios suggested a one-step adsorption process for LL-32 in which the peptide adopts a horizontal orientation with a side-on binding to the membrane while the differences in the ratios detected for LL-37 may be associated to different membrane binding steps: 1) supported by a low amide I/amide II ratio of 0.4, LL-37 attaches parallel on the surface 2) characterized by a ratio increase to 0.6, there is an insertion of peptide followed by a stabilization parallel to the surface. Thus, it is apparent that the VPRTES-tail at the C-terminus is responsible for the much more pronounced binding of LL-37 to the membrane during the initial phase of the adsorption process.

- d. **Dimerization studies carried out by means of MD simulations** by either LL-37 or LL-32 peptides in the presence of the POPC/POPG layer **confirmed that the driving force for the oligomerization is based on the “Hofmeister effect”**^[214] consisting on the hydrophobic effects shielding the hydrophobic face exposed to the solvent and in disagreement to the Shahmiri *et. al.*^[245] suggestions.
- e. **Understanding the role of the LL-37-VPRTES-tail.** The analysis of the MD simulations predicted that the **LL-37-VPRTES-tail** presence **is not only relevant for the immersion of LL-37** into the membrane but **also for peptide oligomerization** (section 9.3.2.2-B and figure 9.3.2.2-B.3). VPRTES-tail in the LL-37 which gathers both monomers via the R34 and E36' side chains and orients them in an energetically favorable manner with their hydrophobic side facing each other in an optimal parallel arrangement. In the case of LL-32, the two monomers form dimers in an antiparallel arrangement.

Appendix

10- Appendix - Bacteriocins

11- Appendix – Cathelicidin LL-37 and fragments

10. Appendix – Bacteriocins

10.1 Force Field Parameters for Dehydroalanine and Dehydrobutyrine

10.1.1 Dha Topology

```

RESI      DHA      0.00      ! 2,3-Didehydroalanine
GROUP
ATOM      N        NH1          -0,47 !
ATOM      HN       H            0,31 !      |
ATOM      CA       CW           0,16 !      HN-N   HB1
GROUP
          !      |      /
ATOM      C        C            0,51 !      CA==CB
ATOM      O        O           -0,51 !      |      \
GROUP
          !      O = C   HB2
ATOM      CB       CEL2        -0,42 !      |
ATOM      HB1      HEL2         0,21 !
ATOM      HB2      HEL2         0,21 !

BOND N HN N CA CA C
BOND C +N CB HB1 CB HB2
DOUBLE O C
DOUBLE CA CB
IMPR N-C CA HN C CA +N O
DONOR HN N
ACCEPTOR O C
AUTO ANGE DIHE

```

10.1.2 Dhb Topology

```

RESI      DBU      0.00      ! DHB (Z)-2,3-didehydrobutyrine
GROUP
ATOM      N        NH1          -0,47 !
ATOM      HN       H            0,31 !
ATOM      CA       CW           0,16 !
GROUP
          !
ATOM      C        C            0,51 !      |

```

```

ATOM      O          O          -0,51 !  HN-N   HB1
GROUP                    !      |   /
ATOM      CB         CEL1       -0,15 !    CA==CB
ATOM      HB         HEL1        0,15 !      |   \
GROUP                    !  O = C   CG--HG1
ATOM      CG         CTL3       -0,27 !      |   / \
ATOM      HG1        HAL3        0,09 !          HG3 HG2
ATOM      HG2        HAL3        0,09 !
ATOM      HG3        HAL3        0,09 !

BOND CG CB CA N HN N
BOND C CA C +N CB HB
BOND CG HG1 CG HG2 CG HG3
DOUBLE O C
DOUBLE CA CB
IMPR N-C CA HN C CA +N O
DONOR HN N
ACCEPTOR O C
AUTO ANGE DIHE

```

10.1.3 Remarks on Atom-Type Assignment

Dehydroamino acids have a double bond between CA and CB, therefore in analogy to R-C=CH₂ alkane groups in lipids, CB in Dha was assigned to a CEL2 atom type while HB1 and HB2 to HEL2. In the case of Dhb, CB and HB were assigned to a CEL1 and HEL1 atom types. The force field parameters associated to these atom types are available in the top_all36_lipid.rtf and the par_all36_lipid.prm files of the CHARMM36 force field. The amide- and carbonyl- groups of the backbones were assigned to the atom types NH1, H and C, O as done for all other canonical amino acids in the CHARMM force field. Finally, the CA atoms were assigned to the CG2D1O atom type from the CHARMM General Force Field v2b4 with a charge of 0.16, as suggested by Paramchem.org server for both molecules. This atom type was renamed to CW.

10.1.4 Force Field Parameters

The FF parameters used during the MD simulations for the dehydroamino acids were taken from the par_all36_protein.prm and par_all36_lipid.prm file of the CHARMM force

field. Additional FF parameters associated to the atom type CW are listed below. Most of these values were generated with the Paramchem.org server or by analogy as indicated below. FF parameters for dihedral contributions modified according to Thormann & Hofmann study are written in italics.

BONDS

CW C 250.000 1.4900 ! from CT1 C Ala Dipeptide
 CW CEL1 440.000 1.340 ! butene, from bond N C
 NH1 CW 320.000 1.4300 ! from NH1 CT1
 CW CEL2 500.000 1.342 ! from CEL1 CEL2

ANGLES

CW NH1 C 50.000 120.0000 ! from CG2D1O-NG2S1-CG201, penalty 21
 H NH1 CW 35.000 117.0000 ! from H-NH1-CT1, NMA Vibrational Modes (LK)
 N C CW 80.000 116.5000 ! from CG2D1O CG201 NG2S0, penalty 4,5
 NH1 CW C 50.000 107.0000 ! from NH1-CT1-C Alanine Dipeptide
 O C CW 80.000 121.0000 ! from O-C-CT1 Alanine Dipeptide
 C CW CEL1 48.00 123.50 ! from CG2DC1-CG2D1O-CG201, penalty 40
 CEL1 CW NH1 60.00 122.00 ! from CG2DC1-CG2D1O-NG311, penalty 66
 CEL2 CW NH1 80.00 128.00 ! from propene CEL2 CEL1 CTL2
 CEL2 CW C 60.00 120.00 ! 1-butene; from CEL2 CEL1 CTL2
 HEL1 CEL1 CW 52.00 119.50 ! 2-butene
 HEL2 CEL2 CW 45.00 120.50 ! propene, from HEL2 CEL2 CEL1

DIHEDRALS

C CW NH1 C 0.4800 2 180.00 ! from X-CUA1-NP-X, Thormann & Hofmann 1998
 CW C N CP1 1.6000 1 0.00 ! from CG2D1O-CG201-NG2S0-CG3C51, penalty 24,4
 CW C N CP1 2.5000 2 180.00 ! from CG2D1O-CG201-NG2S0-CG3C51, penalty 24,4
 CW C N CP3 1.6000 1 0.00 ! from CG2D1O-CG201-NG2S0-CG3C52, penalty 24,4
 CW C N CP3 2.5000 2 180.00 ! from CG2D1O-CG201-NG2S0-CG3C52, penalty 24,4
 CW C NH1 CT1 1.6000 1 0.00 ! from CT1-C-NH1-CT1
 CW C NH1 CT1 2.5000 2 180.00 ! from CT1-C-NH1-CT1
 CW C NH1 CW 1.6000 1 0.00 ! from CT1-C-NH1-CT1
 CW C NH1 CW 2.5000 2 180.00 ! from CT1-C-NH1-CT1
 CT1 C NH1 CW 1.6000 1 0.00 ! from CT1-C-NH1-CT1
 CT1 C NH1 CW 2.5000 2 180.00 ! from CT1-C-NH1-CT1
 CEL1 CW NH1 C 1.8000 1 0.00 ! from CT2-CT1-NH1-C
 CEL1 CW NH1 C 0.4800 2 180.00 ! from CT2-CT1-NH1-C
 H NH1 C CW 2.5000 2 180.00 ! from H-NH1-C-CT1
 H NH1 CW C 0.4800 2 180.00 ! from H-NH1-CT1-C
 HEL2 CEL2 CW C 5.2000 2 180.00 ! from HE2-CE2-CE1-CT2
 NH1 C CW NH1 0.6000 1 0.00 ! from NH1-C-CT1-NH1
 NH1 CW C N 0.5000 2 180.00 ! from NG2S1-CG2D1O-CG201-NG2S0, penalty 111
 NH1 CW C N 0.3500 3 180.00 ! from NG2S1-CG2D1O-CG201-NG2S0, penalty 111
 NH1 CW C N 0.4000 6 0.00 ! from NG2S1-CG2D1O-CG201-NG2S0, penalty 111
 O C CW NH1 0.0000 1 0.00 ! from O-C-CT1-NH1
 O C NH1 CW 2.5000 2 180.00 ! from O-C-NH1-CT1
 O C CW CEL1 0.7000 1 180.00 ! from OG2D1-CG201-CG2DC1-CG2DC1, penalty 40
 O C CW CEL1 1.2000 2 180.00 ! from OG2D1-CG201-CG2DC1-CG2DC1, penalty 40
 O C CW CEL1 0.1000 3 180.00 ! from OG2D1-CG201-CG2DC1-CG2DC1, penalty 40
 O C CW CEL1 0.2000 4 0.00 ! from OG2D1-CG201-CG2DC1-CG2DC1, penalty 40
 C CW CEL1 HEL1 1.0000 2 180.00 ! from CTL3-CEL1-CEL1-HEL1
 NH1 CW CEL1 CTL3 2.5000 2 180.00 ! from NG311-CG2D1O-CG2DC1-CG321, penalty 67
 H NH1 CW CEL1 0.4800 2 180.00 ! from X-CUA1-NP-X, Thormann & Hofmann 1998
 CW CEL1 CTL3 HAL3 0.0300 3 0.00 ! from CEL1-CEL1-CTL3-HAL3
 CEL1 CW C N 0.700 1 0.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
 CEL1 CW C N 1.200 2 180.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
 CEL1 CW C N 0.100 3 0.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
 CEL1 CW C N 0.150 4 0.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41

```

CEL1 CW C NH1 0.700 1 0.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
CEL1 CW C NH1 1.200 2 180.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
CEL1 CW C NH1 0.100 3 0.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
CEL1 CW C NH1 0.150 4 0.00 ! from CG2DC1-CG2DC1-CG201-NG2S1, penalty 41
CEL2 CW C NH1 1.800 2 180.00 ! from CG2R62-CG2R62-CG2R63-NG2R61
C NH1 CW CEL2 1.600 1 0.00 ! from CG2DC1-CG2O1-NG2S1-CG2R61
C NH1 CW CEL2 0.480 2 180.00 ! from CG2DC1-CG2O1-NG2S1-CG2R61, modified
! according to Thormann & Hofmann 1998
NH1 CW CEL2 HEL2 5.000 2 180.00 ! from CG2DC2 CG2DC1 CG2DC3 HGA5
NH1 CW CEL1 HEL1 3.00 2 180.00 ! from OG301-CG2D1O CG2DC1 HGA4
H NH1 CW CEL2 0.480 2 180.00 ! from X-CUA1-NP-X, Thormann & Hofmann 1998
O C CW CEL2 0.700 1 180.00 ! from OG2D1-CG201-CG2DC1-CG2DC1,
O C CW CEL2 1.200 2 180.00 ! from OG2D1-CG201-CG2DC1-CG2DC1
O C CW CEL2 0.100 3 180.00 ! from OG2D1-CG201-CG2DC1-CG2DC1
O C CW CEL2 0.200 4 0.00 ! from OG2D1-CG201-CG2DC1-CG2DC1
C CW CEL1 CTL3 0.56 1 180.00 ! from CG201-CG2DC1-CG2DC1-CG331

```

10.2 Force Field Parameters for Lanthionine and Methyllanthionine

10.2.1 Topology for Methyllanthionine

```

RESI      MLAN      .000      ! path for methyllanthionine
                                     !follow with AUTOgenerate ANGLES DIHEDRALS command

GROUP
ATOM      1CB      CT2      -0,09 !
ATOM      1HB2     HA       0,09 !
!ATOM     1HB3     HA       0,09 !
ATOM      1CG      CT3     -0,27 !
ATOM      1HG1     HA       0,09 !
ATOM      1HG2     HA       0,09 !      1HB2      2HB1
ATOM      1HG3     HA       0,09 !      |          |
GROUP                !      --1CB---2SG---2CB--
ATOM      2CB      CT2     -0,11 !      |          |
ATOM      2HB1     HA       0,09 !      1HG3--1CG--1HG1 2HB2
ATOM      2HB2     HA       0,09 !      |
ATOM      2SG      S       -0,07 !      1HG2
ATOM      2HG1     HS       0,16 !

DELETE ATOM 1HB3
DELETE ATOM 2HG1
BOND 1CB 2SG

```

10.2.2 Topology for Lanthionine

```

RESI      LAN      0.00      ! path for lanthionine
                                     !follow with AUTOgenerate ANGLES DIHEDRALS command

GROUP
ATOM      1CB      CT2      -0,18 !
ATOM      1HB1     HA       0,09 !
!ATOM     1HB2     HA       0,09 !
ATOM      1HB3     HA       0,09 !
GROUP
ATOM      2CB      CT2      -0,11 !           1HB2      2HB1
ATOM      2HB1     HA       0,09 !           |           |
ATOM      2HB2     HA       0,09 !           --1CB---2SG---2CB--
ATOM      2SG      S       -0,07 !           |           |
ATOM      2HG1     HS       0,16 !           1HB1      2HB2
DELETE ATOM 1HB3
DELETE ATOM 2HG1
BOND 1CB 2SG

```

10.2.3 Force Field Parameters

DIHEDRALS

```

HA CT2 S CT2 0.2800 3 0.00
CT2 S CT2 CT1 0.2400 1 180.00 ! from CG321 CG321 SG311 CG321
CT2 S CT2 CT1 0.3700 3 0.00 ! from CG321 CG321 SG311 CG321
S CT2 CT3 HA 0.1600 3 0.00 ! from SG311 CG321 CG331 HGA3
HB CT1 CT2 S 0.1950 3 0.00 ! from HGA1 CG311 CG321 SG311
C CT1 CT2 S 0.2000 3 0.00 ! from CG2O1 CG311 CG321 SG311
NH1 CT1 CT2 S 0.2000 3 0.00 ! from NG2S1 CG311 CG321 SG311
CT1 CT2 S CT2 0.2400 1 180.00 ! from CG321 CG321 SG311 CG321
CT1 CT2 S CT2 0.3700 3 0.00 ! from CG321 CG321 SG311 CG321
CT3 CT2 S CT2 0.2400 1 180.00 ! from CG331 CG321 SG311 CG331
CT3 CT2 S CT2 0.3700 3 0.00 ! from CG331 CG321 SG311 CG331

```


ATOM H195	HGP1	0.3121	0.000	BOND C1 N3	BOND C105 H107	BOND C208 O210
ATOM C196	CG311	0.0281	0.439	BOND C1 C11	BOND C105 C269	BOND C211 H212
ATOM H197	HGA1	0.0901	0.000	BOND C1 H2	BOND H108 C268	BOND C211 H213
ATOM C198	CG331	-0.2701	0.000	BOND C1 C26	BOND C109 H110	BOND C211 C214
ATOM H199	HGA3	0.0901	0.000	BOND N3 H4	BOND C109 H112	BOND C214 H216
ATOM H200	HGA3	0.0901	0.000	BOND N3 C5	BOND C109 H111	BOND C214 H215
ATOM H201	HGA3	0.0901	0.000	BOND C5 C7	BOND C109 C269	BOND C214 C217
ATOM C202	CG201	0.5161	0.307	BOND C5 O6	BOND C113 H114	BOND C217 O218
ATOM O203	OG201	-0.5061	0.000	BOND C7 H10	BOND C113 C116	BOND C217 N219
ATOM N204	NG251	-0.4771	0.320	BOND C7 H8	BOND C113 H115	BOND N219 H220
ATOM H205	HGP1	0.4841	0.000	BOND C7 H9	BOND C113 C269	BOND N219 C221
ATOM C206	CG311	-0.1831	0.287	BOND C11 O12	BOND C116 H117	BOND C221 H222
ATOM H207	HGA1	0.0901	0.000	BOND C11 H14	BOND C116 H118	BOND C221 C223
ATOM C208	CG203	0.6221	0.000	BOND C11 C15	BOND C116 C270	BOND C221 C239
ATOM O209	OG202	-0.7541	0.000	BOND O12 H13	BOND C119 H120	BOND C223 H224
ATOM O210	OG202	-0.7541	0.000	BOND C15 H16	BOND C120 H123	BOND C223 H225
ATOM C211	CG321	-0.1651	0.000	BOND C15 O17	BOND C120 H21	BOND C223 C226
ATOM H212	HGA2	0.0901	0.000	BOND C15 C19	BOND C120 H22	BOND C226 H227
ATOM H213	HGA2	0.0901	0.000	BOND O17 H18	BOND C120 C271	BOND C226 H228
ATOM C214	CG321	-0.2331	0.000	BOND C19 H20	BOND C124 H25	BOND C226 C229
ATOM H215	HGA2	0.0901	0.000	BOND C19 C21	BOND C124 H26	BOND C229 H230
ATOM H216	HGA2	0.0901	0.000	BOND C19 O28	BOND C124 C127	BOND C229 H231
ATOM C217	CG201	0.5251	0.000	BOND C21 H23	BOND C124 C271	BOND C229 C232
ATOM O218	OG201	-0.5071	0.000	BOND C21 H22	BOND C127 H128	BOND C232 H234
ATOM N219	NG251	-0.4151	0.000	BOND C21 O24	BOND C127 H29	BOND C232 H233
ATOM H220	HGP1	0.3051	0.000	BOND O24 H25	BOND C127 C272	BOND C232 N235
ATOM C221	CG311	0.0331	0.333	BOND C26 H27	BOND H130 C272	BOND N235 H238
ATOM H222	HGA1	0.0931	0.000	BOND C26 O29	BOND C131 H132	BOND N235 H236
ATOM C223	CG321	-0.1801	0.000	BOND C26 O28	BOND C131 H133	BOND N235 H237
ATOM H224	HGA2	0.0901	0.000	BOND O29 C30	BOND C131 H134	BOND C239 O240
ATOM H225	HGA2	0.0901	0.000	BOND C30 C32	BOND C131 C273	BOND C239 N241
ATOM C226	CG321	-0.1831	0.000	BOND C30 C53	BOND C135 H136	BOND N241 C243
ATOM H227	HGA2	0.0901	0.000	BOND C30 H31	BOND C135 C138	BOND N241 H242
ATOM H228	HGA2	0.0901	0.000	BOND C32 H33	BOND C135 H137	BOND C243 H244
ATOM C229	CG321	-0.1841	0.000	BOND C32 C34	BOND C135 C273	BOND C243 C245
ATOM H230	HGA2	0.0901	0.000	BOND C38 C39	BOND C138 H140	BOND C243 C249
ATOM H231	HGA2	0.0901	0.000	BOND C34 H35	BOND C138 H139	BOND C245 H247
ATOM C232	CG324	0.1261	0.000	BOND C34 H36	BOND C138 C274	BOND C245 H248
ATOM H233	HGA2	0.0901	0.000	BOND C34 O37	BOND H141 C274	BOND C245 H246
ATOM H234	HGA2	0.0901	0.000	BOND C37 H38	BOND C142 H144	BOND C249 O250
ATOM N235	NG3P3	-0.9041	0.400	BOND C39 C40	BOND C142 H145	BOND C249 N251
ATOM H236	HGP2	0.3311	0.000	BOND C40 H41	BOND C142 H143	BOND N251 H252
ATOM H237	HGP2	0.3311	0.000	BOND C40 O42	BOND C142 C275	BOND N251 C253
ATOM H238	HGP2	0.3311	0.000	BOND C40 C43	BOND C146 H147	BOND C253 C255
ATOM C239	CG201	0.5331	0.307	BOND C42 H43	BOND C146 H148	BOND C253 C259
ATOM O240	OG201	-0.5091	0.000	BOND C43 H44	BOND C146 C149	BOND C253 H254
ATOM N241	NG251	-0.4731	0.320	BOND C43 C53	BOND C146 C275	BOND C255 H256
ATOM H242	HGP1	0.3121	0.000	BOND C43 N45	BOND C149 H150	BOND C255 H257
ATOM C243	CG311	0.0281	0.439	BOND N45 H46	BOND C149 H151	BOND C255 H258
ATOM H244	HGA1	0.0901	0.000	BOND N45 C47	BOND C149 C276	BOND C259 O261
ATOM C245	CG331	-0.2701	0.000	BOND C47 O48	BOND H152 C276	BOND C259 O260
ATOM H246	HGA3	0.0901	0.000	BOND C47 C49	BOND C153 H154	BOND C262 C263
ATOM H247	HGA3	0.0901	0.000	BOND C49 H50	BOND C153 H155	BOND C264 C265
ATOM H248	HGA3	0.0901	0.000	BOND C49 H51	BOND C153 H156	BOND C266 C267
ATOM C249	CG201	0.5161	0.307	BOND C49 H52	BOND C153 C277	BOND C268 C269
ATOM O250	OG201	-0.5061	0.000	BOND C53 H54	BOND C157 H160	BOND C270 C271
ATOM N251	NG251	-0.5041	0.320	BOND C53 O55	BOND C157 H158	BOND C272 C273
ATOM H252	HGP1	0.4851	0.000	BOND C55 C56	BOND C157 H159	BOND C274 C275
ATOM C253	CG311	-0.1631	0.287	BOND C56 H57	BOND C157 C277	BOND C276 C277
ATOM H254	HGA1	0.0901	0.000	BOND C56 C58	BOND C160 H161	BOND C278 C279
ATOM C255	CG331	-0.2701	0.000	BOND C56 C82	BOND C160 H162	BOND C280 C281
ATOM H256	HGA3	0.0901	0.000	BOND C58 H59	BOND C160 C278	BOND C282 C283
ATOM H257	HGA3	0.0901	0.000	BOND C58 H60	BOND H163 C278	IMPR C5 C7 N3 O6
ATOM H258	HGA3	0.0901	0.000	BOND C58 H61	BOND C164 H165	IMPR C47 C49 N45 O48
ATOM C259	CG203	0.6191	0.000	BOND C62 O63	BOND C164 H166	IMPR C62 C56 N194 O63
ATOM O260	OG202	-0.7571	0.000	BOND C62 H164	BOND C164 H167	IMPR C202 C196 N204 O203
ATOM O261	OG202	-0.7571	0.000	BOND O64 P65	BOND C164 C279	IMPR C208 O210 O209 C206
ATOM C262	CG201	-0.1561	7.103	BOND O64 P68	BOND C168 H170	IMPR C217 C214 N219 O218
ATOM C263	CG201	-0.0031	4.508	BOND P65 O66	BOND C168 C171	IMPR C239 C221 N241 O240
ATOM C264	CG201	-0.1451	0.000	BOND P65 O67	BOND C168 H169	IMPR C249 C243 N251 O250
ATOM C265	CG201	-0.0021	0.000	BOND P68 O70	BOND C171 H172	IMPR C259 O260 O261 C253
ATOM C266	CG201	-0.1451	0.000	BOND P68 O69	BOND C171 H173	END
ATOM C267	CG201	-0.0021	0.000	BOND P68 O71	BOND C171 H173	
ATOM C268	CG201	-0.1451	0.000	BOND O71 C72	BOND C171 C281	
ATOM C269	CG201	-0.0021	0.000	BOND C72 H73	BOND H174 C281	
ATOM C270	CG201	-0.1451	0.000	BOND C72 H74	BOND C175 H176	
ATOM C271	CG201	-0.0021	0.000	BOND C72 C282	BOND C175 H177	
ATOM C272	CG201	-0.1451	0.000	BOND H75 C282	BOND C175 H178	
ATOM C273	CG201	-0.0021	0.000	BOND C76 H77	BOND C175 C280	
ATOM C274	CG201	-0.1451	0.000	BOND C76 H78	BOND C179 H181	
ATOM C275	CG201	-0.0021	0.000	BOND C76 H79	BOND C179 C182	
ATOM C276	CG201	-0.1451	0.000	BOND C76 C263	BOND C179 H180	
ATOM C277	CG201	-0.0021	0.000	BOND C80 H81	BOND C179 C280	
ATOM C278	CG201	-0.1451	0.000	BOND C80 H82	BOND C182 H183	
ATOM C279	CG201	-0.0021	0.000	BOND C80 C83	BOND C182 H184	
ATOM C280	CG201	-0.0021	0.000	BOND C80 C263	BOND C182 C282	
ATOM C281	CG201	-0.1451	0.000	BOND C83 H84	BOND H185 C282	
ATOM C282	CG201	-0.1421	0.000	BOND C83 H85	BOND C186 H188	
ATOM C283	CG201	0.0031	0.110	BOND C83 C264	BOND C186 H189	
				BOND H86 C264	BOND C186 H190	
				BOND C87 H88	BOND C188 C283	
				BOND C87 H89	BOND C187 H192	
				BOND C87 H90	BOND C187 H191	
				BOND C87 C265	BOND C187 H193	
				BOND C91 C94	BOND C187 C283	
				BOND C91 H92	BOND N194 H195	
				BOND C91 H93	BOND N194 C196	
				BOND C91 C265	BOND C196 H197	
				BOND C94 H96	BOND C196 C198	
				BOND C94 H95	BOND C198 C202	
				BOND C94 C266	BOND C198 H199	
				BOND H97 C266	BOND C198 H200	
				BOND C98 H99	BOND C198 H201	
				BOND C98 H100	BOND C202 C203	
				BOND C98 H101	BOND C202 N204	
				BOND C98 C267	BOND N204 H205	
				BOND C102 H103	BOND N204 C206	
				BOND C102 H104	BOND C206 H207	
				BOND C102 C105	BOND C206 C208	
				BOND C102 C267	BOND C206 C211	
				BOND C105 H106	BOND C208 O209	

read param card flex append

* Parameters generated by analogy by

* CHARMM General Force Field (CGenFF) program version 0.9.7.1 beta

BONDS

CG2D1	CG2D1	440.00	1.3400 ! LIPID butene, yin_adm jr., 12/95
CG2D1	CG321	385.00	1.5020 ! LIPID butene, from propene, yin_adm jr., 12/95
CG2D1	CG331	383.00	1.5040 ! LIPID butene, yin_adm jr., 12/95
CG2D1	HGA4	360.50	1.1000 ! LIPID propene, yin_adm jr., 12/95
CG201	CG311	250.00	1.4900 ! PROT Ala Dipeptide (5/91)
CG201	CG321	250.00	1.4900 ! PROT Ala Dipeptide (5/91)
CG201	CG331	250.00	1.4900 ! PROT Ala Dipeptide (5/91)
CG201	NG251	370.00	1.3450 ! PROT Alanine Dipeptide ab initio calc's (LK)
CG201	OG2D1	620.00	1.2300 ! PROT Peptide geometry, condensed phase (LK)
CG203	CG311	200.00	1.5220 ! PROT adm jr, 4/05/91, for asn,asp,glu and cters
CG203	OG2D2	525.00	1.2600 ! PROT adm jr, 7/23/91, acetic acid
CG311	CG311	222.50	1.5000 ! PROT alkane update, adm jr., 3/2/92
CG311	CG321	222.50	1.5380 ! PROT alkane update, adm jr., 3/2/92
CG311	CG331	222.50	1.5380 ! PROT alkane update, adm jr., 3/2/92
CG311	NG251	320.00	1.4300 ! PROT NMA Gas & Liquid Phase IR Spectra (LK)
CG311	OG301	360.00	1.4150 ! all34_ethers_1a CG32A OG30A, gk or og
CG311	OG303	340.00	1.4300 ! LIPID phosphate
CG311	OG311	428.00	1.4200 ! PROT methanol vib fit EMB 11/21/89

Appendix

CG311	OG3061	360.00	1.4150	!	Molecu	, from CG321 OG3061, PENALTY= 4
CG311	HGA1	308.00	1.1110	!	PROT alkane update, adm jr	, 3/2/92
CG321	CG321	222.50	1.5300	!	PROT alkane update, adm jr	, 3/2/92
CG321	CG324	222.50	1.5300	!	FLAVOP PIP1,2,3	
CG321	OG303	320.00	1.4400	!	PROTNA serine/threonine phosphate	
CG321	OG311	428.00	1.4200	!	PROT methanol vib fit EMB 11/21/89	
CG321	HGA2	309.00	1.1110	!	PROT alkane update, adm jr	, 3/2/92
CG324	NG3P3	200.00	1.4900	!	MAMA, methylammonium 1.48-> 1.49 based on CCSDT calc (kevo) and xtal survey (pram)	
CG324	HGA2	284.50	1.1000	!	FLAVOP PIP1,2,3	
CG331	HGA3	322.00	1.1110	!	PROT alkane update, adm jr	, 3/2/92
NG2S1	HGP1	440.00	0.9970	!	PROT Alanine Dipeptide ab initio calc's (LK)	
NG3P3	HGP2	403.00	1.0400	!	PROT new stretch and bend; methylammonium (KK 03/10/92)	
OG2P1	PG1	500.00	1.5100	!	MP_1 reorganization, kevo	
OG303	PG1	190.00	1.6500	!	MP_1 reorganization, kevo	
OG304	PG1	330.00	1.6750	!	PP1, PPI2, METP reorganization, kevo ! pulls against attraction	
OG311	HGP1	545.00	0.9600	!	PROT EMB 11/21/89 methanol vib fit; og tested on MeOH EtOH...	
ANGLES						
CG2D1	CG2D1	CG321	48.00	123.50	!	LIPID 2-butene, yin,adm jr
CG2D1	CG2D1	CG331	48.00	123.50	!	LIPID 2-butene, yin,adm jr
CG2D1	CG2D1	HGA4	52.00	119.50	!	LIPID 2-butene, yin,adm jr
CG321	CG2D1	CG331	48.00	123.50	!	RETINOL TMCH
CG321	CG2D1	HGA4	40.00	116.00	!	LIPID 1-butene; propene, yin,adm jr
CG331	CG2D1	CG331	48.00	123.50	!	Molecu
CG311	CG2D1	NG2S1	80.00	116.50	!	PROT Alkane Vib Modes (LK)
CG311	CG2D1	OG2D1	80.00	121.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG321	CG2D1	NG2S1	80.00	116.50	!	PROT Alkane Vib Modes (LK)
CG321	CG2D1	OG2D1	80.00	121.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG331	CG2D1	NG2S1	80.00	116.50	!	PROT Alkane Vib Modes (LK)
CG331	CG2D1	OG2D1	80.00	121.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
NG2S1	CG2D1	OG2D1	80.00	122.50	!	PROT Alkane Vib Modes (LK)
CG311	CG2D1	OG2D1	40.00	116.00	!	50.00 2.35300 ! PROT adm jr. 7/23/91, correction, ACETATE (KK)
OG2D2	CG2D1	OG2D2	100.00	128.00	!	70.00 2.25870 ! PROT adm jr. 7/23/91, correction, ACETATE (KK)
CG2D1	CG311	CG321	52.00	108.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG2D1	CG311	CG331	52.00	108.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG2D1	CG311	NG2S1	50.00	107.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG2D1	CG311	OG301	45.00	110.50	!	Molecu
CG2D1	CG311	HGA1	50.00	109.50	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG2D1	CG311	CG321	52.00	108.00	!	PROT adm jr. 4/05/91, for asn,asp,glu and cters
CG2D1	CG311	CG331	52.00	108.00	!	PROT adm jr. 4/05/91, for ALA cter
CG2D1	CG311	NG2S1	50.00	107.00	!	PROT adm jr. 4/05/91, for asn,asp,glu and cters
CG2D1	CG311	HGA1	50.00	109.50	!	PROT adm jr. 4/05/91, for asn,asp,glu and cters
CG311	CG311	CG311	53.35	111.00	!	8.00 2.56100 ! PROT alkane update, adm jr
CG311	CG311	CG321	53.35	111.00	!	8.00 2.56100 ! PROT alkane update, adm jr
CG311	CG311	NG2S1	70.00	113.50	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG311	CG311	OG301	115.00	109.70	!	Molecu
CG311	CG311	OG303	115.00	109.70	!	Molecu
CG311	CG311	OG311	75.00	110.10	!	PROT MeOH, EMB, 10/10/89
CG311	CG311	OG3061	45.00	111.50	!	Molecu
CG311	CG311	HGA1	34.50	110.10	!	22.53 2.17900 ! PROT alkane update, adm jr
CG321	CG311	NG2S1	70.00	113.50	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG321	CG311	OG3061	45.00	111.50	!	Molecu
CG321	CG311	HGA1	34.50	110.10	!	22.53 2.17900 ! PROT alkane update, adm jr
CG331	CG311	NG2S1	70.00	113.50	!	PROT Alanine Dipeptide ab initio calc's (LK)
CG331	CG311	OG301	45.00	111.50	!	Molecu
CG331	CG311	HGA1	34.50	110.10	!	22.53 2.17900 ! PROT alkane update, adm jr
NG2S1	CG311	HGA1	48.00	108.00	!	PROT Alanine Dipeptide ab initio calc's (LK)
OG301	CG311	OG3061	45.00	110.50	!	Molecu
OG301	CG311	HGA1	60.00	109.50	!	alil34_ethers_1a HCA2 CC32A OC30A, gk or og (not affected by mistake)
OG303	CG311	OG3061	45.00	110.50	!	Molecu
OG303	CG311	HGA1	60.00	109.50	!	PROTNA Ser-Phos
OG311	CG311	HGA1	45.00	108.89	!	PROT MeOH, EMB, 10/10/89
OG3061	CG311	HGA1	45.00	109.50	!	Molecu
CG2D1	CG321	CG321	32.00	112.20	!	LIPID 1-butene; propene, yin,adm jr
CG2D1	CG321	OG303	20.00	99.00	!	Molecu
CG2D1	CG321	HGA2	45.00	111.50	!	LIPID 1-butene; propene, yin,adm jr
CG2D1	CG321	CG321	52.00	108.00	!	PROT adm jr. 5/02/91, acetic acid pure solvent
CG2D1	CG321	HGA2	33.00	109.50	!	30.00 2.16300 ! PROT alanine dipeptide, 5/09/91
CG311	CG321	CG321	58.35	113.50	!	11.16 2.56100 ! PROT alkanes
CG311	CG321	OG311	75.70	110.10	!	PROT MeOH, EMB, 10/10/89
CG311	CG321	HGA2	33.43	110.10	!	22.53 2.17900 ! PROT alkanes
CG321	CG321	CG321	58.35	113.60	!	11.16 2.56100 ! PROT alkane update, adm jr
CG321	CG321	HGA2	26.50	110.10	!	22.53 2.17900 ! PROT alkane update, adm jr
CG324	CG321	HGA2	26.50	110.10	!	22.53 2.17900 ! FLAVOP PIP1,2,3
OG303	CG321	HGA2	60.00	109.50	!	PROTNA Thr-Phos
OG311	CG321	HGA2	45.00	108.89	!	PROT MeOH, EMB, 10/10/89
HGA2	CG321	HGA2	35.50	109.00	!	5.40 1.80200 ! PROT alkane update, adm jr
CG321	CG324	NG3P3	67.70	110.00	!	LIPID ethanalamine
CG321	CG324	HGA2	26.50	111.80	!	22.53 2.17900 ! FLAVOP PIP1,2,3
NG3P3	CG324	HGA2	45.00	107.50	!	35.00 2.10100 ! NA methylammonium
HGA2	CG324	HGA2	35.50	109.00	!	5.40 1.80200 ! PIP1,2,3
CG2D1	CG331	HGA3	42.00	111.50	!	LIPID 2-butene, yin,adm jr
CG2D1	CG331	HGA3	33.00	109.50	!	30.00 2.16300 ! PROT alanine dipeptide, 5/09/91
CG311	CG331	HGA3	33.43	110.10	!	22.53 2.17900 ! PROT alkanes
HGA3	CG331	HGA3	35.50	108.40	!	5.40 1.80200 ! PROT alkane update, adm jr
CG2D1	NG2S1	CG311	50.00	120.00	!	PROT Alkane Vib Modes (LK)
CG2D1	NG2S1	HGP1	34.00	123.00	!	PROT Alkane Vib Modes (LK)
CG324	NG3P3	HGP2	30.00	109.50	!	20.00 2.07400 ! PROT new stretch and bend; methylammonium (KK 03/10/92)
HGP2	NG3P3	HGP2	44.00	109.50	!	PROT new stretch and bend; methylammonium (KK 03/10/92)
CG311	OG301	CG311	95.00	109.70	!	Molecu
CG311	OG303	PG1	20.00	120.00	!	35.00 2.33000 ! Molecu
CG321	OG303	PG1	20.00	120.00	!	35.00 2.33000 ! NA Reorganization: PC and others
PG1	OG304	PG1	45.00	143.00	!	40.00 3.25000 ! PPI2, METP reorganization, kevo
CG311	OG311	HGP1	50.00	106.00	!	og 1/06 EtOH IR fit; was 57.5 106
CG321	OG311	HGP1	50.00	106.00	!	sng mod (qm and crystal data); was 57.5 106
CG311	OG3061	CG311	95.00	109.70	!	Molecu
OG2P1	PG1	OG2P1	104.00	120.00	!	MP_1 reorganization, kevo
OG2P1	PG1	OG303	98.50	107.50	!	MP_1 reorganization, kevo
OG2P1	PG1	OG304	89.50	111.60	!	NA nadiipi, jpr 1/adm jr, 7/85 !Reorganization:PP1, PPI2
OG303	PG1	OG304	48.10	105.00	!	PP1, PPI2, METP reorganization, kevo
DIHEDRALS						
CG321	CG2D1	CG2D1	CG321	0.4500	1	180.00 ! LIPID 2-butene, adm jr
CG321	CG2D1	CG2D1	CG321	8.5000	2	180.00 ! LIPID
CG321	CG2D1	CG2D1	CG331	0.4500	1	180.00 ! LIPID 2-butene, adm jr
CG321	CG2D1	CG2D1	CG331	8.5000	2	180.00 ! LIPID
CG321	CG2D1	CG2D1	HGA4	1.0000	2	180.00 ! LIPID 2-butene, adm jr
CG331	CG2D1	CG2D1	HGA4	1.0000	2	180.00 ! LIPID 2-butene, adm jr
CG2D1	CG2D1	CG321	CG321	0.6000	1	180.00 ! LIPID alkenes
CG2D1	CG2D1	CG321	OG303	1.9000	1	180.00 ! Molecu
CG2D1	CG2D1	CG303	OG303	0.2000	2	180.00 ! Molecu
CG2D1	CG2D1	CG321	OG303	0.6000	3	180.00 ! Molecu
CG2D1	CG2D1	CG321	HGA2	0.3000	3	180.00 ! LIPID alkenes
CG331	CG2D1	CG321	CG321	0.1900	3	0.00 ! RETINOL TMCH
CG331	CG2D1	CG321	HGA2	0.1900	3	0.00 ! RETINOL TMCH
HGA4	CG2D1	CG321	CG321	0.1200	3	0.00 ! LIPID butene, yin,adm jr
HGA4	CG2D1	CG321	OG303	0.2000	3	0.00 ! Molecu
HGA4	CG2D1	CG321	HGA2	0.0000	3	0.00 ! LIPID butene, adm jr
CG2D1	CG2D1	CG331	HGA3	0.3000	3	180.00 ! LIPID alkenes
CG321	CG2D1	CG331	HGA3	0.1600	3	0.00 ! RETINOL TMCH
CG331	CG2D1	CG331	HGA3	0.1600	3	0.00 ! Molecu
NG2S1	CG2D1	CG311	CG321	0.0000	1	0.00 ! PROT ala dipeptide, new C VDW Rmin, 4/10/93 (LK)
NG2S1	CG2D1	CG311	CG331	0.0000	1	0.00 ! PROT ala dipeptide, new C VDW Rmin, 4/10/93 (LK)
NG2S1	CG2D1	CG311	NG2S1	0.6000	1	0.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr
NG2S1	CG2D1	CG311	OG301	0.6000	1	0.00 ! Molecu
NG2S1	CG2D1	CG311	HGA1	0.0000	1	0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
OG2D1	CG2D1	CG311	CG321	1.4000	1	0.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr
OG2D1	CG2D1	CG311	CG331	1.4000	1	0.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr
OG2D1	CG2D1	CG311	NG2S1	0.0000	1	0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
OG2D1	CG2D1	CG311	OG301	0.5500	2	180.00 ! Molecu

OGD21 CG201 CG311 HGA1 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 NG251 CG201 CG321 HGA2 0.0000 3 0.00 ! PROT from NG251 CG201 CG311 CT2, for lactams, adm jr.
 NG251 CG201 CG321 HGA2 0.0000 3 0.00 ! PROT. sp2-methyl, no torsion potential
 OGD21 CG201 CG321 CG321 0.0500 6 180.00 ! PROT For side chains of asp.asn.glu.gln, (n=6) from KK(LK)
 OGD21 CG201 CG321 HGA2 0.0000 3 180.00 ! PROT adm jr., 8/13/90 geometry and vibrations
 NG251 CG201 CG331 HGA3 0.0000 3 0.00 ! PROT. sp2-methyl, no torsion potential
 OGD21 CG201 CG331 HGA3 0.0000 3 180.00 ! PROT adm jr., 8/13/90 geometry and vibrations
 CG311 CG201 NG251 CG3C61 0.6000 3 0.00 ! Molecu, from CG321 CG321 CG321 CG3C61, PENALTY= 8.6
 CG311 CG201 NG251 CG311 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 CG311 CG201 NG251 HGP1 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 CG321 CG201 NG251 CG311 1.6000 1 0.00 ! PROT NMA cis/trans energy difference. (LK)
 CG321 CG201 NG251 CG311 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 CG321 CG201 NG251 HGP1 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 CG331 CG201 NG251 CG311 1.6000 1 0.00 ! PROT NMA cis/trans energy difference. (LK)
 CG331 CG201 NG251 CG311 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 CG331 CG201 NG251 HGP1 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 OGD21 CG201 NG251 HGP1 2.5000 2 180.00 ! PROT Gives appropriate NMA cis/trans barrier. (LK)
 OGD22 CG203 CG311 CG321 0.0500 6 180.00 ! PROT For side chains of asp.asn.glu.gln, (n=6) from KK(LK)
 OGD22 CG203 CG311 CG331 0.0500 6 180.00 ! deleteeme DELETEEME (we want to use wildcarting)
 OGD22 CG203 CG311 NG251 0.0000 6 180.00 ! GA, Glut Acid CDCA Amide, cacha
 OGD22 CG203 CG311 HGA1 0.0500 6 180.00 ! PROT For side chains of asp.asn.glu.gln, (n=6) from KK(LK)
 CG311 CG311 CG311 CG311 0.5000 4 180.00 ! Molecu, from CG311 CG311 CG311 CG321, PENALTY= 0.6
 CG311 CG311 CG311 CG321 0.5000 4 180.00 ! NA bkb
 CG311 CG311 CG311 NG251 0.2000 3 0.00 ! Molecu, from CG321 CG311 CG311 NG251, PENALTY= 0.6
 CG311 CG311 CG311 CG301 0.2000 3 180.00 ! Molecu, from CG321 CG311 CG321 CG302, PENALTY= 5.1
 CG311 CG311 CG311 CG303 0.2000 3 180.00 ! Molecu, from CG321 CG311 CG321 CG303, PENALTY= 4.6
 CG311 CG311 CG311 CG311 0.1400 3 0.00 ! PROT, hydroxyl wild card
 CG311 CG311 CG311 OGC3C61 0.1900 1 180.00 ! Molecu, from CG321 CG321 CG321 OGC3C61, PENALTY= 8.6
 CG311 CG311 CG311 OGC3C61 1.0000 2 180.00 ! Molecu, from CG321 CG321 CG321 OGC3C61, PENALTY= 8.6
 CG311 CG311 CG311 OGC3C61 0.6000 3 0.00 ! Molecu, from CG321 CG321 CG321 OGC3C61, PENALTY= 8.6
 CG311 CG311 CG311 OGC3C61 0.0800 4 180.00 ! Molecu, from CG321 CG321 CG321 OGC3C61, PENALTY= 8.6
 CG311 CG311 CG311 HGA1 0.1950 3 0.00 ! NA, sugar
 CG321 CG311 CG311 OGC301 0.2000 3 180.00 ! Molecu, from CG321 CG311 CG321 OGC301, PENALTY= 4.5
 CG321 CG311 CG311 CG311 0.1400 3 0.00 ! Molecu, from CG311 CG311 CG311 CG311, PENALTY= 0.6
 CG321 CG311 CG311 HGA1 0.1950 3 0.00 ! NA, sugar
 NG251 CG311 CG311 OGC301 0.2000 3 0.00 ! Molecu, from NG251 CG311 CG311 CG311, PENALTY= 15
 NG251 CG311 CG311 OGC303 0.2000 3 0.00 ! Molecu, from NG251 CG311 CG311 CG311, PENALTY= 15
 NG251 CG311 CG311 OGC311 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 NG251 CG311 CG311 OGC3C61 0.2000 3 0.00 ! Molecu, from NG251 CG311 CG311 CG311, PENALTY= 15
 NG251 CG311 CG311 HGA1 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 OGC301 CG311 CG311 OGC301 0.2000 3 0.00 ! Molecu, from OGC301 CG311 CG321 OGC302, PENALTY= 11.5
 OGC301 CG311 CG311 OGC3C61 0.2000 3 0.00 ! Molecu, from OGC301 CG311 CG321 OGC302, PENALTY= 11.5
 OGC301 CG311 CG311 HGA1 0.1950 3 0.00 ! Molecu, from HGA1 CG311 CG321 OGC302, PENALTY= 4.5
 OGC303 CG311 CG311 HGA1 0.1950 3 0.00 ! Molecu, from HGA1 CG311 CG321 OGC303, PENALTY= 4
 OGC311 CG311 CG311 OGC311 0.2000 3 0.00 ! Molecu, from OGC311 CG311 CG321 OGC311, PENALTY= 4
 OGC311 CG311 CG311 OGC3C61 0.2000 3 0.00 ! Molecu, from OGC311 CG311 CG321 OGC303, PENALTY= 12
 OGC311 CG311 CG311 HGA1 0.1950 3 0.00 ! NA, sugar
 OGC3C61 CG311 CG311 HGA1 0.1950 3 0.00 ! Molecu, from OGC3C61 CG321 CG321 HGA2, PENALTY= 8
 HGA1 CG311 CG311 HGA1 0.1950 3 0.00 ! NA, sugar
 CG201 CG311 CG321 CG321 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG201 CG311 CG321 HGA2 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG203 CG311 CG321 CG321 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG203 CG311 CG321 HGA2 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG311 CG311 CG321 CG311 0.2000 3 180.00 ! Molecu, from CG321 CG311 CG321 CG311, PENALTY= 0.6
 CG311 CG311 CG321 HGA2 0.1950 3 0.00 ! NA, sugar
 NG251 CG311 CG321 CG321 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 NG251 CG311 CG321 HGA2 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 OGC3C61 CG311 CG321 CG311 0.1950 3 0.00 ! Molecu, from OGC3C61 CG321 CG321 OGC3C61, PENALTY= 11.5
 OGC3C61 CG311 CG321 HGA2 0.1950 3 0.00 ! Molecu, from OGC3C61 CG321 CG321 HGA2, PENALTY= 4
 HGA1 CG311 CG321 CG321 0.1950 3 0.00 ! NA abasic nucleoside
 HGA1 CG311 CG321 OGC311 0.1950 3 0.00 ! NA, sugar
 HGA1 CG311 CG321 HGA2 0.1950 3 0.00 ! NA, sugar
 CG201 CG311 CG331 HGA3 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG203 CG311 CG331 HGA3 0.1800 3 0.00 ! PROT rotation barrier in Ethane (SF)
 NG251 CG311 CG331 HGA3 0.2000 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 OGC301 CG311 CG331 HGA3 0.1800 3 0.00 ! all34_ethers_1a
 HGA1 CG311 CG331 HGA3 0.1950 3 0.00 ! NA, sugar
 CG201 CG311 NG251 CG201 0.2000 1 180.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr., 3/3/93c
 CG201 CG311 NG251 HGP1 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 CG203 CG311 NG251 CG201 0.2000 1 180.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr., 3/3/93c
 CG203 CG311 NG251 HGP1 1.8000 1 0.00 ! PROT adm jr. 4/05/91, for asp.asn.glu.gln and cters
 CG311 CG311 NG251 CG201 1.8000 1 0.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr., 3/3/93c
 CG311 CG311 NG251 HGP1 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 CG321 CG311 NG251 CG201 1.8000 1 0.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr., 3/3/93c
 CG321 CG311 NG251 HGP1 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 CG331 CG311 NG251 CG201 1.8000 1 0.00 ! PROT ala dipeptide, new C VDW Rmin, adm jr., 3/3/93c
 CG331 CG311 NG251 HGP1 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 HGA1 CG311 NG251 CG201 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 HGA1 CG311 NG251 HGP1 0.0000 1 0.00 ! PROT Alanine Dipeptide ab initio calc's (LK)
 CG201 CG311 OGC301 CG311 0.2000 3 0.00 ! Molecu, from CG203 CG311 OGC301 CG331, PENALTY= 15
 CG311 CG311 OGC301 CG311 0.4000 1 0.00 ! Molecu, from CG331 CG311 OGC301 CG331, PENALTY= 3
 CG311 CG311 OGC301 CG311 0.4000 1 0.00 ! Molecu, from CG331 CG311 OGC301 CG331, PENALTY= 3
 CG331 CG311 OGC301 CG311 0.4900 3 0.00 ! Molecu, from CG331 CG311 OGC301 CG331, PENALTY= 1.5
 CG331 CG311 OGC301 CG311 0.4900 3 0.00 ! Molecu, from CG331 CG311 OGC301 CG331, PENALTY= 1.5
 OGC3C61 CG311 OGC301 CG311 0.5100 1 0.00 ! Molecu, from OGC301 CG301 OGC301 CG331, PENALTY= 16
 OGC3C61 CG311 OGC301 CG311 0.6700 2 0.00 ! Molecu, from OGC301 CG301 OGC301 CG331, PENALTY= 16
 OGC3C61 CG311 OGC301 CG311 0.2800 3 0.00 ! Molecu, from OGC301 CG301 OGC301 CG331, PENALTY= 16
 HGA1 CG311 OGC301 CG311 0.2800 3 0.00 ! Molecu, from HGA1 CG311 OGC301 CG331, PENALTY= 1.5
 CG311 CG311 OGC303 PG1 0.4000 1 180.00 ! Molecu, from CG331 CG311 OGC303 PG2, PENALTY= 2.5
 CG311 CG311 OGC303 PG1 0.3000 2 0.00 ! Molecu, from CG331 CG311 OGC303 PG2, PENALTY= 2.5
 CG311 CG311 OGC303 PG1 0.1000 3 0.00 ! Molecu, from CG331 CG311 OGC303 PG2, PENALTY= 2.5
 OGC3C61 CG311 OGC303 PG1 0.4000 1 180.00 ! Molecu, from CG331 CG311 OGC303 PG2, PENALTY= 46.9
 OGC3C61 CG311 OGC303 PG1 0.3000 2 0.00 ! Molecu, from CG331 CG311 OGC303 PG2, PENALTY= 46.9
 OGC3C61 CG311 OGC303 PG1 0.1000 3 0.00 ! Molecu, from CG331 CG311 OGC303 PG2, PENALTY= 46.9
 HGA1 CG311 OGC303 PG1 0.0000 3 0.00 ! Molecu, from HGA1 CG311 OGC303 PG2, PENALTY= 1
 CG311 CG311 OGC311 HGP1 1.3300 1 0.00 ! PROT PRO2, 2-propanol OH hf/torsional surface, adm jr., 3/2/93
 CG311 CG311 OGC311 HGP1 0.1800 2 0.00 ! PROT PRO2, 2-propanol OH hf/torsional surface, adm jr., 3/2/93
 CG311 CG311 OGC311 HGP1 0.4600 3 0.00 ! PROT PRO2, 2-propanol OH hf/torsional surface, adm jr., 3/2/93
 HGA1 CG311 OGC311 HGP1 0.0000 3 0.00 ! NA backbone.
 CG311 CG311 OGC3C61 CG311 0.5300 1 180.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 5.2
 CG311 CG311 OGC3C61 CG311 0.6800 2 0.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 5.2
 CG311 CG311 OGC3C61 CG311 0.2100 3 180.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 5.2
 CG311 CG311 OGC3C61 CG311 0.1500 4 0.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 5.2
 CG321 CG311 OGC3C61 CG311 0.5300 1 180.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 4.6
 CG321 CG311 OGC3C61 CG311 0.6800 2 0.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 4.6
 CG321 CG311 OGC3C61 CG311 0.2100 3 180.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 4.6
 CG321 CG311 OGC3C61 CG311 0.1500 4 0.00 ! Molecu, from CG321 CG321 OGC3C61 CG321, PENALTY= 4.6
 OGC301 CG311 OGC3C61 CG311 1.0000 3 0.00 ! Molecu, from OGC3C61 CG321 OGC3C61 CG321, PENALTY= 11.1
 OGC303 CG311 OGC3C61 CG311 1.0000 3 0.00 ! Molecu, from OGC3C61 CG321 OGC3C61 CG321, PENALTY= 11.1
 HGA1 CG311 OGC3C61 CG311 0.1950 3 0.00 ! Molecu, from HGA2 CG321 OGC3C61 CG321, PENALTY= 4.6
 CG201 CG321 CG321 CG201 0.2000 3 0.00 ! Molecu, from CG201 CG321 CG321 CG203, PENALTY= 33
 CG201 CG321 CG321 HGA2 0.1950 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG201 CG321 CG321 CG311 0.1950 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG201 CG321 CG321 HGA2 0.1950 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG311 CG321 CG321 CG321 0.5000 3 0.00 ! CARBOCY carboyclic sugars
 CG311 CG321 CG321 CG321 0.5000 6 180.00 ! CARBOCY carboyclic sugars
 CG311 CG321 CG321 HGA2 0.1950 3 0.00 ! NA abasic nucleoside
 CG321 CG321 CG321 CG324 0.1950 3 0.00 ! FLAVOP PIP1,2,3
 CG321 CG321 CG321 HGA2 0.1950 3 0.00 ! LIPID alkanes
 CG324 CG321 CG321 HGA2 0.1950 3 0.00 ! FLAVOP PIP1,2,3
 HGA2 CG321 CG321 HGA2 0.2200 3 0.00 ! LIPID alkanes
 CG321 CG321 CG324 NG3P3 0.1950 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 CG321 CG321 CG324 HGA2 0.1950 3 0.00 ! FLAVOP PIP1,2,3
 HGA2 CG321 CG324 NG3P3 0.1950 3 0.00 ! PROT alkane update, adm jr., 3/2/92
 HGA2 CG321 CG324 HGA2 0.1950 3 0.00 ! FLAVOP PIP1,2,3
 CG2D1 CG321 OGC303 PG1 0.6000 1 180.00 ! Molecu, from CG321 CG321 OGC303 PG1, PENALTY= 70
 CG2D1 CG321 OGC303 PG1 0.6500 2 0.00 ! Molecu, from CG321 CG321 OGC303 PG1, PENALTY= 70
 CG2D1 CG321 OGC303 PG1 0.0500 3 0.00 ! Molecu, from CG321 CG321 OGC303 PG1, PENALTY= 70
 HGA2 CG321 OGC303 PG1 0.0000 3 0.00 ! NA dimp (Reorganization: PG and others
 CG311 CG321 OGC311 HGP1 1.1300 1 0.00 ! og ethanol

```

CG311 CG321 OG311 HGP1 0.1400 2 0.00 ! og ethanol
CG311 CG321 OG311 HGP1 0.2400 3 0.00 ! og ethanol
HGA2 CG321 OG311 HGP1 0.1800 3 0.00 ! og methanol
CG321 CG324 NG3P3 HGP2 0.1000 3 0.00 ! PROT 0.715->0.10 METHYLAMMONIUM (KK)
HGA2 CG324 NG3P3 HGP2 0.1000 3 0.00 ! PROT 0.715->0.10 METHYLAMMONIUM (KK)
CG311 OG303 PG1 OG2P1 0.1000 3 0.00 ! Molecu , from CG321 OG303 PG1 OG2P1, PENALTY= 0.6
CG311 OG303 PG1 OG304 1.2000 1 180.00 ! Molecu , from CG321 OG303 PG1 OG304, PENALTY= 0.6
CG311 OG303 PG1 OG304 0.1000 2 180.00 ! Molecu , from CG321 OG303 PG1 OG304, PENALTY= 0.6
CG311 OG303 PG1 OG304 0.1000 3 180.00 ! Molecu , from CG321 OG303 PG1 OG304, PENALTY= 0.6
CG321 OG303 PG1 OG2P1 0.1000 3 0.00 ! NA dmp !Reorganization:PC and others
CG321 OG303 PG1 OG304 1.2000 1 180.00 ! NA dmp !Reorganization:ADP
CG321 OG303 PG1 OG304 0.1000 2 180.00 ! NA dmp !Reorganization:ADP
CG321 OG303 PG1 OG304 0.1000 3 180.00 ! NA dmp !Reorganization:ADP
PG1 OG304 PG1 OG2P1 0.1000 2 0.00 ! NA ppi2 !Reorganization:PPI2
PG1 OG304 PG1 OG2P1 0.0300 3 0.00 ! NA ppi2 !Reorganization:PPI2
PG1 OG304 PG1 OG303 0.0300 2 0.00 ! NA ppi2 !Reorganization:PPI2
PG1 OG304 PG1 OG303 0.0300 3 0.00 ! NA ppi2 !Reorganization:PPI2

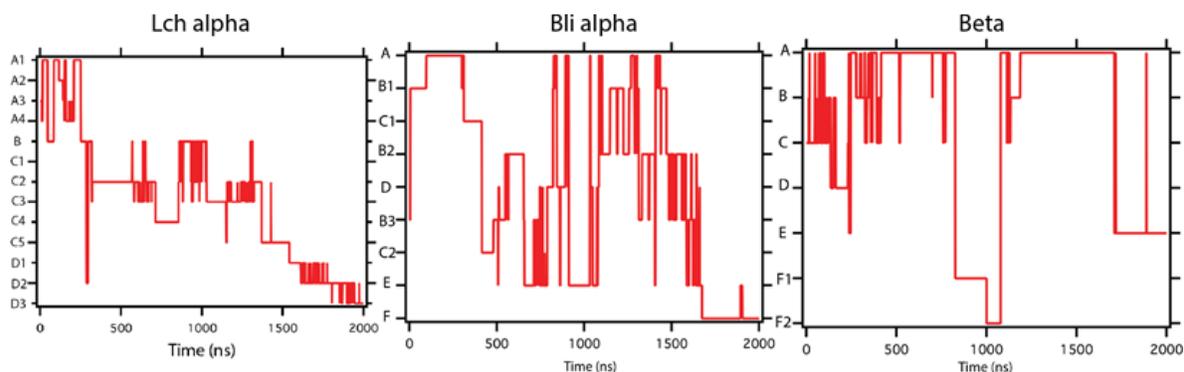
IMPROPERS
CG201 CG311 NG2S1 OG2D1 120.0000 0 0.00 ! PROT NMA Vibrational Modes (LK) WILDCARD
CG201 CG321 NG2S1 OG2D1 120.0000 0 0.00 ! PROT NMA Vibrational Modes (LK) WILDCARD
CG201 CG331 NG2S1 OG2D1 120.0000 0 0.00 ! PROT NMA Vibrational Modes (LK) WILDCARD
CG203 OG2D2 OG2D2 CG311 96.0000 0 0.00 ! PROT 90.0->96.0 acetate, single impr (KK) correct conversion from wildcard CC X X CT1

END
RETURN

```

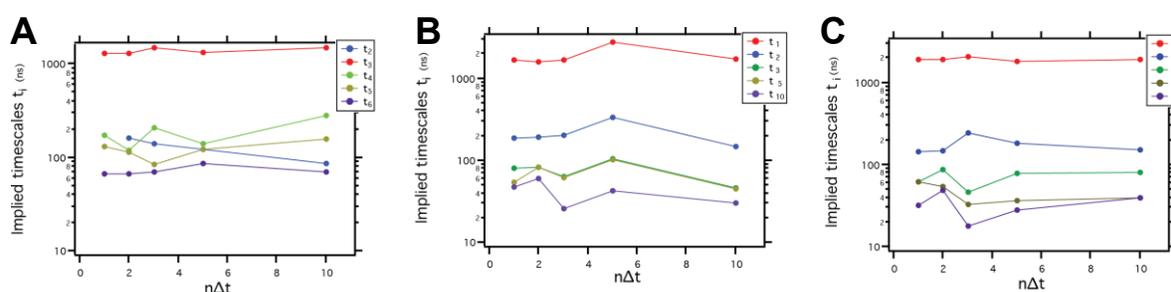
10.4 Markov State Models

10.4.1 Time Series of MSM Macrostates for the 2 μ s MD Simulation



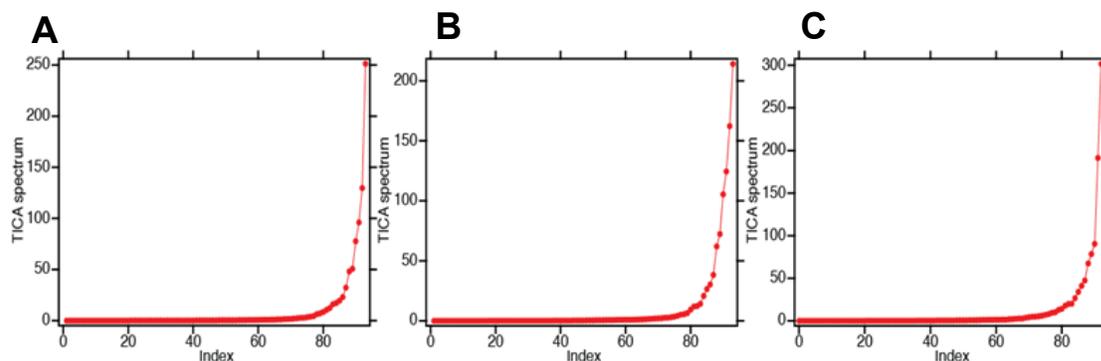
10.4.2 Validation: Chapman-Kolmogorov Test

The Markov state models for the three peptides were validated by plotting the implied timescales for different lag times. The employed timescales are calculated as: $t_i = \tau / \log(\lambda_i)$, where λ_i are the eigenvalues of the transition matrix. The Markov state models in the manuscript are estimated with a lag time of 1 ns, which is the sampling time employed in the MD simulations. Here, we plot the implied timescales for lag times of 1, 2, 3, 5 and 10 ns. Clearly, the timescales level for the employed lag time, validating the markovianity of the calculated models. A) Lch α , B) Bli α C) Lch β .

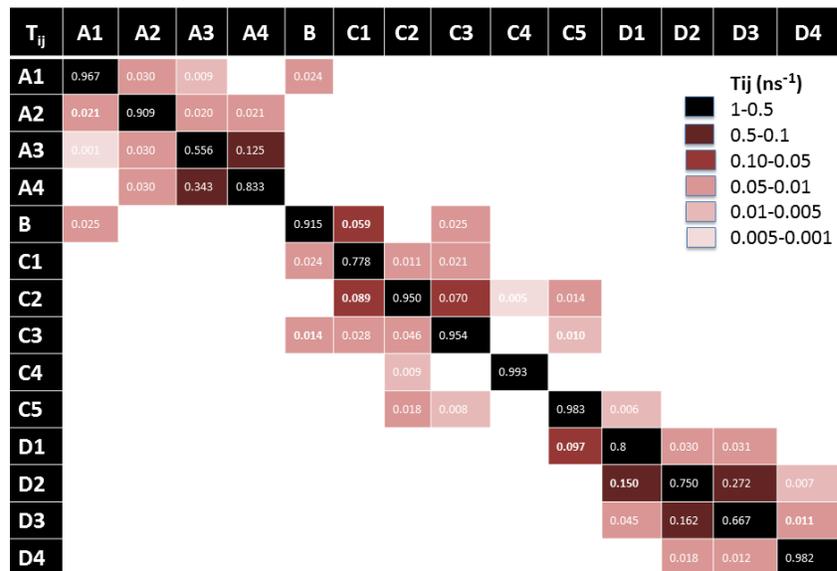


10.4.3 TICA Spectrum for the Three Peptides.

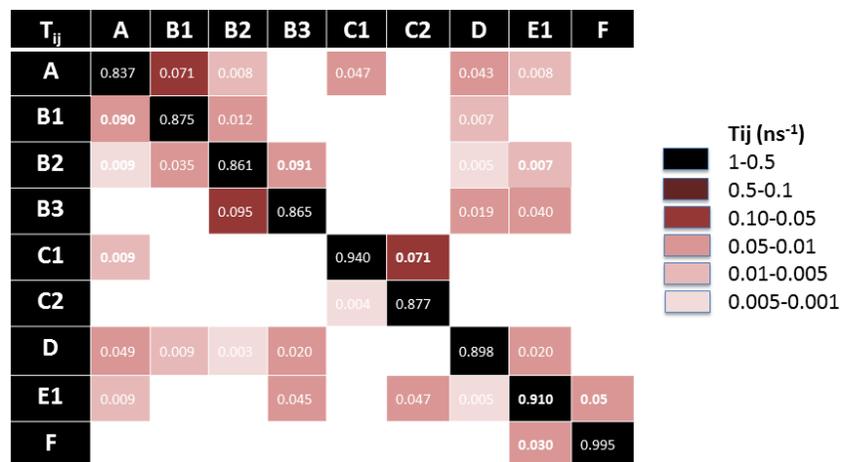
Plot of the TICA eigenvalues for growing magnitudes for Lch α (A), Bli α (B) and Lch β (C). Each system is represented by the linear subspace defined by the first three components. As seen in the plot, that covariance shows a strong nonlinear behavior and these three first components accounts for more than 50% of the fluctuations of the system.



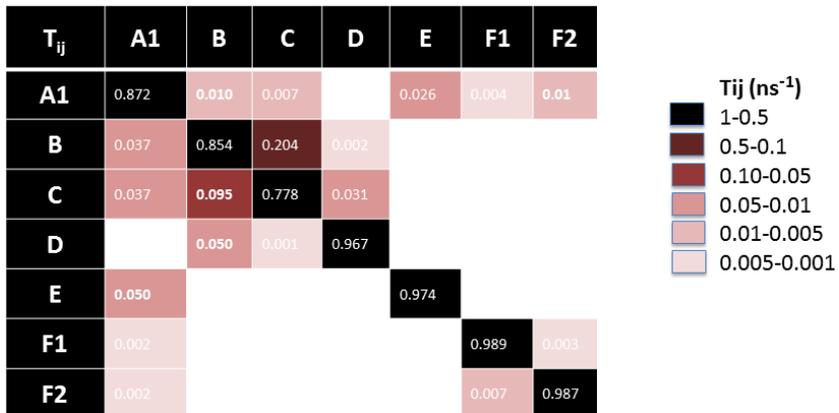
10.4.4 Transition Matrix Lch α



10.4.5 Transition Matrix Bli α



10.4.6 Transition Matrix Lch β

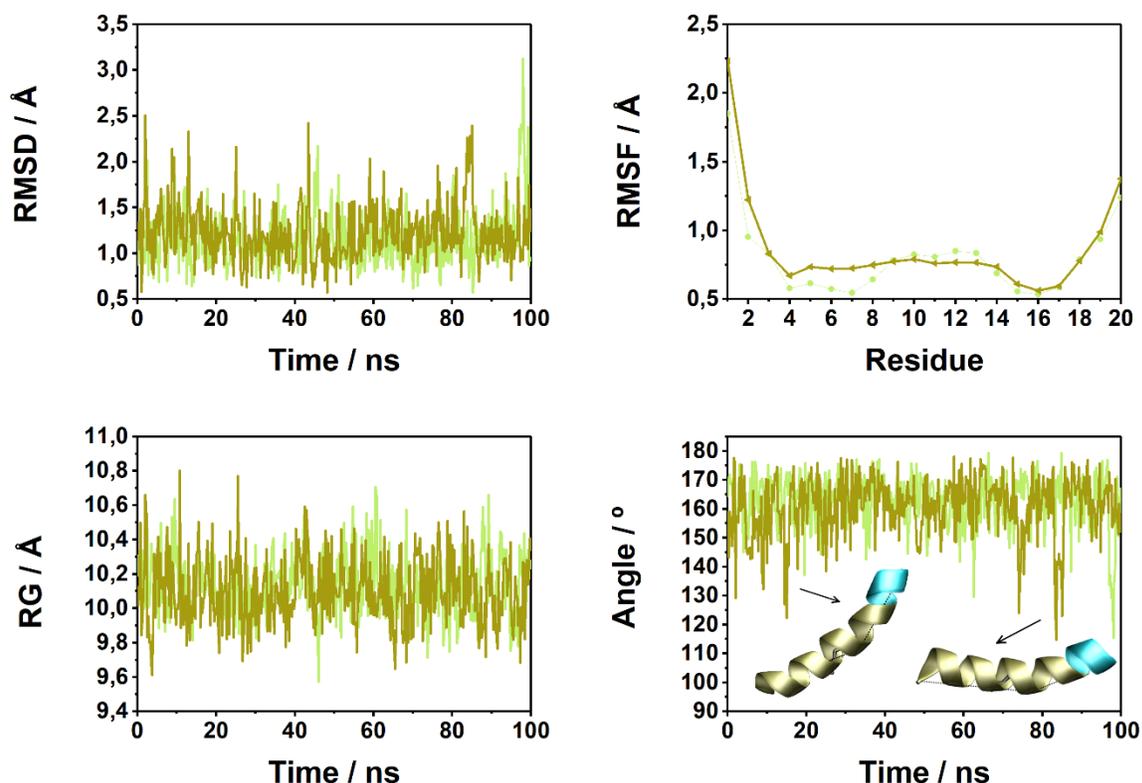


11. Appendix – Cathelicidin LL-37 and fragments

11.1 All-Atom MD Peptide S9V-LL-20

The all-atom MD simulation regarding the peptide S9V-LL-20 was carried out in agreement with the protocol detailed in 9.3.1 Conformational Dynamics of LL-37 and fragment LL-32, LL-20. Despite that the overall behaviour remains quite similar, the S9V-LL-20 is slightly more rigid than the parent peptide, as it is reflected in the steady values observed for the C_{α} -RMSF. The bending around the residues 1, 12 and 20 was also monitored and compared to LL-20, therefore in order to exemplify its behaviour only the snapshot mutant peptide is displayed in the plot.

11.1.1 S9V-LL-20 Comparative Analysis with Fragment LL-20



11.2 Overview of the Infrared Absorption Bands (in cm^{-1}) of the Amino Acid Residues of LL-37 and its Fragments.

Amino acids showing an absorption in the 1800 – 1500 cm^{-1} range are highlighted in bold. The C=O vibration of Asp and Glu are shown in grey, since these absorptions only occur in the protonated form, which is not observed in this work. Intensities of bands relate to extinction coefficients of 500 – 400 $\text{M}^{-1}\text{cm}^{-1}$ (vs), 400 – 200 $\text{M}^{-1}\text{cm}^{-1}$ (s), and 200 – 0 $\text{M}^{-1}\text{cm}^{-1}$ (m). Positions and intensities are obtained from Barth, 2007, *Biochim. Biophys. Acta*, 1767, 1073-1101.

Amino acid	Amide I range			Amide II range	
	> 1700	1700-1650	1650-1600	1600-1550	1550-1500
1 L – Leu					
2 L – Leu					
3 G – Gly					
4 D – Asp	(~1716)			1575 ± 50 (s)	
5 F – Phe					
6 F – Phe					
7 R – Arg		1695 – 1652 (vs)	1663 – 1614 (s)		
8 K – Lys			~ 1627 (m)		~ 1526 (m)
9 S – Ser					
10 K – Lys					
11 E – Glu	(~1712)			1560 ± 50 (vs)	
12 K – Lys			~ 1627 (m)		~ 1526 (m)
13 I – Ile					
14 G – Gly					
15 K – Lys			~ 1627 (m)		~ 1526 (m)
16 E – Glu	(~1712)			1560 ± 50 (vs)	
17 F – Phe					
18 K – Lys			~ 1627 (m)		~ 1526 (m)
19 R – Arg		1695 – 1652 (vs)	1663 – 1614 (s)		
20 I – Ile					
21 V – Val					
22 Q – Gln		1687 – 1668 (s)	1610- 1586 (s)		
23 R – Arg		1695 – 1652 (vs)	1663 – 1614 (vs)		
24 I – Ile					
25 K – Lys			~ 1627 (m)		~ 1526 (m)
26 D – Asp	(~1716)			1575 ± 50 (s)	
27 F – Phe					
28 L – Leu					
29 R – Arg		1695 – 1652 (vs)	1663 – 1614 (s)		
30 N – Asn		1678 – 1677 (s)	1622 – 1612 (m)		

31	L – Leu			
32	V – Val			
33	P – Pro			
34	R – Arg		1695 – 1652 (vs)	1663 – 1614 (s)
35	T – Thr			
36	E – Glu	(~1712)		1560 ± 50 (vs)
37	S – Ser			

12. List of Abbreviations

AMR	Antimicrobial resistance
WHO	World Health Organization
AMPs	Antimicrobial peptides
HDPs	Host defense peptides
MD	Molecular dynamics
MSMs	Markov state models
SEIRA	Surface-enhanced infrared absorption spectroscopy
FF	force field
CHARMM	Chemistry harvard molecular mechanics
GROMOS	Groningen molecular simulation computer program
AMBER	Assisted model building with energy refinement
OPLS-AA	Optimized potentials for liquid simulations all-atom
CMAP	Correction maps
L-J	Lennard-Jones potential
PBC	Periodic boundary conditions
SBC	Stochastic boundary conditions
3D	Three-dimensional
2D	Two-dimensional
Eq.	Equation
PME	Particle mesh Ewald
PDB	Protein data bank
NAMD	Nanoscale molecular dynamics
<i>NVE</i>	Microcanonical ensemble
<i>NVT</i>	Canonical ensemble

<i>NPT</i>	Isobaric-Isothermal ensemble
μPT	Grand canonical ensemble
SD	Steepest descend
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RG	Radius of gyration
μ	Dipolar moment
COM	Center of mass
COG	Center of geometry
IE	Interaction Energy
TdCM	Time-dependent contact map
RDF	Radial distribution function
VMD	Visual molecular dynamics
IR	Infrared radiation
tBLM	Tethered bilayer lipid membrane
EM	Electromagnetic mechanism
6MH	6-mercaptohexanol
WK3SH	Cholestanyl headgroup
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
FT-IR	Fourier-transform infrared transmission
SAM	Tethered mixed self-assembled monolayer
EIS	Electrochemical impedance spectroscopy
ATR-IR	Attenuated total reflection infrared
MCT	Mercury Cadmium Telluride
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

AA	Amino acid
CGenFF	Charmm general force field
PL	Phospholipids
IM	Inner membrane
PM	Plasma membrane (cytoplasmic membrane)
OM	Outer membrane
PGN	Peptidoglycan
LPS	Lipopolysaccharide
LPP	Lipoproteins
GlcNAc	<i>N</i> -acetylglucosamine
MurNAc	<i>N</i> -acetylmuramic
POPS	Phosphatidylserine
C55	Polyisoprenoid carriers C55
TAs	Teichoic acids
WTAs	Wall teichoic acids
LTAs	Lipoteichoic acids
DAP	<i>meso</i> -diaminopimelinacid
LAB	Lactic acid bacteria
RiPPs	Posttranslationally modified peptides
PTM	Posttranslational modifications
Lan	<i>Meso</i> -lanthionine
Mlan	3-methylanthionine
Dha	2,3-dehydroalanine
Dhb	(<i>Z</i>)-2,3-dehydrobutyrine
VdWs	Van der Waals

13. References

- [1] W. H. O. (WHO), “World Health Organization (WHO),” can be found under <http://www.euro.who.int/en/home>, **2017**.
- [2] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, et al., *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- [3] D. A. Case, T. E. Cheatham Iii, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- [4] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- [5] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- [6] N. J. Santora, R. W. Johnson, A. R. Leach, W. Addison, *Molecular Modeling: Principles and Applications*, **1996**.
- [7] C. L. Brooks Iii, M. Karplus, M. Karplus, *J. Chem. Phys.* **1983**, *791*, DOI 10.1063/1.448118.
- [8] G. E. Uhlenbeck, L. S. Ornstein, *Phys Rev* **1930**, *36*, 823–841.
- [9] S. A. Adelman, C. L. Brooks, *J. Phys. Chem* **1982**, *86*, 1511–1524.
- [10] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, **1987**.
- [11] L. Verlet, *Phys. Rev.* **1967**, *159*, 98–103.
- [12] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *981*, 10089–10092.
- [13] M. Karplus, J. A. McCammon, *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- [14] A. Langham, Y. N. Kaznessis, *Methods Mol. Biol.* **2010**, *618*, 267–285.
- [15] “RCSB Protein Data Bank - RCSB PDB,” can be found under <https://www.rcsb.org/pdb/home/home.do>, **2017**.
- [16] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- [17] D. Beeman, *J. Comput. Phy.* **1976**, *20*, 130–139.
- [18] Mark E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, **2010**.
- [19] S. E. Feller, Y. Zhang, R. W. Pastor, B. R. Brooks, *J. Chem. Phys.* **1995**, *1031*, 4613–4621.
- [20] S. Nose, *J. Chem. Phys.* **1984**, *81*, 511–519.
- [21] S. E. Feller, Y. Zhang, R. W. Pastor, B. R. Brooks, *J. Chem. Phys.* **1995**, *103*, 4613.
- [22] G. J. Martyna, D. J. Tobias, M. L. Klein, *J. Chem. Phys.* **1994**, *101*, 4177.

- [23] M. A. Cauchy, *Comptes Rendus Hebd. Séances Acad. Sci. OC I n.d.*, 25, 536–538.
- [24] T. A. Straeter, *Tech. report, NASA Tech. Rep. Serv.* **1971**.
- [25] M. Heinig, D. Frishman, *Nucleic Acids Res.* **2004**, 32, W500-2.
- [26] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, 14, 33–8, 27–8.
- [27] F. Bowman, Gregory R., Pande, Vijay S., Noé, *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, Springer Netherlands, Dordrecht, **2014**.
- [28] A. Amadei, A. Limddrn, *Strmct.Funct.Genet* **1993**, 17, 412–425.
- [29] G. Pérez-Hernández, F. Paul, T. Giorgino, G. De Fabritiis, F. Noé, *J. Chem. Phys.* **2013**, 139, 15102.
- [30] D. Prada-Gracia, J. Gómez-Gardeñes, P. Echenique, F. Falo, *PLoS Comput Biol* **2009**, 5, DOI 10.1371/journal.pcbi.1000415.
- [31] F. Siebert, P. Hildebrandt, *Vibrational Spectroscopy in Life Science*, Wiley-VCH, **2008**.
- [32] E. B. Wilson, J. C. Decius, P. C. Cross, *Molecular Vibrations : The Theory of Infrared and Raman Vibrational Spectra*, Dover Publications, **1980**.
- [33] A. Hartstein, J. R. Kirtley, J. C. Tsang, *Phys Rev* **1980**, 45, 201–204.
- [34] R. Aroca, D. J. Ross, C. Domingo, *Spectrosc. Eur.* **2012**, 24, DOI 10.1366/0003702042475420.
- [35] M. Osawa, *Handb. Vib. Spectrosc.* **2002**, 1–15.
- [36] E. Goormaghtigh, V. Raussens, J.-M. Ruyschaert, *Biochim. Biophys. Acta - Rev. Biomembr.* **1999**, 1422, 105–185.
- [37] J. A. Kozuch, Structure-Function Relationships of Membrane Proteins - Spectroelectrochemical Investigation of Artificial Membranes Vorgelegt von Diplom-Chemiker Jacek Artur Kozuch, **2012**.
- [38] S. Wiebalck, J. Kozuch, E. Forbrig, C. C. Tzschucke, L. J. C. Jeuken, P. Hildebrandt, *J. Phys. Chem. B* **2016**, 120, 2249–2256.
- [39] A. S. Achalkumar, R. J. Bushby, S. D. Evans, *Soft Matter* **2010**, 6, 6036–6051.
- [40] J. Kozuch, C. Weichbrodt, D. Millo, K. Giller, S. Becker, P. Hildebrandt, C. Steinem, *Phys. Chem. Chem. Phys.* **2014**, 16, 9546–55.
- [41] H. Miyake, S. Ye, M. Osawa, *Electrochem. commun.* **2002**, 4, 973–977.
- [42] D. L. Nelson, M. M. Cox, *Lehninger Biochemie*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2001**.
- [43] M. T. Madigan, J. M. Martinko, J. Parker, *Brock Biology of Microorganisms*, Prentice Hall/Pearson Education, Upper Saddle River, NJ, **n.d.**
- [44] Y Tambe, “Weblet Importer,” can be found under <https://commons.wikimedia.org/w/index.php?curid=9999296>, **n.d.**

- [45] N. Malanovic, K. Lohner, *Biochim. Biophys. Acta* **2016**, *1858*, 936–946.
- [46] T. J. Silhavy, D. Kahne, S. Walker, **n.d.**, DOI 10.1101/cshperspect.a000414.
- [47] F. C. Neuhaus, J. Baddiley, *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 686.
- [48] K. Matsuzaki, O. Murase, H. Tokuda, N. Fujii, K. Miyajima, *Biochemistry* **1996**, *35*, 11361–11368.
- [49] M. Dathe, T. Wieprecht, H. Nikolenko, L. Handel, W. L. Maloy, D. L. MacDonald, M. Beyermann, M. Bienert, *FEBS Lett.* **1997**, *403*, 208–212.
- [50] M. Dathe, H. Nikolenko, J. Meyer, M. Beyermann, M. Bienert, *FEBS Lett.* **2001**, *501*, 146–150.
- [51] T. Wieprecht, M. Dathe, E. Krause, M. Beyermann, W. L. Maloy, D. L. Macdonald, M. Bienert, *FEBS Lett.* **1997**, *417*, 135–140.
- [52] R. E. W. Hancock, *Lancet* **1997**, *349*, 418–422.
- [53] L. Zhang, P. Dhillon, H. Yan, S. Farmer, R. E. W. Hancock, **2000**, *44*, 3317–3321.
- [54] M. R. Yeaman, N. Y. Yount, *Pharmacol. Rev.* **2003**, *55*, 27–55.
- [55] S. D. Zakharov, E. A. Kotova, Y. N. Antonenko, W. A. Cramer, **2004**, DOI 10.1016/j.bbamem.2004.07.001.
- [56] L. Yang, T. A. Harroun, T. M. Weiss, L. Ding, H. W. Huang, *Biophys. J.* **2001**, *81*, 1475–1485.
- [57] L. Yang, T. M. Weiss, R. I. Lehrer, H. W. Huang, **n.d.**
- [58] F. Costa, I. F. Carvalho, R. C. Montelaro, P. Gomes, M. C. L. Martins, *Acta Biomater.* **2011**, *7*, 1431–1440.
- [59] D. Drider, S. Rebuffat, *Prokaryotic Antimicrobial Peptides: From Genes to Applications*, **2011**.
- [60] E. M. O'Connor, R. F. Shand, *J. Ind. Microbiol. Biotechnol.* **2002**, *28*, 23–31.
- [61] E. L. De Souza, C. A. Da Silva, C. P. De Sousa, *Brazilian Arch. Biol. Technol.* **2005**, *48*, 559–566.
- [62] J. Cleveland, T. J. Montville, I. F. Nes, M. L. Chikindas, *Int. J. Food Microbiol.* **2001**, *71*, 1–20.
- [63] R. W. Jack, J. R. Tagg, B. Ray, *Microbiol. Rev.* **1995**, *59*, 171–200.
- [64] L. De Vuyst, E. Vandamme, *Bacteriocins of Lactic Acid Bacteria: Microbiology, Genetics and Applications*, **1994**.
- [65] A. Geis, J. Singh, M. Teuber, *Appl. Environ. Microbiol.* **1983**, *45*, 205–211.
- [66] T. Klaenhammer, *FEMS Microbiol. Rev.* **1993**, *12*, 39–85.
- [67] A. Zouhir, R. Hammami, I. Fliss, J. Ben Hamida, *Protein J.* **2010**, *29*, 432–439.
- [68] K. Meindl, T. Schmiederer, K. Schneider, A. Reicke, D. Butz, S. Keller, H. Gühring, L. Vértesy, J. Wink, H. Hoffmann, et al., *Angew. Chemie - Int. Ed.* **2010**, *49*, 1151–1154.

- [69] P. J. Knerr, W. A. van der Donk, *Annu. Rev. Biochem.* **2012**, *81*, 479–505.
- [70] Y. Goto, B. Li, J. Claesen, Y. Shi, M. J. Bibb, W. A. van der Donk, *PLoS Biol.* **2010**, *8*, 4–13.
- [71] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, et al., *Nat. Prod. Rep.* **2013**, *30*, 108–160.
- [72] G. Jung, H.-G. Sahl, *Nisin and Novel Lantibiotics : Proceedings of the First International Workshop on Lantibiotics, April 15-18, 1991, Physikzentrum Bad Honnef, F.R.G., ESCOM, 1991.*
- [73] C. Piper, P. Cotter, R. Ross, C. Hill, *Curr. Drug Discov. Technol.* **2009**, *6*, 1–18.
- [74] P. D. Cotter, C. Hill, R. P. Ross, *Nat. Rev. Microbiol.* **2005**, *3*, 777–788.
- [75] D. Twomey, R. P. Ross, M. Ryan, B. Meaney, C. Hill, *Antonie Van Leeuwenhoek* **2002**, *82*, 165–85.
- [76] U. Pag, H.-G. Sahl, *Curr. Pharm. Des.* **2002**, *8*, 815–833.
- [77] J. M. Willey, W. A. van der Donk, *Annu. Rev. Microbiol.* **2007**, *61*, 477–501.
- [78] N. C. K. Heng, P. A. Wescombe, J. P. Burton, R. W. Jack, J. R. Tagg, in *Bacteriocins*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2007**, pp. 45–92.
- [79] E. Breukink, B. de Kruijff, *Nat. Rev. Drug Discov.* **2006**, *5*, 321–332.
- [80] E. Gross, J. L. Morell, L. C. Craig, **n.d.**
- [81] M. Kashiri, J. P. Cerisuelo, I. Domínguez, G. López-carballo, V. Muriel-gallet, R. Gavara, P. Hernández-muñoz, **2017**, 1–56.
- [82] A. Gharsallaoui, N. Oulahal, C. Joly, P. Degraeve, *Crit. Rev. Food Sci. Nutr.* **2016**, *56*, 1262–74.
- [83] C. Piper, C. Hill, P. D. Cotter, R. P. Ross, *Microb. Biotechnol.* **2010**, *4*, 375–382.
- [84] R. E. Wirawan, N. a Klesse, R. W. Jack, J. R. Tagg, **2006**, *72*, 1148–1156.
- [85] H. S. Rollema, O. P. Kuipers, P. Both, W. M. De Vos, R. J. Siezen, *Appl. Environ. Microbiol.* **1995**, *61*, 2873–2878.
- [86] Q. Zhang, Y. Yu, J. E. Vélasquez, W. A. van der Donk, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 18361–6.
- [87] Z. Wu, W. Wang, M. Tang, J. Shao, C. Dai, W. Zhang, H. Fan, H. Yao, J. Zong, D. Chen, et al., *Gene* **2014**, *535*, 156–164.
- [88] C. M. Guinane, P. D. Cotter, R. P. Ross, C. Hill, **2015**, *81*, 3953–3960.
- [89] H. -G Sahl, R. W. Jack, G. Bierbaum, *Eur. J. Biochem.* **1995**, *230*, 827–853.
- [90] D. Field, P. M. O. Connor, P. D. Cotter, C. Hill, R. P. Ross, *Mol. Microbiol.* **2008**, *69*, 218–230.
- [91] I. Martin, J.-M. Ruyschaerti, D. Sanders, C. J. Giffard, *Eur. J. Biochem.* **1996**, *239*, 156–164.
- [92] D. Field, C. Hill, P. D. Cotter, R. P. Ross, *Mol. Microbiol.* **2010**, *78*, 1077–1087.

- [93] H. S. Rollema, O. P. Kuipers, P. Both, W. M. De Vos, R. J. Siezen, *Appl. Environ. Microbiol.* **1995**, *61*, 2873–2878.
- [94] W. C. Chan, H. M. Dodd, N. Horn, K. Maclean, L. Y. Lian, B. W. Bycroft, L. Lian, *Appl. Environ. Microbiol.* **1996**, *62*, 2966–9.
- [95] E. Breukink, C. Van Kraaij, R. A. Demel, R. J. Siezen, O. P. Kuipers, B. De Kruijff, *Biochemistry* **1997**, *36*, 6968–6976.
- [96] C. J. Giffard, H. M. Dodd, N. Horn, S. Ladha, A. R. Mackie, A. Parr, M. J. Gasson, D. Sanders, *Biochemistry* **1997**, *36*, 3802–3810.
- [97] E. Breukink, B. De Kruijff, *Biochim. Biophys. Acta - Biomembr.* **1999**, *1462*, 223–234.
- [98] K. Matsuzaki, O. Murase, N. Fujii, K. Miyajima, *Biochemistry* **1995**, *34*, 6521–6526.
- [99] C. Van Kraaij, E. Breukink, M. A. Noordermeer, R. A. Demel, R. J. Siezen, O. P. Kuipers, B. De Kruijff, **n.d.**, DOI 10.1021/bi980931b.
- [100] K. A. Brogden, *Nat. Rev. Microbiol.* **2005**, *3*, 238–250.
- [101] P. E. Linnett, J. L. Strominger, *Antimicrob. Agents Chemother.* **1973**, *4*, 231–236.
- [102] M. H. Braff, M. A. Hawkins, A. D. Nardo, B. Lopez-Garcia, M. D. Howell, C. Wong, K. Lin, J. E. Streib, R. Dorschner, D. Y. M. Leung, et al., *J. Immunol.* **2005**, *174*, 4271–4278.
- [103] E. Ruhr, H.-G. Sahl, *Antimicrob. Agents Chemother.* **1985**, *27*, 841–845.
- [104] S.-T. D. Hsu, E. Breukink, E. Tischenko, M. A. G. Lutters, B. de Kruijff, R. Kaptein, A. M. J. J. Bonvin, N. A. J. van Nuland, *Nat. Struct. Mol. Biol.* **2004**, *11*, 963–7.
- [105] M. Thormann, H.-J. Hofmann, *J. Mol. Struct. THEOCHEM* **1998**, *431*, 79–96.
- [106] E. R. Turpin, S. Mulholland, A. M. Teale, B. B. Bonev, J. D. Hirst, S. M. Asaduzzaman, K. Sonomoto, G. Bierbaum, H. G. Sahl, B. B. Bonev, et al., *RSC Adv.* **2014**, *4*, 48621–48631.
- [107] “ParamChem Home,” can be found under <https://www.paramchem.org/>, **n.d.**
- [108] D. J. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb, **2009**.
- [109] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [110] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al., *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- [111] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, et al., *J. Comput. Chem.* **2010**, *31*, 671–690.
- [112] O. M. Becker, A. D. MacKerell, B. Roux, M. Watanabe, *Computational Biochemistry and Biophysics*, M. Dekker, **2001**.
- [113] J.-P. Ryckaert, G. Ciccotti+, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 321–341.

- [114] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, et al., *J. Comput. Chem.* **2014**, *35*, 1997–2004.
- [115] D. C. Koch, T. H. Schmidt, H.-G. Sahl, U. Kubitscheck, C. Kandt, *BBA - Biomembr.* **2014**, *1838*, 3061–3068.
- [116] H. B. Brötz, G. Bierbaum, K. Leopold, P. E. Reynolds, H.-G. Sahl, **1998**, *42*, 154–160.
- [117] H. W. Van Den Hooven, F. M. Lagerwerf, W. Heerma, J. Haverkamp, J. C. Piard, C. W. Hilbers, R. J. Siezen, O. P. Kuipers, H. S. Rollema, *FEBS Lett.* **1996**, *391*, 317–322.
- [118] Z. O. Shenkarev, E. I. Finkina, E. K. Nurmukhamedova, S. V. Balandin, K. S. Mineev, K. D. Nadezhdin, Z. A. Yakimenko, A. A. Tagaev, Y. V. Temirov, A. S. Arseniev, et al., *Biochemistry* **2010**, *49*, 6462–6472.
- [119] T. Caetano, J. M. Krawczyk, E. Mösker, R. D. Süßmuth, S. Mendo, *Chem. Biol.* **2011**, *18*, 90–100.
- [120] H. Holo, Z. Jeknic, M. Daeschel, S. Stevanovic, I. F. Nes, *Microbiology* **2001**, *147*, 643–651.
- [121] N. Sawa, P. Wilaipun, S. Kinoshita, T. Zendo, V. Leelawatcharamas, J. Nakayama, K. Sonomoto, *Appl. Environ. Microbiol.* **2012**, *78*, 900–903.
- [122] L. M. Cintas, P. Casaus, H. Holo, P. E. Hernandez, I. F. Nes, L. S. Håvarstein, *J. Bacteriol.* **1998**, *180*, 1988–1994.
- [123] H. Yonezawa, H. K. Kuramitsu, *Antimicrob. Agents Chemother.* **2005**, *49*, 541–548.
- [124] O. Hyink, M. Balakrishnan, J. R. Tagg, *FEMS Microbiol. Lett.* **2005**, *252*, 235–241.
- [125] M. A. D. B. Navaratna, H. G. Sahl, J. R. Tagg, *Infect. Immun.* **1999**, *67*, 4268–4271.
- [126] M. A. D. B. Navaratna, H. Sahl, R. John, **1998**, *64*, 4803–4808.
- [127] I. Wiedemann, T. Bottiger, R. R. Bonelli, A. Wiese, S. O. Hagge, T. Gutschmann, U. Seydel, L. Deegan, C. Hill, P. Ross, et al., *Mol. Microbiol.* **2006**, *61*, 285–296.
- [128] L. E. Cooper, A. L. McClerren, A. Chary, W. A. van der Donk, *Chem Biol.* **2008**, *15*, 1035–1045.
- [129] L. Huo, W. A. Van Der Donk, *J. Am. Chem. Soc.* **2016**, *138*, 5254–5257.
- [130] B. Xin, J. Zheng, H. Liu, J. Li, L. Ruan, D. Peng, M. Sajid, M. Sun, *Front. Microbiol.* **2016**, *7*, 1–12.
- [131] F. W. J. Collins, P. M. O'Connor, O. O'Sullivan, M. C. Rea, C. Hill, R. P. Ross, *Microbiol. (United Kingdom)* **2016**, *162*, 1662–1671.
- [132] M. S. Gilmore, *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **1996**, *69*, 129–138.
- [133] D. Van Tyne, M. J. Martin, M. S. Gilmore, *Toxins (Basel)*. **2013**, *5*, 895–911.
- [134] W. Haas, B. D. Shepard, M. S. Gilmore, *Nature* **2002**, *415*, 84–87.
- [135] Weixin Tang and Wilfred A. van der Donk, *Seq. Enterococcal Cytolysin Imparts Unusual Lanthionine Stereochem.* **2013**, *9*, 157–159.

- [136] X. Zhao, W. A. Van Der Donk, *Cell Chem. Biol.* **2016**, *23*, 246–256.
- [137] L. Y. Geer, A. Marchler-Bauer, R. C. Geer, L. Han, J. He, S. He, C. Liu, W. Shi, S. H. Bryant, *Nucleic Acids Res.* **2010**, *38*, D492–D496.
- [138] M. W. Rey, P. Ramaiya, B. A. Nelson, S. D. Brody-Karpin, E. J. Zaretsky, M. Tang, A. Lopez de Leon, H. Xiang, V. Gusti, I. G. Clausen, et al., *Genome Biol.* **2004**, *5*, R77.
- [139] S. Hwang, Q. Shao, H. Williams, C. Hilty, Y. Q. Gao, *J. Phys. Chem. B* **2011**, *115*, 6653–6660.
- [140] S. Mendo, N. A. Faustino, A. C. Sarmiento, F. Amado, A. J. G. Moir, *Biotechnol. Lett.* **2004**, *26*, 115–119.
- [141] J. Dennington, Roy; Keith, Todd; Millam, **2009**.
- [142] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, F. Noé, *J. Chem. Phys.* **2011**, *134*, 174105.
- [143] J.-L. Blondel, V.D; Guillaume, E. Lambiotte, Renaud; Lefebvre, “Fast unfolding of communities in large networks,” **2008**.
- [144] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, D. E. Shaw, D. J. Muller, *PLoS One* **2012**, *7*, DOI 10.1371/journal.pone.0032131.
- [145] D. S. Chapple, R. Hussain, C. L. Joannou, R. E. W. Hancock, E. Odell, R. W. Evans, G. Siligardi, *Antimicrob. Agents Chemother.* **2004**, *48*, 2190–8.
- [146] N. Zhou, D. P. Tieleman, H. J. Vogel, *BioMetals* **2004**, *17*, 217–223.
- [147] A. P. Boughton, I. Andricioaei, Z. Chen, **n.d.**, DOI 10.1021/la1024394.
- [148] D. P. Tieleman, H. J. C. Berendsen, M. S. P. Sansom, **n.d.**
- [149] A. Chugunov, D. Pyrkova, D. Nolde, A. Polyansky, V. Pentkovsky, R. Efremov, M. M. Shemyakin, Y. A. Ovchinnikov, **n.d.**, DOI 10.1038/srep01678.
- [150] J. Johansson, G. H. Gudmundsson, M. E. Rottenberg, K. D. Berndt, B. Agerberth, *J. Biol. Chem.* **1998**, *273*, 3718–3724.
- [151] N. I. Martin, E. Breukink, *Future Microbiol.* **2007**, *2*, 513–525.
- [152] L. H. Deegan, S. Suda, E. M. Lawton, L. A. Draper, F. Hugenholtz, A. Peschel, C. Hill, P. D. Cotter, R. P. Ross, *Microb. Biotechnol.* **2010**, *3*, 222–234.
- [153] C. Szekat, R. W. Jack, D. Skutlarek, G. Bierbaum, H. Fa, *Appl. Environ. Microbiol.* **2003**, *69*, 3777–3783.
- [154] P. D. Cotter, L. H. Deegan, E. M. Lawton, L. A. Draper, P. M. O’Connor, C. Hill, R. P. Ross, *Mol. Microbiol.* **2006**, *62*, 735–747.
- [155] R. I. Lehrer, T. Ganz, *Curr Opin Hematol* **2002**, *9*, 18–22.
- [156] M. Zanetti, R. Gennaro, D. Romeo, *FEBS Lett.* **1995**, *374*, 1–5.
- [157] L. Tomasinsig, M. Zanetti, *Curr. Protein Pept. Sci.* **2005**, *6*, 23–34.
- [158] R. G. and D. R. M Zanetti, L Litteri, G Griffiths, *J. Immunol.* **1991**, *146*, 4295–4300.

- [159] I. Nagaoka, Y. Tsutsumi-Ishii, S. Yomogida, T. Yamashita, *J. Biol. Chem.* **1997**, *272*, 22742–22750.
- [160] A. Panyutich, J. Shi, P. L. Boutz, C. Zhao, T. Ganz, *Infect. Immun.* **1997**, *65*, 978–985.
- [161] O. E. Sørensen, A. H. Johnsen, J. Calafat, G. S. Tjabringa, P. S. Hiemstra, N. Borregaard, *Am. Soc. Hematol.* **2001**, *97*, 3951–3959.
- [162] R. L. Gallo, K. J. Kim, M. Bernfield, C. A. Kozak, M. Zanetti, L. Merluzzi, R. Gennaro, *J. Biol. Chem.* **1997**, *272*, 13088–93.
- [163] S. Térmen, M. Tollin, B. Olsson, T. Svenberg, B. Agerberth, G. H. Gudmundsson, *Cell. Mol. Life Sci.* **2003**, *60*, 536–549.
- [164] U. H. N. Dürr, U. S. Sudheendra, A. Ramamoorthy, *Biochim. Biophys. Acta - Biomembr.* **2006**, 1408–1425.
- [165] C. Zhao, T. Nguyen, L. M. Boo, T. Hong, C. Espiritu, D. Orlov, W. Wang, A. Waring, R. I. Lehrer, *Antimicrob. Agents Chemother.* **2001**, *45*, 2695–2702.
- [166] R. Bals, C. Lang, D. J. Weiner, C. Vogelmeier, U. Welsch, J. M. Wilson, *Clin. Diagn. Lab. Immunol.* **2001**, *8*, 370–375.
- [167] J. W. Larrick, J. G. Morgan, I. Palings, M. Hirata, M. H. Yen, *Biochem. Biophys. Res. Commun.* **1991**, *179*, 170–175.
- [168] Y. Sang, M. Teresa Ortega, K. Rune, W. Xiau, G. Zhang, J. L. Soulages, G. H. Lushington, J. Fang, T. D. Williams, F. Blecha, et al., *Dev. Comp. Immunol.* **2007**, *31*, 1278–1296.
- [169] M. Scocchi, D. Bontempo, S. Boscolo, L. Tomasinsig, E. Giulotto, M. Zanetti, *FEBS Lett.* **1999**, *457*, 459–464.
- [170] P. Storici, M. Scocchi, a Tossi, R. Gennaro, M. Zanetti, *FEBS Lett* **1994**, *337*, 303–307.
- [171] A. Tossi, M. Scocchi, M. Zanetti, P. Storici, R. Gennaro, *Eur. J. Biochem.* **1995**, *228*, 941–946.
- [172] G. H. Gudmundsson, K. P. Magnusson, B. P. Chowdhary, M. Johansson, L. Andersson, H. G. Boman, *Genetics* **1995**, *92*, 7085–7089.
- [173] P. Storici, A. Tossi, B. Lenarcic, D. Romeo, *Eur. J. Biochem.* **1996**, *238*, 769–776.
- [174] M. Scocchi, S. Wang, M. Zanetti, *FEBS Lett.* **1997**, *417*, 311–315.
- [175] H. Das, B. Sharma, A. Kumar, *DNA Seq.* **2006**, *17*, 407–414.
- [176] I. G. Fernández de Mera, J. M. Pérez de la Lastra, P. Ayoubi, V. Naranjo, K. M. Kocan, C. Gortazar, J. de la Fuente, *Dev. Comp. Immunol.* **2008**, *32*, 85–91.
- [177] O. Shamova, K. A. Brogden, C. Zhao, T. Nguyen, V. N. Kokryakov, R. I. Lehrer, *Infect. Immun.* **1999**, *67*, 4106–4111.
- [178] K. M. Huttner, M. R. Lambeth, H. R. Burkin, D. J. Burkin, T. E. Broad, *Gene* **1998**, *206*, 85–91.
- [179] M. M. Mahoney, A. Y. Lee, D. J. Brezinski-Caliguri, K. M. Huttner, *FEBS Lett.* **1995**, *377*, 519–522.

- [180] L. Bagella, M. Scocchi, M. Zanetti, *FEBS Lett.* **1995**, *376*, 225–228.
- [181] Y. Xiao, Y. Cai, Y. R. Bommineni, S. C. Fernando, O. Prakash, S. E. Gilliland, G. Zhang, *J. Biol. Chem.* **2005**, *281*, 2858–2867.
- [182] Y. Wang, J. Hong, X. Liu, H. Yang, R. Liu, J. Wu, A. Wang, D. Lin, R. Lai, *PLoS One* **2008**, *3*, e3217.
- [183] T. Uzzell, E. D. Stolzenberg, A. E. Shinnar, M. Zasloff, *Peptides* **2003**, *24*, 1655–1667.
- [184] C.-I. Chang, O. Pleguezuelos, Y.-A. Zhang, J. Zou, C. J. Secombes, *Infect. Immun.* **2005**, *73*, 5053–5064.
- [185] M. Kopitar, A. Ritonja, T. Popovic, D. Gabrijelcic, I. Krizaj, V. Turk, *Biol. Chem. Hoppe. Seyler.* **1989**, *370*, 1145–1152.
- [186] a. Ritonja, M. Kopitar, R. Jerala, V. Turk, *FEBS Lett.* **1989**, *255*, 211–214.
- [187] B. Korkmaz, T. Moreau, F. Gauthier, *Biochimie* **2008**, *90*, 227–242.
- [188] A. E. Shinnar, K. L. Butler, H. J. Park, *Bioorg. Chem.* **2003**, 425–436.
- [189] M. Zanetti, *Curr. Issues Mol. Biol.* **2005**, 179–196.
- [190] R. M. E. Editor, *Host Defense Peptides and Their Potential as Therapeutic Agents*, **2016**.
- [191] M. Z. Renato Gennaro, *Biopolymers* **2000**, *55*, 31–49.
- [192] R. Bals, J. M. Wilson, *Cell. Mol. Life Sci.* **2003**, 711–720.
- [193] U. H. N. Dürr, U. S. Sudheendra, A. Ramamoorthy, *Biochim. Biophys. Acta* **2006**, *1758*, 1408–1425.
- [194] R. Gennaro, B. Skerlavaj, D. Romeo, *Infect. Immun.* **1989**, *57*, 3142–3146.
- [195] G. H. Gudmundsson, B. Agerberth, J. Odeberg, T. Bergman, B. Olsson, R. Salcedo, *Eur. J. Biochem.* **1996**, *238*, 325–332.
- [196] Y. Wang, G. Walter, E. Herting, B. Agerberth, J. Johansson, *Antimicrob. Agents Chemother.* **2004**, *48*, 2097–2100.
- [197] M. E. Selsted, M. J. Novotny, W. L. Morris, Y. Q. Tang, W. Smith, J. S. Cullor, *J. Biol. Chem.* **1992**, *267*, 4292–4295.
- [198] D. Romeo, B. Skerlavaj, M. Bolognesi, R. Gennaro, *J. Biol. Chem.* **1988**, *263*, 9573–9575.
- [199] V. N. Kokryakov, S. S. L. Harwig, E. A. Panyutich, A. A. Shevchenko, G. M. Aleshina, O. V. Shamova, H. A. Korneva, R. I. Lehrer, *FEBS Lett.* **1993**, *327*, 231–236.
- [200] B. Agerberth, H. Gunne, J. Odeberg, P. Kogner, H. G. Boman, G. H. Gudmundsson, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 195–9.
- [201] J. B. Cowland, A. H. Johnsen, N. Borregaard, *FEBS Lett.* **1995**, *368*, 173–176.
- [202] J. W. Larrick, M. Hirata, R. F. Balint, J. Lee, J. Zhong, S. C. Wright, *Infect. Immun.* **1995**, *63*, 1291–1297.

- [203] O. Sørensen, K. Arnljots, J. B. Cowland, D. F. Bainton, N. Borregaard, *Am. Soc. Hematol.* **1997**, *90*, 2796–2803.
- [204] B. Agerberth, J. Charo, J. Werr, B. Olsson, F. Idali, L. Lindbom, R. Kiessling, H. Jö, H. Wigzell, G. H. Gudmundsson, *Am. Soc. Hematol.* **2000**, *96*, 3086–3093.
- [205] A. Di Nardo, A. Vitiello, R. L. Gallo, *J. Immunol.* **2003**, *170*, 2274–2278.
- [206] Y.-H. Yang, G.-G. Zheng, G. Li, B. Zhang, Y.-H. Song, K.-F. Wu, *Leuk. Res.* **2003**, *27*, 947–950.
- [207] M. F. Nilsson, B. Sandstedt, O. Sørensen, G. Weber, N. Borregaard, O. L. E. Sørensen, N. Weber, M. Ståhle-ba, *Infect. Immun.* **1999**, *67*, 2561–2566.
- [208] R. Bals, X. Wang, M. Zasloff, J. M. Wilson, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 9541–9546.
- [209] K. Hase, L. Eckmann, J. D. Leopard, N. Varki, M. F. Kagnoff, *Infect. Immun.* **2002**, *70*, 953–963.
- [210] M. Murakami, T. Ohtake, R. A. Dorschner, B. Schitteck, C. Garbe, R. L. Gallo, *J Invest Dermatol* **2002**, *119*, 1090–1095.
- [211] J. Malm, O. Sørensen, T. Persson, M. Frohm-Nilsson, B. Johansson, A. Bjartell, H. Lilja, M. Ståhle-Bäckdahl, N. Borregaard, A. Egesten, *Infect. Immun.* **2000**, *68*, 4297–4302.
- [212] M. Frohm, H. Gunne, a C. Bergman, B. Agerberth, T. Bergman, a Boman, S. Lidén, H. Jörnvall, H. G. Boman, *Eur. J. Biochem.* **1996**, *237*, 86–92.
- [213] E. Andersson, O. E. Sørensen, B. Frohm, N. Borregaard, a Egesten, J. Malm, *Hum. Reprod.* **2002**, *17*, 2529–2534.
- [214] P. Li-, J. Johansson, G. H. Gudmundsson, M. E. Rottenberg, K. D. Berndt, B. Agerberth, E. Rottenberg, *J. Biol. Chem.* **1998**, *273*, 3718–3724.
- [215] J. Turner, Y. Cho, N.-N. Dinh, A. J. Waring, R. I. Lehrer, *Antimicrob. Agents Chemother.* **1998**, *42*, 2206–2214.
- [216] S. Dosler, E. Karaaslan, *Peptides* **2014**, *62*, 32–37.
- [217] J. Overhage, A. Campisano, M. Bains, E. C. W. Torfs, B. H. A. Rehm, R. E. W. Hancock, *Infect. Immun.* **2008**, *76*, 4176–4182.
- [218] E. Piktel, K. Niemirowicz, U. Wnorowska, M. Wątek, T. Wollny, K. Głuszek, S. Gózdź, I. Levental, R. Bucki, *Arch. Immunol. Ther. Exp. (Warsz)*. **2016**, 33–46.
- [219] D. Xindoli, F. Morgera, U. Zinth, R. Rizzo, S. Pacor, A. Tossi, *Biochem. J.* **2015**, *3*, 443–457.
- [220] Z. Oren, J. C. Lerman, G. H. Gudmundsson, B. Agerberth, Y. Shai, *Biochem. J* **1999**, *341*, 501–513.
- [221] R. Sood, Y. Domanov, M. Pietiäinen, V. P. Kontinen, P. K. J. Kinnunen, *Biochim. Biophys. Acta - Biomembr.* **2008**, *1778*, 983–996.
- [222] C.-C. Lee, Y. Sun, S. Qian, H. W. Huang, *Biophysj* **2011**, *100*, 1688–1696.
- [223] E. T. Verjans, S. Zels, W. Luyten, B. Landuyt, L. Schoofs, *Peptides* **2016**, *85*, 16–26.

- [224] F. K. Rembert Koczulla, Georges von Degenfeld, Christian Kupatt, P. B. Stefan Zahler, Torsten Gloe, Katja Issbrücker, Pia Unterberger, Mohamed Zaiou, Corinna Lebherz, Alexander Karl, Philip Raake, Achim Pfosser, R. L. G. Ulrich Welsch, Pieter S. Hiemstra, Claus Vogelmeier, and R. B. Matthias Clauss, *J. Clin. Invest.* **2003**, *111*, 1665–1672.
- [225] A. Nijnik, R. E. Hancock, *Curr. Opin. Hematol.* **2009**, *16*, 41–47.
- [226] A. Girnita, H. Zheng, A. Groberg, L. Girnita, M. Stale, *Oncogene* **2012**, *31*, DOI 10.1038/onc.2011.239.
- [227] M. Hemshekhar, V. Anaparti, N. Mookherjee, *Pharmaceuticals* **2016**, *9*, DOI 10.3390/ph9030040.
- [228] J. S. Mader, N. Mookherjee, R. E. W. Hancock, R. C. Bleackley, *Mol. Cancer Res.* **2009**, *7*.
- [229] S. X. Ren, A. S. L. Cheng, K. F. To, J. H. M. Tong, M. S. Li, J. Shen, C. C. M. Wong, L. Zhang, R. L. Y. Chan, X. J. Wang, et al., *Cancer Res.* **2012**, *72*, 6512–6523.
- [230] S. X. Ren, J. Shen, A. S. L. Cheng, L. Lu, R. L. Y. Chan, Z. J. Li, X. J. Wang, C. C. M. Wong, L. Zhang, S. S. M. Ng, et al., *PLoS One* **2013**, *8*, e63641.
- [231] S. B. Coffelt, F. C. Marini, K. Watson, K. J. Zvezdaryk, J. L. Dembinski, H. L. Lamarca, S. L. Tomchuck, K. Honer Zu Bentrup, E. S. Danka, S. L. Henkle, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 3806–3811.
- [232] L. Ouyang, Z. Shi, S. Zhao, F. T. Wang, T. T. Zhou, B. Liu, J. K. Bao, *Cell Prolif.* **2012**, *45*, 487–498.
- [233] G. Wang, *Pharmaceuticals* **2014**, *7*, 545–594.
- [234] X. Li, Y. Li, H. Han, D. W. Miller, G. Wang, *J. Am. Chem. Soc.* **2006**, *129*, 5776–5785.
- [235] N. Malanovic, R. Leber, M. Schmuck, M. Kriechbaum, R. A. Cordfunke, J. W. Drijfhout, A. De Breij, P. H. Nibbering, D. Kolb, K. Lohner, *BBA - Biomembr.* **2015**, *1848*, 2437–2447.
- [236] A. A. Strömstedt, M. Pasupuleti, A. Schmidtchen, M. Malmsten, *Antimicrob. Agents Chemother.* **2009**, *53*, 593–602.
- [237] F. Porcelli, R. Verardi, L. Shi, K. A. Henzler-Wildman, A. Ramamoorthy, G. Veglia, *Biochemistry* **2008**, *47*, 5565–5572.
- [238] G. Wang, *Publ. JBC Pap. Press* **2008**, DOI 10.1074/jbc.M805533200.
- [239] G. Wang, M. Elliott, A. L. Cogen, E. L. Ezell, R. L. Gallo, R. E. W. Hancock, *n.d.*, DOI 10.1021/bi2016266.
- [240] Malmsten, *Curr Top Med Chem* **2016**, *16*, 16–24.
- [241] K. A. Henzler-Wildman, G. V Martinez, M. F. Brown, A. Ramamoorthy, *Biochemistry* **2004**, *43*, 8459–8469.
- [242] G. Wang, Y. Li, X. Li, *J. Biol. Chem.* **2004**, *280*, 5803–5811.
- [243] K. A. H. Wildman, D.-K. Lee, A. Ramamoorthy, *Biochemistry* **2003**, *42*, 6545–6558.
- [244] G. Wang, B. Mishra, R. F. Epand, R. M. Epand, *Biochim. Biophys. Acta - Biomembr.* **2014**,

1838, 2160–2172.

- [245] M. Shahmiri, M. Enciso, C. G. Adda, B. J. Smith, M. A. Perugini, A. Mechler, *Nat. Publ. Gr.* **2016**, DOI 10.1038/srep38184.
- [246] A. Barth, C. Zscherp, *Q. Rev. Biophys.* **2002**, 35, S0033583502003815.
- [247] L. K. Tamm, S. a Tatulian, *Q. Rev. Biophys.* **1997**, 30, 365–429.

Danksagung

First, I want to thank my supervisor **Prof. Dr. M. A. Mroginski** for giving me the opportunity of developing my doctoral thesis in her group. Thank you for the trust you have placed and the freedom during these last four years.

Prof. Dr. K. Lohner, vielen Dank für die Begutachtung meiner Arbeit und **Prof. Dr. R. Süßmuth** danke sehr für die Übernahme des Prüfungsvorsitzes.

Prof. Dr. P. Hildebrandt und **Prof. Dr. T. Gutschmann** für die schöne Zusammenarbeit im Cathelicidin Projekt.

Dr. T. Utesch, vielen Dank für deine Hilfe, Ratschläge und sehr geduldige Einführung in die MD. **Dr. M. Horch** vielen Dank für deine Ratschläge. Thanks to **Dr. R. Tapia- Rojo** for the successful collaboration in the development of the MSM of Lichenicidin. Additionally, my deepest thanks to **Dr. G. Pérez-Hernández**, and your aid with the MSMs explanation.

Thanks to my dear AMP family **E. Forbrig**, for the nice cathelicidin and defensin collaboration, and to **B. Gonzalez** because the discussions/ conferences would be less productive without you!

I am deeply grateful to **N. Malanovic** and to **J. Kozuch** for giving me their permissions.

To all members of AG Mroginski for an uncomparable work atmosphere, special thanks to those with whom I shared my doctoral period: **D. Tombolelli**, **A. Takiden**, **Y. Rippers**, **D. Belger** and **N. Herder**. Thanks to all my students: **R. Weimar**, **C. Suazo**, **R. Atton**, **C. Jonas** and **P. B. Sapir**.

To the **HLRN** for the innumerable hours of computing together with the **UniCat** excellence cluster and the **DFG** for the funding.

Thanks to **N. Miguel** for taking care of my mental health and her effective *dm-treatments*. Many thanks to **I. Glizer** and **J. Dragelj** for being always there in times of need. Thanks to my **Señoras** for the ubiquitous support. To all my friends, that have supported me from Spain. Special thanks to **E. Sanjurjo** who has provided me the main injection of happiness, laughs and crazyness. Extensible to **A. Uchima**, **N. Olmedo**, **A. Lozano** who have also contributed to draw a smile in my face at any time despite the distance between us.

To conclude, I want to thank my family. To my grandparents **Miguel**, **Carmen** and my uncles **Teresa** and **Miguel** who have always valued my work. To my uncle **Julio**, who helped me to chase my dreams. To my siblings **Jaime** and **Kiko** because without them my life would not be the same. To my grandma **Alejandrina**, thanks for your love, your strength and how you have filled my heart during all this years. To my parents, **Maria Teresa** and **Francisco** for the great effort that you have made to provide me with an education.

To my husband **Javier**. Thank you very much for always encouraging me to reach my goals as well as for being my support and life partner. To my granddad **Ángel**. The love that you gave me together with grandma, undoubtedly defined my life. Your hard work and your desire for self-improvement have been always present for me. I hope that wherever you are, you smile as much as we would be together.

Statement of Authorship

This doctoral thesis is submitted for the degree of Doctor rerum naturalium (Dr. rer. nat.). I certify that this thesis was written by myself and enclosed my work during the doctoral research period from February 2013 to February 2017 under the supervision of Prof. Dr. M. A. Mroginski at the Physical Chemistry/ Biophysical Chemistry Department of Technical University Berlin.

Collaborations have been detailed and specifically acknowledged. Furthermore, all references used have been quoted. I promise that all my research herein reported have not been extracted for any thesis, research paper or previously presented for another doctoral degree from this or any other university.

Berlin, 24.08.2017

Alejandra de Miguel Catalina

