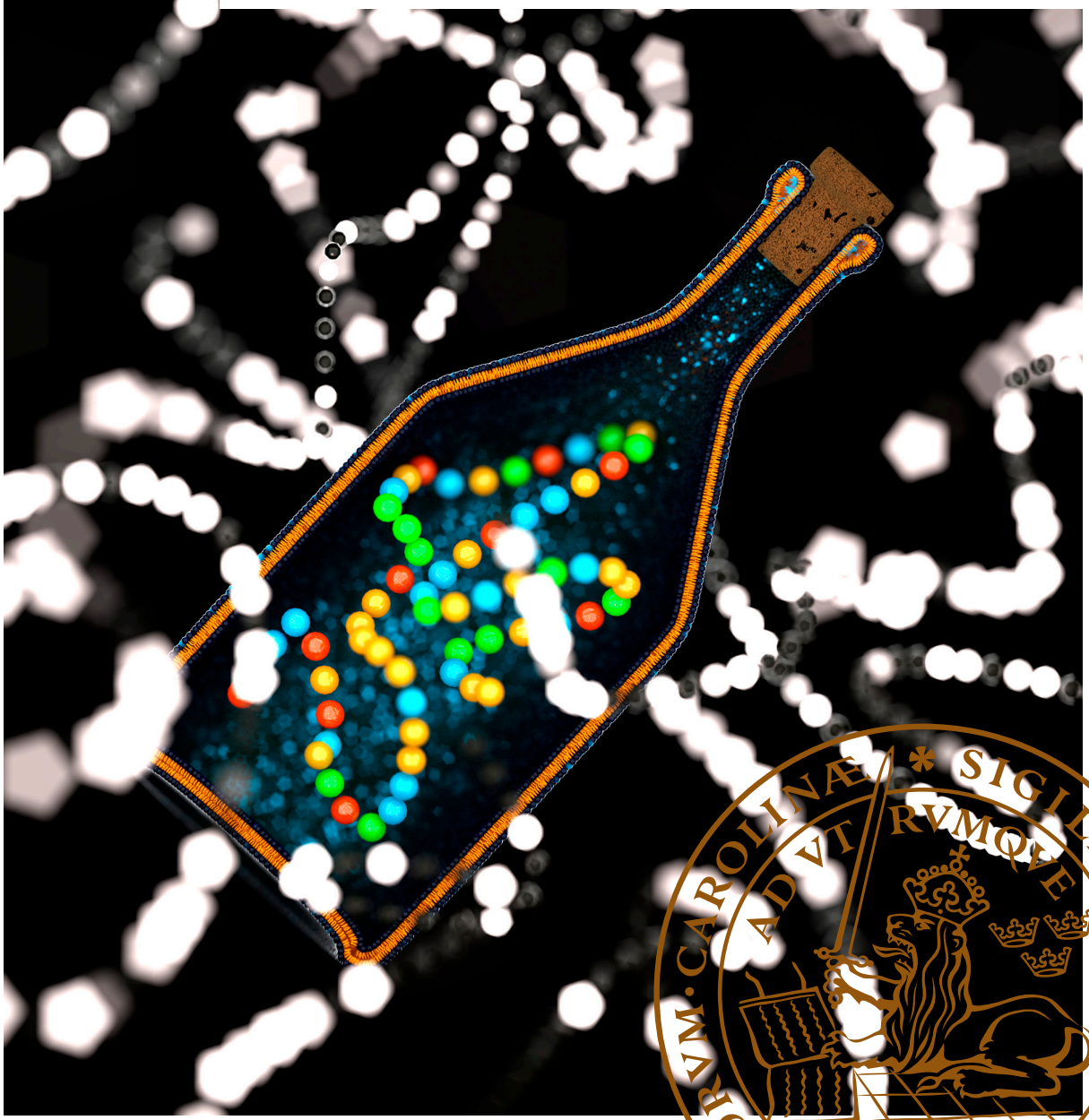
Interactions of intrinsically disordered peptides with phospholipid bilayers

Eriksson Skog, Amanda

2024

Interactions of intrinsically disordered peptides with phospholipid bilayers

AMANDA ERIKSSON SKOG | COMPUTATIONAL CHEMISTRY | LUND UNIVERSITY



Interactions of intrinsically disordered peptides with phospholipid bilayers

by

Amanda Eriksson Skog

LUND
UNIVERSITY

Doctoral Dissertation

by due permission of the Faculty of Science at Lund University.
To be defended on the 6th of December 2024 at 9:00, in lecture hall A at the Centre for Chemistry and Chemical Engineering at Lund University.

Faculty opponent

Dr. Martin Cramer Pedersen

Niels Bohr Institute, University of Copenhagen

Organization	Document name
LUND UNIVERSITY Department of Chemistry Box 124 SE–221 00 LUND Sweden	**DOCTORAL DISSERTATION** Date of disputation 2024-12-06 Sponsoring organization
Author(s) Amanda Eriksson Skog	

Title and subtitle

Interactions of intrinsically disordered peptides with phospholipid bilayers

Abstract

The main goal of this thesis has been to investigate the interaction of the intrinsically disordered peptide Histatin 5 (Hst5), with phospholipid bilayers, using a mixture of experimental and computational techniques, such as, small-angle X-ray scattering, circular dichroism, neutron reflectometry, quartz-crystal microbalance with dissipation monitoring, atomistic molecular dynamics simulations, and coarse-grained Monte Carlo simulations. Hst5 is of particular interest due to its known antimicrobial effects, where it acts as the first defense against fungal infections in the mouth. With antimicrobial resistance being an increasingly and serious threat against world health, alternative treatments to common antimicrobial agents are needed, where antimicrobial peptides being one option. It is therefore important to understand the mechanism behind their effect, to be able to utilize their properties correctly in pharmaceuticals. In the first study of this thesis, Hst5 was found to spontaneously translocate a phospholipid bilayer, without affecting the structural integrity of the bilayer, thus resulting in a peptide cushion between the supported bilayer and the solid surface. This formation is in line with the antimicrobial effect of Hst5, and has therefore been used as an indication of antimicrobial effect throughout the work conducted in this thesis. The experimental conditions needed for the cushion formation was further determined. In the second study, the role of the amino acid histidine, which is frequently found in Hst5, was investigated to determine its role in the translocation process. The results showed that the penetration depth into the bilayer increases with increasing number of histidines. In this study, it was also suggested that not only the number of histidines, but also their position were important, therefore, a study regarding the order of amino acids was conducted as the ongoing study presented in this thesis. The results indicate a difference in interaction, dependent on the sequence order, however, the underlying explanation is not yet understood. Furthermore, the sequence length of Hst5 was investigated in the third study of this thesis, which only showed small differences at 10 mM NaCl concentration, while in 150 mM NaCl, both the shorter and longer variant display interactions with the bilayer, while Hst5 does not. The final study included in this thesis concerns a different peptide called KEIF, which is the intrinsically disordered N-terminal of magnesium transporter A found in *Escherichia coli*. This study was conducted with the aim to investigate the hypothesis that surface active intrinsically disordered peptides gain structure upon adsorption, which is related to their function. The peptide became more structured upon adsorption, supporting the presented hypothesis.

Key words

intrinsically disordered peptides, antimicrobial peptides, peptide-lipid interaction, NR, QCM-D, CD, MC, MD, SAXS

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language
	English
ISSN and key title	ISBN
	978-91-8096-070-0 (print) 978-91-8096-071-7 (pdf)
Recipient's notes	Number of pages: 241 / Price
	Security classification

Interactions of intrinsically disordered peptides with phospholipid bilayers

by

Amanda Eriksson Skog

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover illustration front: Illustration by Lucas Nilsson-Villoresi

Funding information: The thesis work was financially supported by Nordforsk, the Crafoord Foundation, Vinnova, the Royal Physiographic Society of Lund, the Hanseatic League of Science (HALOS) Cross Border Research, and the Magnus Bergvall foundation. This thesis has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806 (MOSBRI).

Faculty of Science at Lund University, Department of Chemistry

ISBN: 978-91-8096-070-0 (print)
ISBN: 978-91-8096-071-7 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2024

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

A little nervous breakdown can really work wonders for a girl
- Rory Gilmore

Contents

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Spontaneous formation of cushioned model membranes promoted by an intrinsically disordered protein**

Y. Gerelli, **A. E. Skog**, S. Jephthah, R. J.L. Welbourn, A. Klechikov, M. Skepö
Langmuir, 2020, 36(15), pp. 3997-4004

II **Interaction of a histidine-rich antimicrobial saliva peptide with model cell membranes: The role of histidines**

A. E. Skog, G. Corucci, M. D. Tully, G. Fragneto, Y. Gerelli, M. Skepö
Langmuir, 2023, 39(22), pp. 7694-7706

III **Translocation of Antimicrobial Peptides across Model Membranes: The Role of Peptide Chain Length**

A. E. Skog, N. Paracini, Y. Gerelli, M. Skepö
Mol. Pharmaceutics, 2024, 21(8), pp. 4082–4097

IV **Assessing the interaction between the N-terminal region of the membrane protein magnesium transporter A and a lipid bilayer**

A. E. Skog, N. C. Jones, L. K. Månsson, J. P. Morth, S. Vrønning Hoffmann Y. Gerelli, M. Skepö
Manuscript submitted to J. Colloid Interface Sci.

Author contributions

Paper I: Spontaneous formation of cushioned model membranes promoted by an intrinsically disordered protein

I performed all QCM-D experiments and analysis, and participated in most NR experiments, participated in discussion and contributed to the writing and revision of the paper.

Paper II: Interaction of a histidine-rich antimicrobial saliva peptide with model cell membranes: The role of histidines

I planned the study together with my supervisor, prepared all the samples, and performed the experiments. I analyzed the data with input from the co-authors, and wrote the manuscript with support from the co-authors.

Paper III: Translocation of Antimicrobial Peptides across Model Membranes: The Role of Peptide Chain Length

I planned the study together with my supervisor, prepared all the samples, and performed the experiments and the simulations. In the simulations, I extended the adsorption analysis to include two surfaces. I analyzed the data with input from the co-authors, and wrote the manuscript with support from the co-authors. I was responsible for the submission and revision process.

Paper IV: Assessing the interaction between the N-terminal region of the membrane protein magnesium transporter A and a lipid bilayer

I planned the study together with my supervisor, prepared all the samples, and performed most of the experiments and all the simulations. I analyzed the data with input from the co-authors, I wrote the first draft of the manuscript independently, and with support from the co-authors finalized the manuscript I was responsible for the submission process.

List of abbreviations

IDP	intrinsically disordered protein
IDR	intrinsically disordered region
AMP	antimicrobial peptide
Hst5	Histatin 5
MgtA	magnesium transporter A
KEIF	N-terminal part of magnesium transporter A
LUV	large unilamellar vesicles
SUV	small unilamellar vesicles
POPC	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
POPS	1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine
LLC	lamellar liquid crystalline
CD	circular dichroism
SRCD	synchrotron radiation circular dichroism
OCD	oriented circular dichroism
SAXS	small-angle x-ray scattering
QCM-D	quartz-crystal microbalance with dissipation monitoring
NR	neutron reflectometry
SLD	scattering length density
VFP	volume fraction profile
MC	Monte Carlo
MD	molecular dynamics
PBC	periodic boundary conditions
R_g	radius of gyration

Acknowledgements

My PhD journey started back in Mach 2020, and as we all remember, so did the Covid-19 pandemic. A lot (if not most) of the plans for my PhD work got thrown out the window. Booked measurements at large scale facilities abroad, interacting irl with fellow PhD students, and going to conferences were simply not feasible. Luckily, I have a lot of amazing people around me who helped me through both the pandemic, but more importantly, the rest of my PhD studies.

I want to start of by expressing my gratitude to *Marie* for taking me in as a summer student back in 2017. Allowing me to continue with a bachelor degree thesis, masters degree thesis, and finally accepting me as a PhD student. It has truly been an honor working with you, and I am so greatful for everything you have done for me, as well as all the opportunities you have given me. Thank you.

Secondly, I am very grateful to have been supervised by, and working with *Yuri*. Your experimental expertise, the time you spent teaching me about the measurements and analysis of the data has really been invaluable. Thank you for putting up with all my questions and answering them over and over again.

To all collaborators and co-authors in various projects, *Nico, Nyk, Søren, Giaccomo, Giovanna, Mark*, it has been a pleasure collaborating with you, and I am grateful for all the scientific discussions we have had.

I am very grateful to all the people who helped proof-reading this document, your comments, and the discussions following them, have been invaluable! Special thanks to *Mona*, your comments and our discussions was very helpful!

The past and present members of the *Skepö group*, thank you for all the fikas, and helping each other out with various tasks! I am also grateful for all the people at the *Division of Computational Chemistry* and *Division of Physical Chemistry* I had the privilege of working alongside with, and also have an occasional cup of coffee with. Special thanks to *Maria* and *Chris* for helping me with everything lab related, without you, the experiments would not run as smooth! Also thanks to *Maria* for helping with all administrative tasks!

Thank you *Sophie* for trusting me with making changes in the KEMA20 lab course, and supporting my ideas. I learned a lot from teaching this course! Thank you for always keeping you door open, it has been invaluable!

Without all the amazing people around me this work would have driven me mad, so thank you *Rebecka* and *Lucas* for being my family in Lund, always having a cup of tea ready for me. *Julia* and *Johan* for always keeping you door open, and for always having

me when I need to escape the city! *Linda*, for being the colleague I needed when I started this journey, which turned into a great friendship. To all my *friends* and *family*, my biggest supporters, thank you for always cheering on me.

Slutligen, tack Wilmer för att du varit en klippa när allt stormat, för att du står ut med mig alla dagar. Att komma hem till din famn varje kväll har varit ovärderligt.

After showing gratitude to all these amazing people around me, I want to show some gratitude towards myself, because without me all this work would not have been done the way it has. So thanks me for doing all of this, I did great.

Populärvetenskaplig sammanfattning på svenska

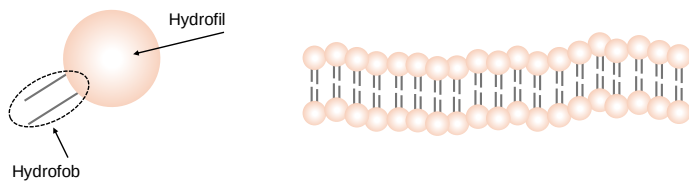
Sedan 1928 när Alexander Fleming upptäckte penicillin i sitt laboratorium har det använts flitigt för att bota diverse sjukdomar orsakade av bakterier. Bakterier är en mikroorganism, eller mikrob som det också kallas, vilket innebär att de har en storlek på mikro-skalan, och är så små att vi inte kan se dem med blotta ögat. Det finns också andra typer av mikrober, till exempel virus och svamp. Alla dessa olika typer av mikrober kan (men måste inte) leda till sjukdom, vilket troligen har påverkat människor genom hela historien. Över det senaste århundradet har det därför lagts mycket kraft på att utveckla mediciner, likt penicillin, för att bota sjukdomar orsakade av mikrober. Exempelvis kallas ämnen verksamma mot bakterier för antibiotika, vilket penicillin är en typ av. Är de verksamma mot virus kallas de antivirala ämnen. Slutligen refererar man ofta till medel mot svamp helt enkelt som svampdödande medel. Som en försvarsmekanism mot dessa medel har mikroberna utvecklat resistens, där de på olika sätt ändrar sin uppbyggnad som gör att medicinerna inte längre är lika effektiva, eller alls verksamma.

Candida albicans är en sorts svamp som finns naturligt hos oss människor, bland annat på huden och i munnen, utan att orsaka sjukdom i normaltillstånd. Om den får grogrund kan den dock föröka sig och på så vis bli patogen, alltså att den orsakar ett sjukdomstillstånd. När detta händer, exempevis i munnen eller på huden, är det ofarligt och kan behandlas lokalt ganska enkelt. Det verkliga problemet uppstår när svampen tar sig in i till exempel blodomloppet eller i det centrala nervsystemet, och blir en så kallad invasiv candidainfektion. Situationen blir då kritisk och denna typ av svampinfektion leder i 20% till 50% av fallen till döden, trots behandling med svampdödande medel. Det är därför enkelt att förstå att resistens vid en invasiv infektion är förödande.

En alternativ behandlingsmetod till klassiska antimikrobiella medel är så kallade antimikrobiella peptider. Peptider, eller proteiner som de ofta kallas, är en livsviktig komponent i våra kroppar där de är byggstenar i alla kroppens vävnader, transportörer av livsnödvändiga ämnen, och en del av vårt immunförsvar. Proteiners bestämda struktur har länge ansetts vara ett krav för att proteinet ska ha en funktion. För att utmana den uppfattningen, har det sedan 1990-talet praktiskt taget exploderat i studier kring proteiner som saknar en väldefinerad struktur, så kallade oordnade proteiner. Dessa proteiner har visat sig vara delaktiga i flera biologiska processer och därmed ha en funktion, trots bristen av struktur. Antimikrobiella peptider är naturligt förekommande i alla biologiska riken, och är en naturlig del av det mänskliga immunförsvaret. Eftersom dessa peptider utvecklats tillsammans med mikrober under lång tid utan att påvisa samma problematik med resistensutveckling, kan de vara en del av lösningen mot den ökade resistensen. De kan användas antingen själva, eller i kombination med klassiska antimikrobiella medel.

I den här avhandlingen har små, oordnade peptider studerats, där det huvudsakliga fokuset har varit på en antimikrobiell peptid som finns i saliv, *Histatin 5*, som är en del av kroppens första försvar mot svampinfektion orsakad av *Candida albicans*. Det har gjorts flera studier där effekten av peptiden studerats, men mekanismen bakom hur den dödar svampen är fortfarande inte helt känd. Det som är känt, är att den på något vis tar sig in i svampcellen och attackerar mitokondrien (cellens kraftverk). Målet med mina studier har därför varit att, utifrån att studera olika typer av växelverkan, förstå hur peptiden tar sig in i cellen.

Cellmembranet (skalet) runt celler, är till stor del uppbyggt av fosfolipider, men innehåller även andra beståndsdelar som proteiner och kolesterol. Fosfolipider är molekyler som består av två delar, där en del älskar vatten (hydrofil), kallad huvudet och en del hatar vatten (hydrofob), kallad svansarna. En bra minnesregel inom kemi är att *lika löser lika*, det innebär att hydrofila molekyler löser sig bra i vatten, medan hydrofoba molekyler löser sig dåligt i vatten, men löser sig istället bättre i olja. Därför organiserar sig lipidmolekylerna med svansarna mot varandra i en vattenlösning för att skydda dem mot vattnet. Ett exempel för hur det kan se ut demonstreras i Figur 1 där lipiderna organiserats till ett plant bilager. Detta är en viktig del i studiens frågeställning; Hur kan en hydrofil peptid som Histatin 5 kan ta sig igenom ett membran med ett lager av hydrofoba svansar? För att förstå det har jag studerat både växelverkan mellan Histatin 5 och ett bilager bestående av enbart fosfolipider. Detta bilager är menat att vara ett modellsystem för ett verkligt cellmembran. Olika varianter av Histatin 5, där olika egenskaper hos peptiden har ändrats, har också undersökts för att kunna utröna hur de olika egenskaperna hos peptiden bidrar till möjligheten att korsa bilagret.



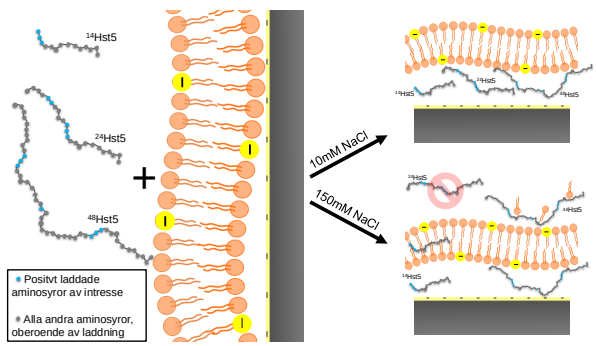
Figur 1: En förenklad skiss av hur en lipidmolekyl ser ut, samt ett exempel på hur lipidmolekyler kan organisera sig i vattenlösning.

Växelverkan mellan peptid och membran har gjorts med flera olika metoder, både experimentella och beräkningsbaserade, för att kunna besvara olika typer av frågor. Bland annat har strukturen av alla Histatin 5 varianter studerats med två olika metoder, *SAXS*[1] och *CD*[2]. SAXS ger information om flexibilitet och storlek, medan CD ger information om hur peptiden veckar sig. Vidare har själva växelverkan mellan peptiden och

[1] småvinkelspridning, eller small-angle X-ray scattering på engelska
[2] cirkulär dikroism

bilagret gjorts med två ytkänsliga tekniker, *QCM-D*[3] och *NR*[4]. QCM-D har svarat på frågor som hur mycket av peptiden som adsorberar (fastnar) på ytan, hur lång tid det tar för peptiden att adsorbera, och om peptiden adsorberar som ett fast lager, eller om det kan deformeras. NR har främst använts för att ta reda på var i systemet olika komponenter befinner sig, vilket är grunden till att kunna svara på forskningsfrågan. I datorsimuleringar har systemet representerats med olika mycket detaljer, antingen är alla atomer i systemet representerade som en egen partikel, eller så klumpas flera atomer med liknande egenskaper ihop, och representeras med en partikel. Hur mycket detaljer man väljer att ha i sin beräkning beror på den specifika frågeställningen och kan skilja sig åt mellan mindre frågor inom samma projekt. Det är därför möjligt att kombinera resultat från beräkningar med olika mycket detaljer för att besvara den ursprungliga frågeställningen.

Det vi har visat i arbetet, presenterat i den här avhandlingen, är att när Histatin 5 introduceras för ett negativt laddat bilager, som deponerats på en negativt laddad yta, kan peptiden spontant transporteras genom bilagret (trots att det är hydrofobt!) och ackumuleras sedan mellan den negativt laddade ytan och bilagret. Detta har vi använt som modellsystem för att Histatin 5 tar sig in i en cell och dödar den. I efterföljande studier har vi undersökt specifika egenskaper hos peptiden, som visat bland annat att aminosyran histidin är viktig för peptidens möjlighet att ta sig igenom bilagret, även har längden betydelse för växelverkan, men där visade det sig att en längre och en kortare variant av Histatin 5 faktiskt interagerade med bilagret när Histatin 5 inte gjorde det. Ett exempel på det senare visas i Figur 2. Sammanfattningsvis; vi har hittat egenskaper hos Histatin 5 som är viktiga för dess förmåga att transporteras över bilagret, men vidare studier behövs för att hitta ytterliga viktiga egenskaper hos peptiden. En kunskap som kan användas till att tillverka effektivare läkemedel.



Figur 2: En skiss av hur interaktionen mellan Histatin 5 och två längdvarianter av den med ett negativt laddat bilager ser ut. Undersökningen gjordes vid låg (10 mM) och hög (150 mM) salthalt.

[3]kvartsmikrovåg med dissipationsmätning, eller quartz crystal microbalance with dissipation monitoring på engelska

[4]neutronreflektometri

Part A

Introductory chapters

Trying is the first step towards failure.
- Homer Simpson

1 | Introduction

Contents

This chapter aims to give a brief introduction on the topic of antimicrobial resistance. This will include a definition of what a microbial is, as well as some brief history about the fight against diseases. The aim of the thesis is presented. Lastly, the outline of the thesis is given.

1.1 Microbes, diseases, and antimicrobial resistance

1.1.1 Microbes and diseases

The general definition of a microbe, or microorganism, is an organism of microscopic size, that is, it cannot be seen with the naked eye, rather a microscope is needed. Most microbes belong to one of the four major groups, namely, bacteria, viruses, fungi, or protozoa[1]. Species from these different groups may cause disease in humans (almost all viruses and protozoa do). However, many bacteria and fungi are naturally occurring at different places in our body helping it function properly. Despite most bacteria and fungi being harmless, diseases caused by microbes have (probably) always troubled humans.[2] It was therefore a great medical breakthrough when Alexander Fleming in 1928 discovered penicillin,[3] a fungi able to kill certain types of bacteria. It was first in 1941 Abraham and co-workers purified enough penicillin to investigate the clinical effectiveness of it.[4] Penicillin is a type of antibiotics, that is, a compound inhibiting the growth of, or killing bacteria. Over the last century, the development of antibiotics has increased the average human life expectancy with over 20 years.[5] In addition to antibiotics, there are compounds to treat infections caused by the other types of microbes as well, namely antifungal, antiviral and antiprotozoal agents. According to the World Health Organization (WHO), these agents are the foundation of modern medicine,[6] meaning that occurrences compromising the effectivity of these agents are paramount for world health.

1.1.2 Antimicrobial resistance

Antimicrobial resistance is when microbes no longer respond to treatment using antimicrobial agents. It is a natural process occurring over time by changes in genetics of microbes. The process is however accelerated, mainly by incorrect and excessive use of these agents. This means that diseases, which could previously be treated with medicines, are now difficult, or even impossible to treat. In his Nobel Prize speech in 1945, Fleming warned about bacterial resistance against these compounds, and that the spread of resistance causes the antibiotics to be less effective, or even ineffective.[7] In 2019, almost 5 million deaths were associated with bacterial resistance against antibiotics.[8] In recent years, the growing problem of antibiotic resistance has received considerable attention, however, a growing, significantly less noticed case, is antifungal resistance.[9]

[1] a large and diverse group of single-cell,eukaryotic microorganisms, which belong to the kingdom Protista.[1]

Fungal infections are a general increasing health concern, however, individuals who are immunocompromised are at higher risk of getting severe fungal infections, also called invasive fungal disease. The rapid emergence of antifungal resistance increases the global health threat posed by these infections. There are only a few classes of antifungal agents currently in clinical use to treat fungal infections affecting internal organs, and although they are effective, they are known to give rise to an abundance of unwanted side effects.[9] These increased levels of resistance is an enormous threat on global world health, and it is therefore paramount to find alternative treatments.

1.1.3 Candida albicans

Candida albicans is a part of the healthy human microbiome where it is found in the mouth, throat, gut, vagina, and on the skin. However, if multiplied in these areas, it becomes pathogenic. It then produces diseases such as cutaneous candidiasis, affecting the skin, or oropharyngeal candidiasis, affecting the mouth and throat.[9] Infections caused by *Candida* are the most common fungal infections in both the mouth and on skin, where the majority of the infections are caused by the species *Candida albicans*.[10] In addition to the more harmless infections, it can also produce invasive candidiasis, affecting for example the blood and central nervous system, just to mention two. The overall mortality of invasive candidiasis ranges between 20% to 50%, despite active antifungal treatment. With the risk of rising resistance, the situation is concerning.[9]

1.2 Aim of the thesis

With the problem outlined above, it is clear that alternatives are needed to treat and prevent infectious diseases. Antimicrobial peptides (AMPs) have received considerable attention as an alternative to traditional antibiotics as they are a natural part of our innate immune system. An introduction of this is given in Chapter 2. In this thesis, a peptide, previously known to be antimicrobial, called Histatin 5 (Hst5),[11–14] is investigated from a physiochemical point of view. The peptide's interaction with model membranes have been studied with the aim to better understand the mechanism behind the interaction. Histatin 5 has some important characteristics, discussed more in detail in Chapter 2, and the hypothesis is that these characteristics are important for its mechanism to kill microbes. The problem was also approached from another perspective, where instead the interaction of a part of a bacterial protein, the N-terminal of magnesium transporter A, referred to as KEIF,[15] with a model membrane was investigated. This peptide was also used to investigate if surface active disordered peptides gain structure upon adsorption to surfaces, and thereby also their function, which is

the hypothesis.

1.3 Outline of the thesis

The work conducted over the past almost five years is summarized in this thesis, and focuses on understanding the interaction between disordered peptides with lipid bilayers by altering the sequence of the peptide.

In **Chapter 2**, an introduction to peptides and lipids, the main characters of this thesis, is given.
In **Chapter 3**, a brief introduction to statistical thermodynamics, as well as intermolecular interactions are given.
In **Chapter 4**, the experimental techniques are introduced.
In **Chapter 5**, the simulation methods and models are presented.
In **Chapter 6**, a summary of the main results of the research papers, and one ongoing study, are presented.
In **Chapter 7**, the concluding remarks and future perspectives are presented.

Enjoy!

If mathematical analysis should ever hold a prominent place in chemistry - an aberration which is happily almost impossible - it would occasion a rapid and widespread degeneration of that science.

- Aguste Comte

2 | Biological macromolecules

Contents

There are four major classes of biological macromolecules; nucleic acids, carbohydrates, lipids, and proteins. In this thesis, the two latter are investigated. This chapter introduces proteins, and the subcategories; intrinsically disordered proteins, and antimicrobial peptides, as well as the two peptides studied in this work; Histatin 5 and KEIF. This is followed by an introduction and description of lipids, with the focus on phospholipids.

2.1 Proteins and peptides

Proteins are biological macromolecules with important functions in almost all biological processes. They are built up by a linear chain of *amino acids*, linked by amide bonds. In this thesis, a protein will be defined as a chain of amino acids longer than 50 residues, anything shorter will be referred to as a peptide. An amino acid consists of an α-carbon, which is the central carbon atom. Linked to this atom is an amino group ($-NH_2$), a carboxylic acid group ($-COOH$), a hydrogen atom (H), as well as a side group (-R), which is distinctive of each amino acid. At a neutral pH (pH 7), which also roughly corresponds to physiological pH, the amino group is protonated ($-NH_3^+$), and the carboxyl group is deprotonated ($-COO^-$), making the amino acid *zwitterionic*. There are 20 different amino acids that make up the majority of peptides and proteins. These amino acids can be grouped into, positively charged, negatively charged, polar, and hydrophobic, depending on the side chain.[16] There are four different levels describing protein structure:

Primary structure, describes the order in which the amino acids are linked, that is, the sequence. This sequence is unique for each protein.

Secondary structure, occurs due to freedom in rotation over bonds within each amino acid, allowing the protein to fold in many different ways. This folding allows for regular patterns of hydrogen bonds, giving rise to local structures such as α-helices and β-strands (see Figure 2.1).

Tertiary structure is further folding of the local structures into a well-defined three-dimensional shape. This folding is mainly driven by the hydrophobic interaction for a protein in aqueous solution, where the hydrophobic residues are hidden from the surrounding water.[16]

Quaternary structure, describes the fourth level of structure of a protein with multiple sub-units. It describes the arrangement and interaction between these sub-units, each having a three-dimensional structure.[16]

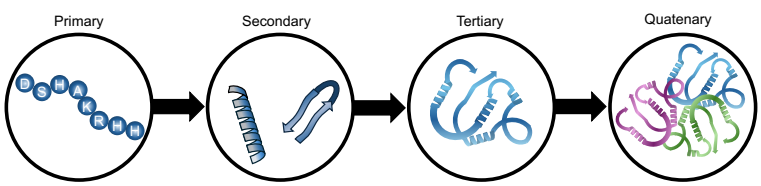


Figure 2.1: Illustration of the different levels of protein structure.

There is a traditional view that the function of a protein is critically dependent on a

well-defined, folded three-dimensional structure, referred to as the structure-function paradigm.[17,18] This view has been challenged by the field of *intrinsically disordered proteins* (IDPs), which has been rapidly evolving since the late 1990's.[19]

2.1.1 Intrinsically disordered proteins

IDPs and *intrinsically disordered regions* (IDRs) refers to whole proteins or segments that do not self-fold into fixed 3D structures, with such disorder sometimes existing in the native state at physiological conditions. This means that these proteins and regions are more flexible compared to other proteins, and they can interchange rapidly between several different conformations. The disorder can often be recognized already in the primary sequence, as these proteins and regions are generally enriched in charged polar amino acids, and scarce in bulky, hydrophobic residues.[20] They also range in size and net charge, however, they have in common a high value of net charge and a low hydrophobicity.[21,22] These proteins are involved in many vital reactions and pathways, and there are also several IDPs related to diseases, e.g., Parkinson's disease, Alzheimer's disease, and different types of cancer.[18]

2.1.2 Antimicrobial peptides

With the emerging antimicrobial resistance (described in the introduction), alternative treatments to classical antimicrobial agents are needed. For this, *antimicrobial peptides* are of interest as potential agents.[23–25] These peptides can be found in all kingdoms of life: eukaryotes, bacteria, and archaea, and they are a natural part of the human innate immune system.[23,24] AMPs are characterized by a short sequence, between 10 to 60 residues, and the majority of them are cationic. Since the natural AMPs are very diverse, they can be classified according to several different characteristics: (i) the source, that is, where they come from, e.g., mammalians or plants, (ii) their activity, meaning what they are active against, e.g., antibacterial or antifungal, (iii) their structure, e.g., α-helical or β-sheet, and (iv) which amino acids are present in abundance in the sequence, e.g., arginine or histidine.[25] If the peptides are classified according to (iii), they can be further divided into four major groups, namely α-helical, β-structures, α- and β-structures, as well as extended structure.[23,24] As this thesis focuses on natively disordered peptides, naturally the focus has been shifted towards AMPs belonging to the group of extended structure. This group are usually rich in the amino acids glycine, arginine, or histidine, and lack secondary structure. The AMP Hst5 belongs to this group,[23] which will be extensively discussed throughout this thesis.

The antimicrobial mechanism varies widely with the different AMPs, but are known

to either target the outer membrane of the microbe, or to target intracellular processes. The exact mechanism is not determined. However, for membrane targeting mechanism, there are three suggested models: (i) toroidal pore model, displayed in Figure 2.2C, the AMPs are first vertically embedded in the cell membrane, followed by accumulation of AMPs which bends the membrane to form a hole. (ii) Barrel-stave model, where the AMPs first accumulate, and then insert into the membrane, forming a channel, displayed in Figure 2.2C. In the last model, referred to as (iii) the carpet model, the AMPs adsorb to the top of the membrane and destroy it with a detergent-like effect, where the membrane is disintegrated, which is displayed in Figure 2.2A. AMPs could also traverse the cell membrane without causing extensive damage to the membrane, and instead target cellular processes. The target varies between different AMPs and varies from affecting the transcription and translation of proteins to inhibiting the DNA replication.[24,25]

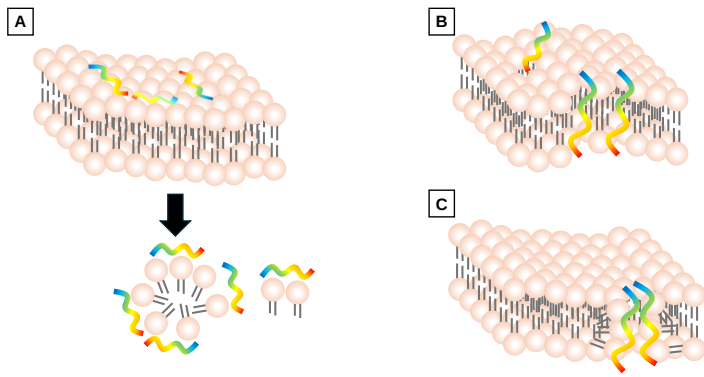


Figure 2.2: Schematic illustration of the different modes membrane targeting mechanisms of AMPs. **(A)** The carpet model, **(B)** the barrel stave model, and **(C)** the toroidal pore model.

2.1.3 Histatin 5

The histidine family, found in saliva, is a group of cationic peptides, rich in the amino acid histidine has shown to be both antibacterial and antifungal. The group consists of 12 members out of which Hst5 is the most potent antifungal agent against *Candida albicans*.[11–14] Hst5 is a 24 amino acid long peptide,[26] with a net cationic charge of +5 in physiological pH, and seven out of the amino acids are histidines, constituting 29% of the sequence. This amino acid is known to charge regulate, that is, change its charge in the vicinity of other charged molecular species, such as a lipid bilayer.[27] The high content of this amino acid enables for charge regulating of Hst5, which has been proposed as a key feature in the antimicrobial mechanism of the peptide.[28] Hst5 is intrinsically disordered in aqueous solution,[26] but adopts an α-helical structure in

2,2,2-trifluoroethanol (TFE).[13,29,30]

The candicidal activity, i.e. its ability to inhibit or kill Candida fungi, of histatins has been shown to be inversely proportional to both the concentration of divalent ions and the ionic strength of the medium,[13] and it has also been shown that they are unable to disrupt lipid membranes.[14,31] Studies have shown that Hst5 instead translocates the cell membrane[32] and accumulates in the mitochondria, an action dependent on a mitochondrial transmembrane potential.[33,34] Active transporters over the membrane have been suggested,[14,35] however, studies have shown Hst5 to be able to translocate phospholipid liposomes without active transport.[36,37] It has been suggested that this direct translocation across the cell membrane leads to cell membrane disruption, and that this is the initial fungicidal action of Hst5.[34,37] This process is determined to be dependent on Hst5 concentration.[37]

Variants of Hst5

To better understand the translocation of Hst5 across a lipid bilayer with respect to the primary sequence of the peptide, several variants have been designed by us. The sequences are presented in Table 2.1, together with the sequence of Hst5. Four variants were designed with a varying number of histidine residues. These variants have previously been investigated with respect to their zinc-binding abilities.[38] Hst5 possesses several metal binding motifs for both copper, nickel[39] and zinc (HEKHH) and (HAKRHH), where the latter was established through investigation of these four variants.[38,40]

Table 2.1: The amino acid sequences of the peptides included in this thesis. Positively charged residues are presented in blue, negatively charged ones in red, and histidine residues are presented in green.

Peptide	Sequence
Hst5 and variants	
Hst5	DSHAKRHHGYKRKFHEKHHSHRGY
$Hst5_0$	DSQAKRQQGYKRKFQEKQQSQRGY
$Hst5_2$	DSQAKRQHGYKRKFQEKQHSQRGY
$Hst5_3$	DSQAKRQQGYKRKFHEKHHSQRGY
$Hst5_4$	DSHAKRHHGYKRKFQEKQQSHRGY
^{14}Hst5	KRKFHEKHHSHRGY
^{48}Hst5	DSHAKRHHGYKRKFHEKHHSHRGYDSHAKRHHGYKRKFHEKHHSHRGY
$Hst5^{rand1}$	HHYARKSDKHSFHYGRHKHERGKH
$Hst5^{rand2}$	HGHSYKKAGYKEHSKHHHHDRRRF
N-terminal of MgtA	
KEIF	MFKEIFTRLIRHLPSRLVHRDPLPGAQQTVNTV

The importance of sequence length was also of interest to investigate, and for this, two additional peptides were studied, namely the tandem-repeat of Hst5, referred to as ^{48}Hst5, as well as the last 14 amino acids of Hst5, referred to as ^{14}Hst5. ^{48}Hst5 was previous studied by Fagerberg *et al.*,[41] where the peptide was compared to Hst5 both in dilute concentrations, as well as in more concentrated systems where the aggregation behavior was compared. The shorter variant, ^{14}Hst5, has been extensively studied before,[26,42–46] and determined to maintain almost the same candicidal effect as Hst5, which is the reason why we selected this part of the peptide to investigate.

The last set of Hst5 variants was designed to investigate the importance of the position of the amino acids, hence, two sequences with the same amino acid content as Hst5 was designed, however, the sequence was randomized, presented as Hst5^{rand1} and Hst5^{rand2} in Table 2.1. Hst5^{rand1} was previously investigated by Cragnell *et al.*,[38] with the purpose to investigate how the random sequence affected the peptides zinc-binding properties. Hst5^{rand2} was investigated to compare its re-entrant condensation behavior compared to Hst5.[47]

2.1.4 Magnesium ion transport in bacteria

In prokaryotes, three classes of magnesium ion transporters have been identified as responsible for translocation of Mg^{2+} across the cell membrane: the *CorA* magnesium transporter, *magnesium transporter E* (MgtE), and *magnesium transporter A* and *B* (MgtA and MgtB).[48] The two latter, MgtA and MgtB are P-type ATPases mediating the *influx* of Mg^{2+} down the electrochemical gradient, as opposed to the more common function of ATPases that use ATP to mediate *efflux* of cations against an electrochemical gradient.[48,49]

MgtA in *Escherichia coli* (*E. coli*)[50] consists of 898 amino acids, with a molecular weight of 99.466 kDa. The N-terminal region of this protein was predicted to be intrinsically disordered in a study by Subramani *et al.*.[15] The same study found that this part of the protein was not required for activation by *cardiolipin*, which is essential for MgtA activation *in vitro*. It was also determined not to be involved in membrane trafficking, which sparked the interest to investigate the function of this N-terminal IDR in MgtA.

The N-terminal IDR of *E. coli* MgtA constitutes the first 33 amino acids of the protein, and in this work, it is referred to as *KEIF*, which is the first redable part in the amino acid sequence, as underlined in Table 2.1. An extensive study to characterize the physiochemical properties of KEIF was performed by Jephthah *et al.*,[51] which confirmed the suggested intrinsic disorder of the peptide. Interactions of KEIF with lipid

vesicles displayed an increase in ordered structure upon adsorption. In a more recent study by Koder Hamid *et al.*[52] a similar behavior of KEIF was observed upon adsorption to the clay Laponite®. In this study, the arginine residue in position 16 was singled out as an important residue in adsorption.

To fight microbial diseases, it can be advantageous to not only understand the antimicrobial agents, but also the pathogenic microbe. The hope is therefore that an understanding of the biological function of KEIF in MgtA can favor the development of future antibiotics.

2.2 Lipids

All cells are enclosed by biological membranes, composed of lipids and proteins, where the lipids prevents molecules from leaking out, and keep unwanted molecules out of the cell. The proteins on the other hand act as transporters of specific molecules both in and out of the cell. *Lipids* can be defined as biomolecules that are not soluble in water, but are highly soluble in organic solvents, such as chloroform.[16] There are different kind of lipids with different characteristics, however, in this thesis, only *phospholipids* will be discussed. Phospholipids are *amphipathic molecules*, that is, they are composed of a *hydrophilic*, "water loving" part referred to as the head group and a *hydrophobic*, "water hating" part, made up by two fatty acid "tails".[16] This is illustrated in Figure 2.3A. The head group of biological phospholipids are either anionic or zwitterionic, and the chains usually contain some level of unsaturation.[53] Due to the characteristics of these molecules, in an aqueous environment, they self-assemble into different structures to minimize the interaction of the hydrophobic part with the solvent. The aggregation shape depends on the opposing forces of the system where the hydrophobic effect drives the self-assembly, while there is an electrostatic or steric repulsion of the head groups.[53,54] This happens when the lipids are in the *lamellar liquid crystalline* (LLC) phase,[54] which is above the *transition temperature*, T_t of the lipid molecule. Below T_t, the lipids form gels.[55] In this thesis, vesicles (or liposomes) and planar bilayers will be used as model systems to mimic cell membranes. A schematic representation of these self-assemblies are presented in Figure 2.3B and C, for vesicles and planar bilayers, respectively.

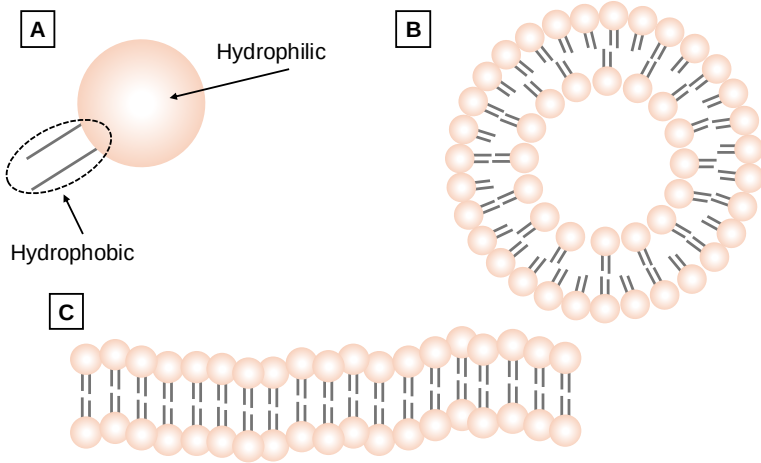


Figure 2.3: Schematic illustration of **(A)** a lipid molecule with its hydrophilic head group and its hydrophobic tails, **(B)** a lipid vesicle, and **(C)** a planar bilayer.

2.2.1 Phospholipids as model membranes

As mentioned, phospholipids were used to model cell membranes in this thesis. While biological membranes are very diverse in both structure and function, they do have some characteristics in common, including (but not limited to) that they contain lipids and protein, they are asymmetric, meaning that the inner part differs from the outer. They have fluid structures, and they have a negative potential. There are three main kinds of lipids in biological membranes are phospholipids, *glycolipids* and *cholesterol*, where phospholipids are the main group of the three.[16] Most membranes contain roughly 50% lipids, and 50% protein (w/w), but varies with different types of membranes. The lipid composition also varies between different cell membranes, where the plasma membrane of *E. coli* contains mostly phosphatidylethanolamine (PE), while mammalian plasma membranes are more complex, containing four major phospholipids; phosphatidylcholine (PC), phosphatidylserine (PS), phosphatidylethanolamine (PE), and sphingomyelin (SM).[56]

Using a system composed of only phospholipids as a model for a biological membrane is a good model from the perspective that this group of lipids is the largest contributor to the lipid fraction of these membranes. However, roughly 50% of the membrane weight comes from proteins, which is not considered in the work of this thesis. Adding membrane proteins would indeed make the model system more realistic, however, it would also add significant complexity, and therefore, a very simple model system of only phospholipids was used, to pin-point the effects caused by the peptides only.

2.2.2 Vesicle solutions

Liposomes are artificial vesicles which can be made of phospolipids. They are spherical in shape and their size can vary. In addition to this, they can be composed of either one, or several bilayers, referred to as unilamellar vesicles and multilamellar vesicles, respectively.[57] The unilamellar vesicles can be further divided into *giant unilamellar vesicles* (GUV, $>$ 1 µm), *large unilamellar vesicles* (LUV, 100 nm–1 µm), or *small unilamellar vesicles* (SUV, 20 nm–100 nm), depending on their size. However, this definition is not well defined, and can therefore vary between authors.[58] In this thesis, a vesicle of 100 nm will be referred to as a LUV, while in Paper I, a vesicle with a diameter of approxiamtely 150 nm is referred to as a SUV, which only points to the lack of a fixed definition.

There are a variety of methods to form vesicles, but they can be divided into two classes, solvent free methods, which are performed without organic solvents, and solvent displacement methods, in which the lipids are first dissolved in organic solvent, which is later switched to an aqueous one, and the organic solvent is removed.[58] In this thesis only solvent free methods will be discussed. For all solvent free methods, the first step is to form a lipid film. Despite the definition of the method, a lipid film is obtained by dissolving the lipids in organic solvent and then evaporating the solvent to leave behind a lipid film. The definition of the solvent free method refers to the absence of organic solvent in all subsequent steps. Once a lipid film is obtained, it is rehydrated in aqueous solution, which causes the lipid film to swell and form a polydisperse solution of vesicles. This step can be promoted by mechanical stirring or sonication. To obtain a monodisperse vesicle solution, this step is usually followed by extrusion, where the vesicle solution is passed through a polycarbonate membrane with small pores.[58] Both sonication and extrusion is used for the work in this thesis, and is described more in detail in Chapter 4.

2.2.3 Supported lipid bilayers

Supported lipid bilayers, or SLBs for short, are composed of a single lipid bilayer, which is deposited on top of a solid surface. Several different materials, including, but not restricted to, gold, mica, silicon, and sapphire can be used as a solid surface, depending on the nature of the sample, as well as what technique the system will be investigated with. The SLBs can be produced with several different methods, however, the most common are the *vesicle fusion approach* and the *Langmuir-Blodgett* and *Langmuir-Schaefer deposition approach*.[59] For the work included in this thesis the vesicle fusion approach has been used, and the description of the practical steps of this approach are given in Chapter 4. In short, vesicles adsorb to the solid surface and are either spontaneously

ruptured to form a planar bilayer, or the rupture is induced by osmotic shock. This method produces symmetric bilayers where both the inner and the outer leaflet have the same composition. If asymmetric bilayers are required, the Langmuir-Blodgett and Langmuir-Schaefer deposition method has to be used, where each leaflet of the bilayer is formed separately in the air-liquid interface, allowing for control of the content of each leaflet.[59]

There are both advantages and disadvantages using SLBs as model systems for biological membranes, where their robustness and reproducibility are prominent advantages. Additionally, working with SLBs allows for the use of several surface-sensitive techniques such as *quartz-crystal microbalance with dissipation monitoring* (QCM-D) and *ellipsometry*, among others, to obtain information about interactions with these model membranes, which otherwise would not be accessible with these techniques. The most prominent disadvantage, however, is that the SLB is in close contact with the solid surface, which affects both the diffusion of lipids and proteins as well as the fluctuations of the bilayer, which are suppressed by the proximity to the solid surface.[59]

To overcome the problem with the effect of the solid surface, while still maintaining the ability to use the surface-sensitive techniques, more advanced models for membranes can be used, where the interaction between the bilayer and the underlying surface is minimized. This can be done in different ways, which can broadly be characterized into three groups: floating, tethered, and cushioned bilayers.[59] In this thesis, only cushioned bilayers will be discussed.

In cushioned bilayers, a polymer used should act as a lubricating layer between the surface and the bilayer. It is important that the interfaces between the surface and the hydrated polymer, and between the hydrated polymer and the bilayer are completely wetted, to ensure stability of the system. Additionally, it is important for stability that there are repulsive interaction between the bilayer and the solid surface.[60] A cushioned bilayer can spontaneously form by addition of a polymer of weak positive charge after successfully forming an SLB. The polymer then diffuses beneath the bilayer, cushioning it from the underlying surface.[61] The cushion can also be formed by adding vesicles to a dried layer of polymers already dispersed on the solid surface.[62] The formation of a cushioned bilayer will be discussed throughout this thesis.

There are three kinds of men. The one that learns by reading. The few who learn by observation. The rest of them have to pee on the electric fence for themselves.

- Will Rogers

3 | Theoretical background

Contents

In this chapter, a short background to the theoretical parts relating to molecular simulations of this thesis is given, including statistical thermodynamics, as well as intermolecular interactions. These two topics are the basics of the computations performed in this thesis, and a brief introduction of both will be given here.

3.1 Statistical thermodynamics

From *thermodynamics*, we are able to connect experimental properties to macroscopic systems in equilibrium. What is not provided by thermodynamics is an explanation, or interpretation, on a molecular level of these properties at equilibrium. Instead, that is given by *statistical mechanics*. When statistical mechanics is applied to explain thermodynamics, it is usually referred to as *statistical thermodynamics*. In this thesis, a brief introduction of the key concepts will be provided, and the interested reader is therefore

encouraged to seek a more in-depth description, which, for example, can be found in the book by Hill.[63]

The first important concept in statistical mechanics is *an ensemble*, which is a mental collection of a very large number of systems. These systems are all equivalent on a thermodynamic (macroscopic) level, but differ on a microscopic level. The ensembles can be classified using the macroscopic properties they represent, which are outlined below.

Microcanonical (NVE) **ensemble** corresponds to an isolated system where the energy, E, the volume, V, and the number of particles, N, are constant.

Canonical (NVT) **ensemble** represents a thermodynamic system where the volume, the number of molecules, and the temperature, T, are all kept constant. It is important to note that this system is not isolated, thus even though the above mentioned observables are constant, the energy can fluctuate. The possibility to find the system in an energy state E_i is

$$P_i(N, V, T) = \frac{\exp\left[-E_i(N, V)/kT\right]}{Q(N, V, T)}, \tag{3.1}$$

where Q is the canonical *partition function*, given by

$$Q(N, V, T) = \sum_i \exp\left[-E_i(N, V)/kT\right]. \tag{3.2}$$

$\exp\left[-E_i(N, V)/kT\right]$ is known as the *Boltzmann weight*. The partition function describes the equilibrium statistical properties of the system. This can express the *Helmholtz free energy, A*, which is the characteristic function for the canonical ensemble in thermodynamics and can be used to derive other important variables, including entropy, pressure, and the total energy. Helmholtz free energy is given by

$$A = -kT \ln Q. \tag{3.3}$$

Grand canonical (μVT) **ensemble** corresponds to an open, isothermal thermodynamic system where the volume, the temperature and the chemical potential, μ, are controlled and kept constant.

Isothermal-isobaric (NpT) **ensemble** has a constant number of molecules, a constant pressure, p, and constant temperature.

In this thesis, simulations have been performed with the canonical, as well as the isothermal-isobaric ensemble.

When conducting an experiment, a time average is taken over the observable, giving a *time average*. If we wanted to obtain the same observable from molecular properties, a large number of molecules would have to be treated, in addition to observe them for sufficiently long time. This would in practice be very difficult, however, there is another approach, namely the *first postulate of statistical mechanics*. This postulate states that a (long) time average of a mechanical variable in a thermodynamic system is equal to the ensemble average of the variable in the limit of an infinitely large ensemble, provided that the ensemble replicate the thermodynamic state and environment. What it actually says is that the time average on one system may be replaced with an instantaneous average over a large number of systems representing the actual system. This postulate is valid for all ensembles, and is the basis of molecular simulations. The *second postulate of statistical mechanics* only applies to the microcanonical ensemble, and states that for an infinitely large ensemble of an isolated thermodynamic system, the systems of the ensemble are uniformly distributed over the possible states consistent with the specified values of N, V and E. This is also referred to as the *principle of equal a priori probability*, meaning that all microscopic states are equally probable.

In the description of the partition function, Equation 3.2, it was introduced in a quantum mechanical formulation with discrete energy states. Many simulation methods are however based on classical mechanics. Then, the microstates are so close in energy that they are instead treated as a continuum. Treating the canonical partition function classically gives

$$Q_{classical} = \frac{1}{N!h^{3N}} \int \exp\left[-H\left(\mathbf{p}^N, \mathbf{r}^N\right)/kT\right] \mathrm{d}\mathbf{p}^N \mathrm{d}\mathbf{r}^N, \tag{3.4}$$

where h is Planck's constant and the integration is performed over all momenta $\mathbf{p}^N$ and all coordinates $\mathbf{r}^N$ for all N particles. $H\left(\mathbf{p}^N, \mathbf{r}^N\right)$ is the Hamiltonian of the system, which consists of two parts, one for kinetic energy, which depends on temperature, and one for potential energy, which depends on interactions. The kinetic part can be integrated directly, which simplifies the partition function to

$$Q_{classical} = \frac{Z_N}{N!\Lambda^{3N}}, \tag{3.5}$$

where Z_N is the *configurational integral* calculated from the potential energy, U_{pot}, and

is given by

$$Z_N = \int_V \exp\left[-U_{pot}(\mathbf{r}^N)/kT\right] \mathrm{d}\mathbf{r}^N, \tag{3.6}$$

and Λ is the *de Broglie wavelength* given by

$$\Lambda = \frac{h}{(2\pi mkT)^{1/2}}. \tag{3.7}$$

If the configurational integral is known, the ensamble average of an observable, X, can be calculated according to

$$\langle X(\mathbf{r}^N)\rangle = \frac{\int_V X(\mathbf{r}^N) \exp\left[-U_{pot}(\mathbf{r}^N)/kT\right] \mathrm{d}\mathbf{r}^N}{Z_N}. \tag{3.8}$$

Solving the integrals requires numerical tools, such as the Monte Carlo method or molecular dynamics, which will be discussed in Chapter 5.

3.2 Polymer physics

Before discussing the intermolecular interactions important for the work in this thesis, a short introduction to polymer physics will be given, which is based on the book by Evans and Wennerström,[54] where the interested reader can find further information. A *polymer* is a molecule made up of several smaller units, *monomers*, connected by covalent bonds. These polymers can be *homopolymers*, consisting of only one kind of monomers, *heteropolymers*, made up of different kinds of monomers, or *block copolymers*, which contains a block of repeating monomers, followed by one or several block(s) of other monomers. Proteins and peptides are heteropolymers made up of different amino acids, as described in Chapter 2. It is therefore possible to explain protein interactions using the same theories as applied to polymers.

The solubility of polymers depends on how good the solvent is, which can be described as if the monomer-monomer intermolecular interactions are more favorable than the monomer-solvent intermolecular interaction, more on this in the next section. It also depends on the *entropy* of the system, which has two contributions; the mixing of the polymer with the solvent taking the polymers from a pure state where it is only surrounded by other polymer molecules, mixing it with water, increasing the entropy of the system. The second one is that the polymer can assume many more conformations

in solution, which also increases the entropy of the system, increasing the solubility. The solubility dependence described here is valid for dilute samples.

Charged polymers are referred to as *polyelectrolytes*, and, as stated in Chapter 2, both Hst5 and KEIF, has a net positive charge, and can therefore be considered as polyelectrolytes. If the geometry of a polyelectrolyte is described as a cylinder, a phenomenon referred to as *ion condensation* appears. The counterions can then approximately be divided into two groups. The first group is called condensed ions, and these counterions are electrostatically bound to the charged cylinder. The other group of counterions are diffusely bound.

There are many more characteristics that are important for polymers, however, what is presented here are the most relevant parts for the work of this thesis.

3.3 Intermolecular interactions

On the molecular level, there are two classes of forces responsible for interactions: (i) the *intra*molecular forces acting *within* molecules. These can, for example, connect atoms to form molecules. (ii) The *inter*molecular forces acting *between* atomic and molecular species, without creating chemical bonds between them. The latter ones will be discussed in this section, and are central in explaining both the structural and conformational properties of proteins and peptides, and to explain the interaction between peptides and lipids. The intermolecular interactions included in this chapter can be classified according to their distance dependence, to either long-ranged, or short-ranged. The descriptions and equations presented are based on those given by Israelachvili.[64]

3.3.1 Coulomb interactions

The *Coulomb force* acts between two charged atoms, or ions, and is the strongest of the intermolecular forces considered here. It is a long-ranged interaction, with a $1/r$-dependence. For two charges, Q_1 and Q_2, separated by a distance r, the magnitude of the force, $F(r)$, is given by,

$$F(r) = \frac{Q_1 Q_2}{4\pi\varepsilon_0\varepsilon_r r^2}, \tag{3.9}$$

where ε_0, and ε_r denotes the vacuum electric permittivity, and the dielectric constant,

respectively. The *Coulomb interaction* between the two charges is obtained from

$$w(r) = \int_{\infty}^{r} -F(r)dr = \frac{Q_1 Q_2}{4\pi\varepsilon_0\varepsilon_r r} = \frac{z_1 z_2 e^2}{4\pi\varepsilon_0\varepsilon_r r}. \tag{3.10}$$

The switch from Q_i to the ion valency, z_i, multiplied by the elementary charge, e (1.602×10^{-19} C), is done for systems in solution. This interaction is long-ranged. However, if the charges are surrounded by ions, as they are in an aqueous salt solution, the interaction is screened. This reduces the range of the interaction. The screened Coulomb potential can be described as

$$V(r) = \frac{z_i z_j e^2}{4\pi\varepsilon_0\varepsilon_r r} \exp\left(-\kappa r\right), \tag{3.11}$$

according to the Debye-Hückel theory.[65] $V(r)$ is the potential energy, and κ is the *Debye length*, defined by

$$\kappa^{-1} = \sqrt{\frac{\varepsilon_0\varepsilon_r kT}{2N_A e^2 I}}, \tag{3.12}$$

where k is the Boltzmann constant, T is the temperature, N_A the Avogadro constant, and I refers to the ionic strength, which is defined as

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2. \tag{3.13}$$

Here, n denotes the number of different ion species, and c_i is the concentration of ion i with valency z_i.

3.3.2 Van der Waals interactions

The *van der Waals interactions* are a collection of three types of interactions, (i) the *Debye interaction*, which is only active when one of the molecules is a permanent *dipole*. A dipole referrers to that even if a molecule is not charged, it can possess an *electric dipole*, meaning there is an asymmetry in the electron distribution within the molecule. The second contribution is (ii) the *Keesom interaction*, where both interacting molecules have to be permanent dipoles, and (iii) the *London dispersion interaction*,

which is present between all types of molecules. The basis of the dispersion interaction can be explained as follows; even if the time averaged dipole moment of a molecule is zero, it possesses instantaneous dipole moments due to electron movements. This can generate induced dipoles in surrounding molecules, giving rise to interactions in the system. This interaction is usually attractive, but can also be repulsive in a system containing more than two particles.

The interaction free energy of two different polar molecules is given by

$$\begin{aligned} w(r) &= -\Big[C_{Debye} + C_{Keesom} + C_{London}\Big] \\ &= -\left[(u_1^2\alpha_{02} + u_2^2\alpha_{01}) + \frac{u_1^2u_2^2}{3kT} + \frac{3\alpha_{01}\alpha_{02}h\nu_1\nu_2}{2(\nu_1+\nu_2)}\right] \Big/ (4\pi\varepsilon_0)^2 r^6, \end{aligned} \tag{3.14}$$

where h is the Planck constant, and ν_i is the orbiting frequency of each atom. As shown in Equation 5.5, this interaction is short-ranged, with a $1/r^6$-dependence.

3.3.3 Hydrogen bonds

When a hydrogen atom is covalently bonded to an electronegative atom, such as nitrogen, oxygen, or fluorine, and gets in close proximity to another such electronegative atom, a *hydrogen bond* is formed. It is primarily of electrostatic origin, and can be thought of as a strong dipole-dipole interaction. The bond is directional, and can also orient surrounding molecules. An example of a hydrogen bond is depicted in Figure 3.1.

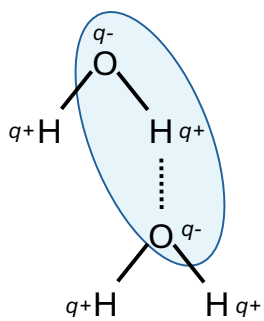


Figure 3.1: Illustration of a hydrogen bond (dotted line between the hydrogen and oxygen atoms) between two water molecules.

3.3.4 The hydrophobic interactions and the hydrophobic effect

Water molecules can form hydrogen bonds between each other, making the water-water interaction strong. If a non-polar molecule is introduced in water, it is more beneficial for the water molecules to interact with other water molecules, compared to interacting with the non-polar molecule. The water molecules therefore rearranges around the non-polar molecule in a way that maintains as many hydrogen bonds as possible. For small non-polar molecules, no bonds need to be broken. The new arrangement is however more ordered than the original state, leading to a decrease in entropy, which is unfavorable for the system. The separation between non-polar molecules and water, to maintain as many hydrogen bonds as possible, is referred to as the *hydrophobic effect*. The attraction between non-polar molecules is often stronger in an aqueous solution, compared to in free space. This is referred to as the *hydrophobic interaction*.

3.3.5 Charge regulation

As described for the amino acid histidine in Chapter 2, *charge regulation* is when a molecule changes it charge state in the vicinity of another molecular matter, which is possible due to migration of protons to and from titratable sites. The molecule can then adapt to the changes in the chemical environment. The charge regulation mechanism is present for all macromolecules with titratable sites. This concept will be discussed further in this section, and is based on the work by Lund and Jönsson,[27] which can be referred for more detailed descriptions.

Charge regulation can be explained from basic thermodynamic arguments, where we first consider an equilibrium process for a single ionizable site

$$\mathrm{HA} \rightleftharpoons \mathrm{H^+} + \mathrm{A^-}, \tag{3.15}$$

for which, the corresponding equilibrium constant is

$$\begin{aligned} K_a &= \frac{\gamma_{H^+}\gamma_{A^-}}{\gamma_{HA}} \frac{\rho_{H^+}\rho_{A^-}}{\rho_{HA}} \\ &\Leftrightarrow -\log \frac{\rho_{HA}}{\rho_{A^-}} = \mathrm{pH} - \mathrm{p}K_a + \log \frac{\gamma_{HA}}{\gamma_{A^-}}. \end{aligned} \tag{3.16}$$

ρ denotes the average concentrations, and γ are activity coefficients. The protons in the system migrate to and from the site due to thermal motion, and the fraction ρ_{HA}/ρ_{A^-} defines the average charge state of the site. Charge regulation can be described thorough

a situation where pH equals pK_a, Equation 3.16 then shows that the average charge state is determined by the activity coefficients. In close proximity to other molecules, γ will change, as it accounts for interactions with other molecules in solution. Due to the fact that, at equilibrium, the chemical potential or activity, $a_i = \rho_i \gamma_i$ must be constant throughout the medium, if the activity coefficient changes, so must the concentrations. The charge regulation mechanism can also be described in a more microscopic way, with a statistical mechanical description. This is extensively done in the work by Lund and Jönsson,[27] which the interested reader is referred to. For the purpose of this thesis, the most important parts will be described and discussed, avoiding largely the equations and relations.

It is commonly only the average charge of a titratable site that is considered, when in reality, it is several values of different probabilities, distributed by a Gaussian distribution around the average charge. It is this distribution that gives rise to the charge regulation mechanism, and the most important molecular parameter is the *capacitance*, defined as the variance of the mean charge. When a molecule is exposed to an external electric potential, the capacitance is a measure of how much charge can be induced. This parameter strongly depends on the surrounding pH solution, and the pK_a-values of titratable groups in the molecule.

In proteins/peptides, which this thesis is focused around, this mechanism applies to amino acids with titratable sites, such as aspartic acid, histidine, etc., which all have different pK_a-values. This interaction applies for a number of interactions, such as protein-protein interaction, ligand binding, and protein-membrane interactions, which is of interest in this thesis.

3.3.6 Surface adsorption

In this section, adsorption of polymers to surfaces will be discussed, based on the book by Evans and Wennerström.[54]

A polymer in solution will adsorb to a surface if the adsorption energy is large enough. Even though each monomer interacts with a relatively small energy, the fact that polymers make many contacts with the surface makes the resulting adsorption energy high. A polyelectrolyte will adsorb on a oppositely charged surface due to electrostatic forces. There is also an entropic contribution to this interaction, where the number of conformations the polyelectrolyte can assume will significantly decrease upon adsorption, decreasing the entropy, which makes the interaction less favorable. This decrease is however small compared to the increase in entropy when the polyelectrolyte can act as a giant counterion on the surface, releasing the small counterions at the surface, while simultaneously releasing its own counterions. The adsorption of a polyelectrolyte to

an oppositely charged surface is mainly electrostatically driven, and the conformation of the polymer at the surface is highly dependent on salt concentration in an aqueous solution. With low or no salt present, the polyelectrolyte will adsorb so that the charges match optimally, and assume a flat conformation on the surface. As the salt concentration is increased, the effect of the electrostatic interaction is decreased, as indicated in Equation 3.11, and the polyelectrolyte can obtain a higher conformational entropy by releasing some parts from the surface, where the ions of the electrolyte can act as counterions instead.

If there are more than one surface present, and not enough polymers to saturate both surfaces, a polymer can adsorb to both surfaces, resulting in a attractive force called *bridging*.

There's no way you can rely on an experiment.
- Gerrit Groenhof

4 | Experimental methods

Contents

In this chapter, the experimental techniques used in the thesis are described. Circular dichorism (CD) and small angle X-ray scattering (SAXS) were used to determine the size, shape, and structure of the peptides. Quartz-crystal microbalance with dissipation monitoring (QCM-D) and neutron reflectometry (NR) are surface active techniques, which were used to determine the characteristics of the interaction of the peptides with SLBs.

4.1 Protein purification and concentration determination

Hst5 and variants thereof, as well as KEIF, used in this work, were purchased as lyophilized powders. Before the peptides were further used in any experiments, they were purified by dialysis to remove any impurities and buffer remains. The dialysis were performed either against MilliQ water, or directly against the desired buffer for the planned experiment, at 6°C using membranes with a 0.5–1 kDa cutoff. The dialysis was performed during 24 to 96 h, with exchange of dialysis buffer two to three times per 24 h. The peptide sample was then either used directly in measurement, or, if dialyzed against MilliQ water, the sample was lyophilized so that it could be dissolved in any buffer or solvent.

For all experiments performed for the work presented in this thesis, the experimental technique has depended upon the peptide concentration in the sample, or it has been of general interest to be able to compare results acquired for different samples or at different occasions. For this reason, for almost all samples, the protein concentration has been determined using a Nanodrop 2000 Spectrophotometer. For all peptides, measurements were performed at 280 nm, and the same extinction coefficient of 2650 M^{-1} cm^{-1} was used for Hst5 and all variants of the same length. For the shorter ^{14}Hst5 peptide, a value of 1490 M^{-1} cm^{-1} was used, and for the longer ^{48}Hst5, an extinction coefficient of 5960 M^{-1} cm^{-1} was utilized, all obtained from the Expasy tool ProtParam.[66] The exception of this concentration determination method is for samples dissolved in 2,2,2-trifluoroethanol (TFE). For these samples, the concentration was determined only by weight to volume ratio.

For KEIF, the concentration of the samples used in synchrotron radiation circular dichroism (SRCD) was determined from the absorbance at 205 nm,[67] which corresponded well with the expected concentration from the known weight-to-volume ratio. Therefore, for the other techniques, the concentration was set by weight-to-volume ratio of the peptide powder and buffer solution.

4.2 Vesicle preparation and vesicle fusion protocol

Two different lipids were used in this study, namely 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine (POPS), as well as their partially deuterated homologues d_{31}POPC (1-palmitoyl-d31-2-oleoyl-sn-glycero-3-phosphocholine) and d_{31}POPS (1-palmitoyl-d31-2-oleoyl-sn-glycero-3-phospho-L-serine), which all were purchased as lyophilized powders. As the first step in the preparation of vesicles, stock solutions were prepared in a 3:7 meth-

anol:chloroform mixture, using the following lipid molar ratios (POPC:POPS) 100:0, 9:1, 6:4, and 3:1. These ratios are indicated in the acronyms of the lipid mixtures as $PC_{POPC}PS_{POPS}$. Use of partially deuterated lipids is also indicated in the acronyms (usually as d_{31}). The T_t of POPC and POPS are -2°C, and 14°C, respectively.[68]

The methanol:chloroform mixture was evaporated under nitrogen flow to form a lipid film, and any remaining solvent was evaporated under reduced pressure, 0.8 bar. The lipid films were hydrated in either 500 mM NaCl, and 20 mM TRIS buffer at pH 7.4 for the charged lipid mixtures, or for the pure zwitterionic lipids, in MilliQ water. Two methods were utilized to form LUVs, either *tip sonication*, where the lipid solution was treated for a total of 30 min with 30% maximum amplitude, by pulsing with 2 s ON and 3 s OFF, or *extrusion* where the lipid solution was pushed 31 times through a 0.1 µm polycarbonate membrane filter.

To obtain supported lipid bilayers, SLBs, the vesicle fusion[69,70] protocol were optimized by means of QCM-D measurements, using silicon dioxide coated sensors. The injection of LUV solution into the flow module was followed by a series of rinsing steps, dependent of the system. Pure zwitterionic vesicles composed of only POPC was deposited following published protocols,[69,71] however, for the charged vesicles, a different method was required. It was found that the most efficient way to obtain high-quality charged SLBs on silicon dioxide was to first fill the flow module with buffer containing 500 mM NaCl, followed by injection of vesicles prepared in the same buffer, and lastly rinse the flow module with MilliQ water after an incubation time of 60 min to induce osmotic shock.

For depositions on sapphire surfaces, no protocol development was necessary, since zwitterionic and charged vesicles fused spontaneously without any necessity of osmotic shock, nor use of salt solutions.

4.3 Circular dichroism spectroscopy

CD spectroscopy is a method used to determine the secondary structure (far-UV CD, $\lambda \leq 240$ nm), fingerprints of the tertiary structure (near-UV CD, 250 –320 nm), and structural changes of proteins/peptides. In this thesis, far-UV CD is utilized to determine secondary structure elements of the studied peptides in different environments. The CD method is, compared to other structure determining techniques, such as X-ray crystallography and nuclear magnetic resonance spectroscopy, a low resolution technique, however, it is less demanding than the other techniques, both considering amount of sample and time.[72]

4.3.1 Basic principle

In the task of describing the basic principles of CD, it is appropriate to start by describing and defining polarized light, which is a type of *electromagnetic radiation*, comprised of an electric field and a magnetic field. These fields oscillate in perpendicular planes, which are perpendicular to the direction of propagation of the light. Normally, light oscillates in all possible directions, meaning it is *unpolarized*. However, if the oscillations are restricted to one direction, the light is *linearly polarized*, which is illustrated in Figure 4.1A. If the electric vector instead rotates around the direction of propagation, undergoing a full revolution per wavelength, the light is *circularly polarized*. A clockwise rotation corresponds to a right circularly polarized light, and an anticlockwise rotation to a left circularly polarized light.[73]

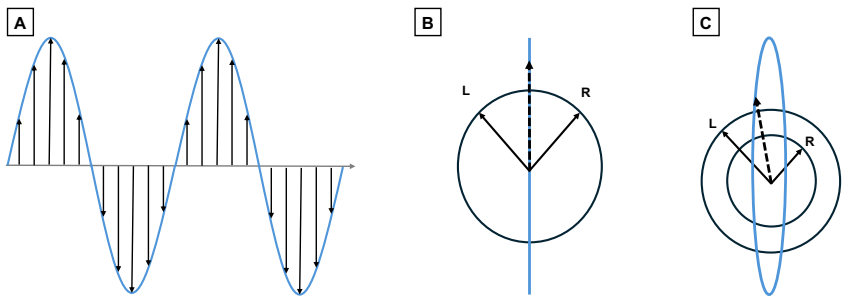


Figure 4.1: (A) Schematic illustration of linearly polarized light. The direction of propagation is indicated by the grey arrow, and the black arrows corresponds to the electric vector at different points along the propagation. **(B)** Circularly polarized light L and R rotating in opposite directions making up linearly polarized light. The electric vector is represented by the dashed arrow and corresponds to the sum of the two components, which is always oriented along the blue line. **(C)** The two componentes have different amplitude, causing the electric vector (dashed line) to trace and ellipse.

One can think of linearly polarized light as made up by two components of circularly polarized light of equal magnitude, where one rotates clockwise (right handed, R) and the other one counter-clockwise (left handed, L), illustrated in Figure 4.1B. If the two components are not of the same amplitude, the light is instead *elliptically polarized*, and the electric vector traces an ellipse, as shown in Figure 4.1C. During a CD spectroscopy experiment, when the circularly polarized light passes through an *optically active* sample, the left and right polarized light is adsorbed to a different extent, resulting in that the radiation is elliptically polarized.[72]

A *chromophore*, that is, a light-absorbing molecule, is optically active if it is (i) inherently *chiral*, (ii) covalently attached to a chiral center, or (iii) in a chiral environment due to the three-dimensional structure of the molecule. The chromophore of main interest in proteins/peptides is the backbone group, where the C_α is inherently chiral

with the exception of glycine. This gives information of the secondary structure of the protein/peptide. The CD spectrum obtained from a measurement is the summation of signals from all the peptide chromophores in the sample. Characteristic patterns for each secondary structure element is presented in Figure 4.2.[73]

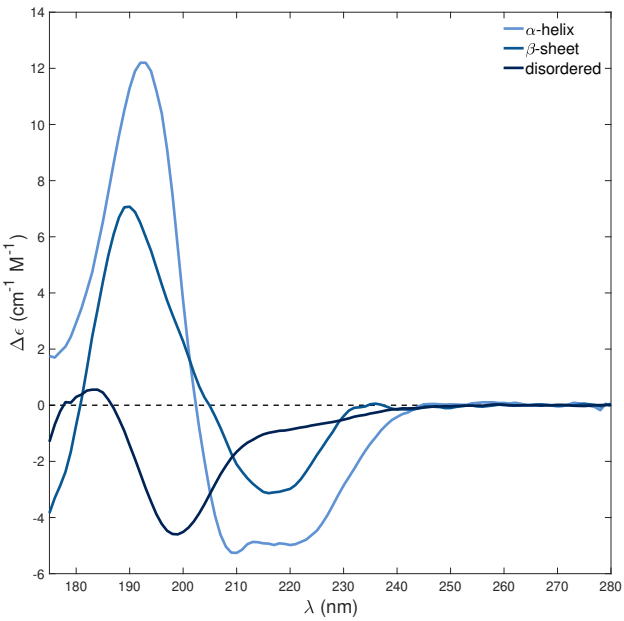


Figure 4.2: CD spectra obtained for proteins with different secondary structure elements. The spectra are obtained from the Protein Circular Dichroism Data Bank[74] with the following spectrum id: CD0000117000 (α-helix)[75], CD0000118000 (anti-parallel β-sheet)[75], and CD0006124000 (disordered).[76]

4.3.2 Measurement setup and data analysis

The CD spectroscopy measurements conducted for the projects included in this thesis were performed either with a lab based instrument (JASCO J-715 spectropolarimeter with a photomultiplier tube detector), or at the AU-CD beamline at the synchrotron light source ASTRID2 (Department of Physics and Astronomy, Aarhus University, Denmark). In both cases, spectra were recorded every 1.0 nm, and each scan was repeated at least three times. Subtraction of appropriate reference spectra (buffer/solvent/lipid solution, depending on the investigated system) was performed on all spectra. For the measurements performed on the lab based instrument, a quartz cuvette with a 1 mm path length (110-1-40, Hellma Analytics) was used, while, at the AU-CD beamline, a 0.1 mm path length quartz cuvette (121-0.10-40, Hellma Analytics) was

used.

During the measurement, the difference in absorption of the R and L circularly polarized light is measured for different wavelengths. Despite measuring the absorbance, the results are most commonly reported as ellipticity (θ) in degrees, which is related to the absorbance, ΔA, through $\theta = 32.98\Delta A$.[72] In order to compare different measurements, the signal has to be normalized, as the magnitude of the CD signal depends on the sample concentration and the path length. For this, the signal is often expressed as either the mean residue ellipticity, given in the unit deg cm^2 dmol^{-1}, which is calculated as

$$[\theta]_{MRW} = \frac{\theta \cdot MRW}{10 \cdot d \cdot c}, \tag{4.1}$$

where θ is the observed ellipticity (mdeg), d is the path length of the cuvette (cm), c is the protein/peptide concentration (mg mL^{-1}), and MRW is the molecular weight (Da) divided by the number of peptide bonds.[72] Or the data is presented as delta epsilon, $\Delta\varepsilon$, which is related to the molar ellipticity simply by a factor of 3298.[77]

$$\Delta\varepsilon = \frac{[\theta]_{MRW}}{3298} \tag{4.2}$$

To obtain a good signal from the protein/peptide, one has to ensure that the absorbance of the buffer is low. Chloride ions are known to have a high absorbance in the far UV region, where it would interfere with the protein/peptide signal.[72] In the experiments conducted for this thesis, sodium fluoride was used to set the salt concentration rather than sodium chloride, and phosphate buffer was used rather than TRIS, to avoid the problem chloride ions can cause to the data collection.

It is usually possible to determine the dominating secondary structure elements present within a sample by simply observing the shape of the CD spectrum. This is also a straightforward method to detect conformational and structural changes as a consequence of changes in environment, such as temperature, salt concentration, or solvent. In the work included in this thesis, the effect of the latter two were investigated, using aqueous buffers of different salt concentrations, which was compared to TFE. In addition to this simple method, one can obtain a quantified determination of the composition of the secondary structure elements through several available methods, e.g. BeStSel[78,79] and DichroWeb[80], which are both available online. All these methods are based on the approximation that a CD spectrum can be expressed as a linear combination of spectra with different secondary structure elements.[77] For this reason, a good set of reference data is needed for the results, where a large reference

set is often favorable to account for structural variability within a secondary structure element. Results can vary depending on the method used and which reference set is included. Since unordered structure is not a structural element, but rather the lack of other elements, the variability for this structure is especially large. This means that determining the secondary structure element content of IDPs from CD data is quite challenging and, in general, one should be careful when deducing a quantitative conclusion about a sample, and to primarily use CD spectroscopy as an indicative tool for changes in secondary structure elements.

4.3.3 Oriented circular dichroism

Oriented CD (OCD) is circular dichroism spectroscopy of oriented samples which can be used to determine the orientation of a peptide with respect to a lipid membrane.

The OCD measurements included in this thesis were performed at the AU-CD beamline on ASTRID2, Aarhus University, Denmark. The sample film was formed by pipetting a small amount of sample onto a quartz window (202-40QS, Hellma Analytics). The sample was then dried under a protective cover, followed by a rehydration step where the film was enclosed in a container with saturated K_2SO_4 to achieve a relative humidity of $\sim$ 98%. After rehydration, the quartz window with the film was mounted in the sample holder with a few drops of saturated K_2SO_4 to keep the sample hydrated during the measurement. Windows with the films were placed in a holder, with a second clean window enclosing an area with the hydrating solution within. The holder was then placed into a rotation stage, controlled by a computer, fitting into a temperature controlled chamber. Linear dichroic effects due to alignment of the films were avoided by measuring the CD spectrum at four angles, by rotating the sample perpendicular to the beam path. The data was acquired every 90°, with the 0° measurement repeated to check film stability during measurement. The data was averaged over all angles, and presented in terms of mdeg.

OCD analysis has mainly been performed on α-helical peptides, and is based on Moffitt's theory,[81] which will be explained in brief. Depending on the orientation of the peptide, with respect to the membrane normal, a negative Gaussian band around 208 nm is present. The highest amplitude represents the peptide oriented perpendicular to the membrane normal, that is, laying flat on the surface. When the α-helical peptide is instead parallel with the membrane normal, that is, fully inserted into the membrane, the negative band disappears. The transition between the two states is a partially inserted state called a tilted state. For this state, the amplitude of the negative band around 208 nm is between the surface state and the inserted state. A more in-depth description of the analysis can be found elsewhere.[82,83] OCD was included in Paper IV, where a comparison was done with solution CD measurements, as well

as a comparison between a film composed of only KEIF molecules and mixed films composed of PC_3PS_1 and KEIF.

4.4 Small angle X-ray scattering

SAXS is a method to analyze the structure of condensed matter. The technique can be applied on a variety of fields, including, for example, emulsions, porous materials, and nanoparticles. However, the focus of this thesis will be on biological macromolecules in solution, more specifically, peptides/proteins in aqueous solution. The technique allows for investigation of the structure of non-crystalline biochemical systems, where one obtains low-resolution structural information on the overall shape and internal structure. It is also possible to study assembly, and large-scale conformational changes in real time of particles in the size-range from one to a few hundred nanometers. It is important to remember that one obtains the ensemble average of the scattered particles in the sample, rather than the specific properties of each particle.[84]

4.4.1 Basic principle

As suggested by the name *small angle X-ray scattering*, the fundamental principle of this technique is the scattering of X-rays. The sample of interest is irradiated by a focused and collimated beam of X-rays, that interacts with the electrons in the atoms, which scatters the radiation. The scattering can be either *inelastic*, where energy is exchanged or transferred, or *elastic*, without energy exchange. The latter is important in SAXS. In SAXS, the scattered waves will interfere with each other, giving rise to a 2D interference pattern at the detector, from which structural information about the sample can be extracted. A schematic representation of a SAXS measurement is presented in Figure 4.3.

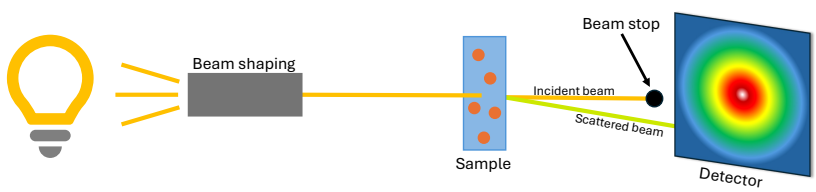


Figure 4.3: A schematic illustration of the main components in a SAXS experiment.

The momentum transferred, more commonly referred to as the scattering vector, $\mathbf{q}$, is defined as the difference between the incident and scattered wave vectors, $\mathbf{k_i}$ and $\mathbf{k_s}$,

as shown in Figure 4.4. The magnitude of the incident wave vector is $\|\mathbf{k_i}\| = 2\pi/\lambda$, where λ is the wavelength of the incident beam. Since the scattering is elastic and there is no loss of energy, $\|\mathbf{k_s}\| = \|\mathbf{k_i}\|$, and the magnitude of $\mathbf{q}$ can be expressed as

$$q = |\mathbf{q}| = \frac{4\pi}{\lambda} \cdot \sin\theta, \tag{4.3}$$

where 2θ is the angle between the incident and scattered wave vector.[85]

As mentioned above, the X-rays are scattered due to interactions with the electrons in the atoms in the sample, which means the more electrons a sample contains, the stronger the scattering signal. Contrast is therefore created by a difference in electron density throughout the sample, which for biological macromolecules is not very different from the aqueous solution, as they contain mostly light elements such as hydrogen and carbon. Since the electron density difference is small, the resulting signal is especially weak.[84]

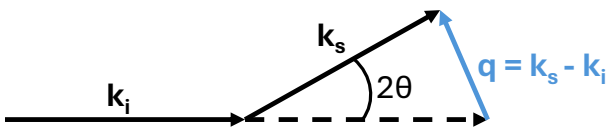


Figure 4.4: A schematic representation of the scattering vector, $\mathbf{q}$, defined by the difference between the incident wave vector, $\mathbf{k_i}$, and the scattered wave vector, $\mathbf{k_s}$.

The measured data is presented in terms of the *scattering intensity*, $I(q)$, which is a function of the magnitude of the scattering vector. It depends on both the volume of the particle and the electron density, hence, larger particles give rise to larger scattering intensities. The scattering intensity can be expressed as

$$I(q) = A \times P(q)S(q), \tag{4.4}$$

for a system of identical homogeneous spheres. A is a constant term containing, among other things, the particle contrast and the concentration. $P(q)$ is the *form factor*, which represents the *intra*-particle distances, from which size and shape of the individual particle can be determined. $S(q)$ is called the *structure factor*, which instead represents *inter*-particle distances, and for the peptide systems investigated in this thesis, gives information if there is attraction or repulsion between the particles. This shows up in the low q region as an increase in the intensity in the scattering curve (attraction) or a decrease (repulsion). In systems which are dilute and weakly interacting the structure

factor is a constant, and the form factor can be determined. When working with dilute samples it is important to keep the samples pure, as the scattering from small molecules will be completely dominated by larger particles. To minimize this effect for the work in this thesis, the samples were all centrifuged before measurements, to remove potential aggregates and larger particles. In addition, the lowest concentration samples were performed in replicas, which could be averaged in the analysis process.

4.4.2 Measurement setup and data analysis

The SAXS measurements included in this thesis were all performed at the BM29 BioSAXS beamline at the ESRF, Grenoble, France. All samples were loaded into a flow-through quartz capillary using an autosampler robot. The dialysis buffer solution was measured before, and after each sample's acquisition. For all samples, including buffers, ten consecutive frames with an exposure time of 1 s each were recorded.

There are standard analyses for proteins and peptides which are usually performed. These do not require any modeling and provide information about particle shape and size, and are also a way of checking the quality of the data. Interesting parameters obtainable from SAXS includes (but are not limited to) the *radius of gyration*, R_g the *maximal chord length*, D_{max}, and the *molecular weight*, M_w. In addition, information about aggregation is possible, if present.

Kratky plot The scattering profile can be transformed and presented in a *Kratky plot* ($I(q)q^2$ vs q), which is a direct report of the shape and flexibility of the protein chain; if it is compact or unstructured. For a globular, well-folded protein, the Kratky plot is bell-shaped with a well-defined maximum. On the contrary, the Kratky plot for a fully unfolded protein displays a plateau in the high q region, which can be followed by an increase as q increases. This depends on the local rigidity of the chain. However, representing the curves in this way does not distinguish between fully folded and partially folded proteins. It is therefore common to instead present the dimensionless, or normalized Kratky plot, where the intensity, $I(q)$ is normalized with the forward scattering intensity, I_0, while q is normalized by the R_g of the protein. These additions allow to compare proteins of different sizes and molecular weights.[86]

The Gunier approximation One way of obtaining the R_g is through the *Gunier approximation*,[84,86,87] which states that at low values of q, the scattering profile can be

approximated, in log-scale, to

$$\ln I(q) = \ln I_0 - (R_g)^2/3, \tag{4.5}$$

where I_0 is the forward scattering (the scattering signal extrapolated to $q = 0$). For well-folded proteins $\ln I(q)$ is usually linear with respect to q^2 for small q, in the region $qR_g < 1.3$. For IDPs, the region is commonly reduced to $qR_g < 0.8$.[86] If the used q-range is too large, the estimated R_g value usually turns out smaller than the true value. It is also important to note if there are signs of attraction or repulsion in the low q values, as that will compromise the quality of the data. If either is present, detailed analysis with this method should be avoided.

Pair distance distribution function The *pair distance distribution function*, $P(r)$, gives information about the shape, and the intra-particle distances, of the particles in the sample, as it shows the distribution of pair distances within the molecules, in this case, the peptides. It is expressed in real space, as opposed to the scattering pattern which contains information in inverse space.[85] This is achieved by applying a Fourier transformation on the entire scattering profile[88]

$$P(q) = 4\pi \int_0^\infty P(r) \frac{\sin qr}{qr}\, dr. \tag{4.6}$$

From $P(r)$ it is also possible to obtain R_g, as well as the forward scattering,[86,87] using the following equations

$$R_g^2 = \frac{\int P(r) r^2\, dr}{2 \int P(r)\, dr}, \tag{4.7}$$

$$I_0 = 4\pi \int_0^{D_{max}} P(r)\, dr. \tag{4.8}$$

$P(r)$ is assumed to be zero at $r = 0$ and at D_{max}, which is the maximum distance within the peptide chain.

The analyzes performed for the work in this thesis, the PRIMUS software, a part of the ATSAS package, version 3.2.1[89] was utilized.

4.5 Quartz-crystal microbalance with dissipation monitoring

QCM-D is a powerful technique that can be used to monitor adsorption processes and changes in the viscoelastic properties of films at solid interfaces.[90] It is an attractive method to study adsorption processes of biological systems, including protein/peptide or polymer adsorption,[91,92] and lipid bilayer formation,[93] where the interfacial and trapped water is also included.

4.5.1 Basic principle

The technology of QCM exploits the measurement of changes in the vibrational oscillations of a quartz crystal when an alternating potential is applied to it, as illustrated schematically in Figure 4.5. Changes in the natural resonant frequency of the crystal, and its odd overtones, is used to characterize adsorbed layers on the top of the crystal. By measuring several overtones of the frequency and dissipation shifts, the sensitivity of the measurements is increased. The probed depth decreases with increasing overtone number, which means that a higher overtone can be more sensitive to the deposited layer. For the work in this thesis, the 11^{th} overtone was displayed when all overtones were overlapping, which means that all of them provide equal information. In other words, the system behaves the same on the different depths probed by the different overtones. When the overtones did not overlap, all overtones were displayed. The quartz crystals can be coated with different materials, according to the experimental requirements (*e.g.* SiO_2, gold, etc.).

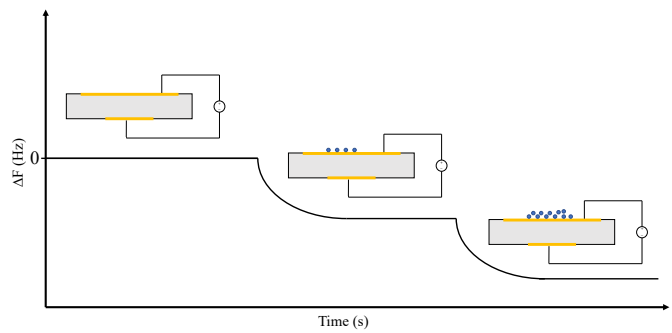


Figure 4.5: A schematic illustration of the QCM principle. A voltage is applied over the quartz crystal (AT-cut) making it oscillate at a resonance frequency. Upon adsorption of additional mass to the crystal, the frequency will decrease, indicated in the illustration by a change in ΔF.

Since the QCM technique was used almost exclusively for gas-phase and vacuum applications at first, additional information of the adsorbed layer was required when mov-

ing to liquid-phase applications. Problems, such as highly hydrated adsorbed layers, or water molecules trapped between adsorbed species, as well as non-rigid adsorbed layers, leads to dampening of the crystal's oscillation due to frictional losses in the deposited layer. To handle this additional complexity, changes in the energy dissipation (denoted as ΔD) is also recorded,[94] adding the "D" in the QCM-D abbreviation. The dissipation is a measurement of the loss of oscillation over a time period, and is defined as

$$D = \frac{E_{\text{dissipated}}}{2\pi E_{\text{stored}}}, \tag{4.9}$$

where E_{stored} is the stored energy in the oscillating system, and $E_{\text{dissipated}}$ is the dissipated energy during one period of oscillation.[94] By using both the ΔF and ΔD values measured, it is possible to obtain both quantitative information on how much is adsorbed and qualitative information such as the rigidity of the adsorbed layer. If the adsorbed layer is not rigid or coupled to the surrounding solution, for example intact lipid vesicles, it will have viscoelastic properties, usually indicated by a high dissipation value. If the dissipation value instead is low, that indicates an adsorption of a rigid layer, for example a lipid bilayer. It is important to note that the shift in both ΔF and ΔD is relative to the baseline, hence, a change in the surrounding environment, that could for example be in temperature, density of liquid, or going from liquid to air, will affect both values.[94]

4.5.2 Measurement setup and data analysis

QCM-D measurements were performed on an E4 apparatus (Biolin Scientific, Sweden), which is equipped with four thermally controlled flow modules. All experiments were performed on SiO_2-coated AT-cut 5 MHz quartz sensors (Biolin Scientific, Sweden), and a peristaltic pump was used to inject sample into the flow modules, and rinse the system with buffer.

Changes in ΔF and ΔD of the oscillation can be related to the mass associated with the deposited material and its viscoelastic properties. The changes in mass per unit area Δm (ng cm^{-2}) is, according to the *Sauerbrey equation*, related to ΔF through[95,96]

$$\frac{\Delta F_n}{n} = -\frac{1}{C_f} \cdot \Delta m, \tag{4.10}$$

where C_f is the mass sensitivity constant, $C_f = 17.7$ ng cm^{-2} Hz^{-1} for an AT-cut quartz crystal with 5 MHz resonant frequency,[97] which was used in the experiments in this thesis. n is the overtone number and ΔF_n is the frequency response at the n^{th}

overtone (Hz). This equation holds true for an even distribution of adsorbed mass, which needs to be rigid and small in comparison to the mass of the crystal (or small frequency change in comparison with the resonance frequency). When adding a mass to the crystal surface, a decrease in resonant frequency is observed ($\Delta F < 0$), whereas an increase in the resonant frequency ($\Delta F > 0$) generally indicates that something has desorbed from the surface, hence, a decrease of mass from the crystal.

In Paper III of this thesis, we wanted to effectively compare results obtained from QCM-D and neutron reflectometry (the technique described in the next section). To do this, we converted the adsorbed mass to an equivalent thickness through

$$t^{QCM} = \frac{\Delta m}{\rho_m}, \tag{4.11}$$

where ρ_m is the mass density of the peptide obtained from the ratio of molecular mass, M_m, and molecular volume, M_v, of the peptide species under investigation. This can only be used when Equation 4.10 holds.

4.6 Neutron reflectometry

NR is a technique to measure thickness and composition. It is a low resolution technique compared to other scattering techniques such as X-ray reflectivity or diffraction. Due to the high penetration capacity of neutrons toward many materials, it is possible to probe buried interfaces using NR. It is being used to investigate a wide variety of topics, including surface chemistry, surface magnetism, and solid films.[98]

The phenomenon of light reflecting from surfaces is well-known, and is caused by the change of refractive index across interfaces. Newton realized already in 1675 that one could measure the thickness of a thin film by illuminating it with a parallel beam of white light, and observing the color of the reflected light. X-ray reflection follows the same laws as light, however, with different refractive indices depending on the number of electrons per unit volume. This was shown in experiments by Compton in the 1920's.[98] The first mirror reflection of neutrons was demonstrated in 1944 by Fermi and Zinn[99] to which the same fundamental equations can be applied as optical reflectivity but again with different refractive indices.

4.6.1 Basic principle

Neutrons are scattered by the nucleus of atoms. The quantity describing how strong a neutron interacts with a nucleus is the *coherent scattering length*. The scattering length

of neutrons is seemingly random across the periodic table, as opposed to the interaction between X-rays and matter, where the increase in scattering length is linear with atomic number. This is useful in biological systems which are rich in light elements such as hydrogen (H), which has the lowest electron density of all elements, and is therefore not easily detected by X-rays. The possibility to use isotropic substitution, e.g. ^{1}H to ^{2}H (denoted D for deuterium) can increase the coherent scattering of the biological materials.[100]

The *scatting length density*, ρ (or SLD), is defined as

$$\rho = \frac{\sum_{j=1}^{N} b_j}{V_m}, \tag{4.12}$$

where b_j is the coherent scattering length of a nucleus, and V_m is the volume containing the N nuclei considered. Scattering lengths can have different signs, so can the SLD, and moreover, different scattering lengths can be combined, in the same volume V_m, to tune the value of the SLD. This feature is the base of the *contrast variation method*.[101] The most notable case for biological samples is that of H and D; where b_j is negative for H, and positive for D. By substituting some or all of the H atoms naturally occurring in a biomolecule by D atoms, the biomolecule's SLD can be tuned. The same method can be applied to the solvent, for which the use of different ratios of H_2O and D_2O can be used to highlight or hide different parts of the sample, where the latter is illustrated schematically in Figure 4.6.

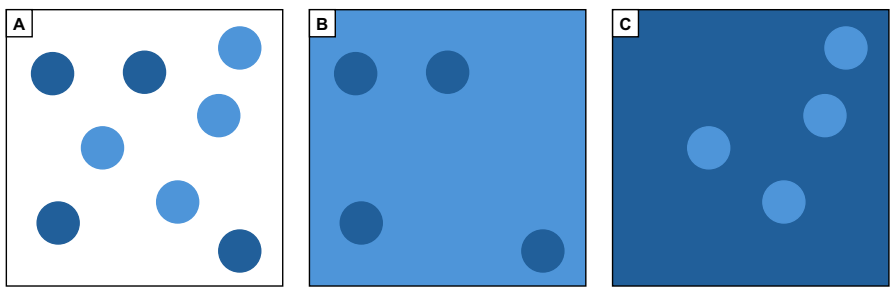


Figure 4.6: A schematic figure illustrating contrast variation, where the SLD of the solvent matches **(A)** neither of the particles, **(B)** the light blue particles, and **(C)** the dark blue particles. In contrast match condition, the SLD of the solution and that of part of the sample equals and the contrast term is $\Delta\rho = 0$.

Neutron reflection can be either specular or nonspecular, where the former provides information about the sample structure perpendicular to the interface, while the latter provides information on the lateral structure (such as in-plane correlations) of the interface.[100] In this thesis, only specular reflectivity will be considered. The following

basics of NR are mainly based on concepts presented in the work by Fragneto,[100] and Cubitt and Fragneto[98], which provides a more in-depth description for the interested reader.

Neutron reflection theory can be approached using quantum mechanics, where a neutron is treated as a wave. The wavelength, λ of a neutron with mass m_n is given by the *de Broglie equation*

$$\lambda = \frac{h}{m_n v}, \tag{4.13}$$

where h is Planck's constant and v is the velocity of the neutrons. Once a neutron crosses a medium, it is subject to an average interaction potential, that in the continuum limit is described as

$$V = \frac{h^2}{2\pi m_n}\rho. \tag{4.14}$$

ρ denotes the previously defined scattering length density. In Equation 4.12, it replaces the coherent neutron scattering length, which is originally present in the interaction potential to describe the individual neutron-nucleus interaction. Since specular NR probes the structure of the sample along the perpendicular direction with respect to the interface, the information in the reciprocal space is expressed as a function of the momentum transfer perpendicular to the interface, denoted q_z, is given by

$$q_z = 2k_i = \frac{4\pi}{\lambda}\sin\theta, \tag{4.15}$$

where λ is the wavelength of the incoming beam, and θ is the angle of the beam with the interface. This is schematically illustrated in Figure 4.7.

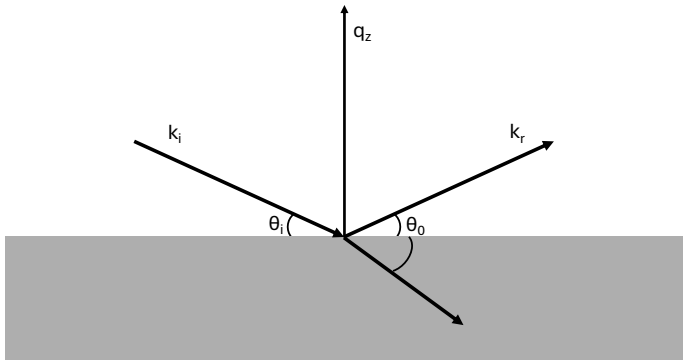


Figure 4.7: A schematic figure displaying the reflection from an (ideally flat) interface. k_i and k_r are the incident and scattered wave vectors. $\theta_i = \theta_0 = \theta$ are the angles of the wave vectors in the incident plane, and q_z is the momentum transfer.

Reflectivity for an ideal and perfectly flat surface can be approximated, at large values of q, by

$$R_k \approx \frac{16\pi^2}{q_z^4}\Delta\rho^2, \tag{4.16}$$

which shows that the reflectivity above the critical angle decreases sharply with q^4. If a uniform layer is added on top of the substrate, the reflectivity is instead given by

$$R_k \approx \frac{16\pi^2}{q_z^4}[(\rho_1-\rho_0)^2+(\rho_2-\rho_1)^2+2(\rho_1-\rho_0)(\rho_2-\rho_1)\cos q_z d], \tag{4.17}$$

where d is the thickness of the layer, which is directly related to the change in q as $d \propto \frac{2\pi}{\Delta q}$. ρ_0, ρ_1, and ρ_2 are the scattering length densities of the two bulk phases (ρ_0 and ρ_2) and the layer (ρ_1). The reflection is schematically displayed in Figure 4.8. This calculation can be performed for any number of layers, but becomes cumbersome for larger number of layers, where other methods have instead been established, e.g. Parratt's recursive algorithm.[102]

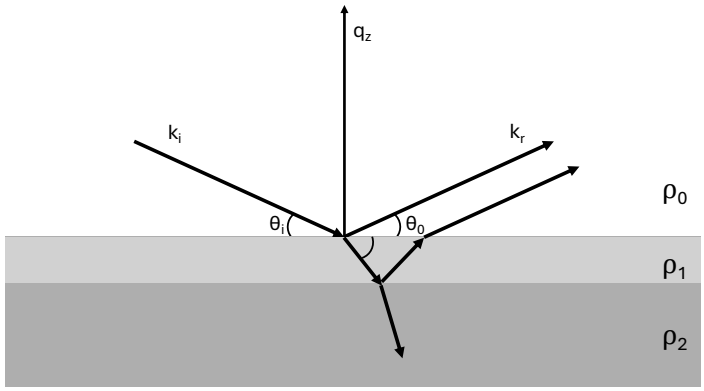


Figure 4.8: A schematic figure displaying the reflection from one layer on top of a substrate. A sequence of several reflection and refraction events take place every time the neutron beam crosses an interface and until the neutron wave is not distinguishable. These events are not indicated for clarity.

4.6.2 Measurement setup

The basic setup of a reflection experiment contains the following components in the given order: a radiation source, followed by either choppers (for time of flight mode) or a monochromator (for monochromatic mode), a collimation system, the sample, and finally a detector. Neutrons can be produced in two ways, either by a nuclear reactor (fission), which produces a continuous neutron beam, or though spallation, which produces a pulsed neutron beam. The experiments in this thesis were performed at four neutron reflectometers: D17, FIGARO and SuperADAM at ILL (Grenoble, France), and INTER at ISIS (Didcot, UK), where the measurements at INTER were performed by other members of the team before I joined the project. The focus will therefore be on the reflectometers at ILL.

In a NR experiment, the reflectivity is measured as a function of the wave vector perpendicular to the reflecting surface, q_z. It follows from the definition of q_z in Equation 4.15 that measurements can be performed either in monochromatic mode, where the glancing incident angle is varied and the intensity of reflection is recorded for each incident angle for a single wavelength, or the time of flight (TOF) mode, where the angle is fixed, while several wavelengths are measured simultaneously, probing several values of q_z. The latter configuration can also be used to probe kinetics, due to the fact that a large range of q_z can be measured at a short timescale (in the order of minutes), for the samples used in this thesis. SuperADAM is operated as a monochromatic reflectometer, while both D17 and FIGARO are TOF reflectometers. According to Bragg's law, Equation 4.15, neutrons with different wavelengths leads to a range in q_z like the variation in incident angle. Because of this, a couple (1-3) angles are often used and

combined to obtain a larger range in q_z-space.

Because of incoherent scattering due to any material the neutron encounters while passing through the sample, including the sample itself, the sample holder, the substrates, as well as the water phase, a q_z-independent background is commonly present. At large q_z, reflectivity becomes very small (see Equation 4.17) for very thin films as those investigated in the thesis, leading to a deterioration of the signal to noise ratio. This background needs to be subtracted from the measured reflectivity, and it also marks a lower limit of reflectivity which can be measured.

All experiments were performed using solid/liquid sample cells, which were developed at ILL. In the setup the silicon block ($8 \times 5 \times 1.5$ cm^3, cut along the III plane, polished with 3 Å root-mean-square (RMS) roughness, purchased from Sil'tronix ST, Archamps, France) is sandwiched between an aluminum plate and a PEEK (polyether ether ketone) trough, which contains the liquid solution in contact with the sample formed on top of the surface of the solid substrate. To ensure that the neutron beam reaches the detector, the incoming beam is transmitted through the most transparent material, which, in the case of the setup used in this thesis, is the solid silicon substrate, rather than the aqueous solution. The silica shows a low adsorption and incoherent scattering cross section. The sample cell holds approximately 2 mL of liquid and is connected to tubes for solvent inlet and outlet. An HPLC pump controlled by a computer is used for solvent exchange, however, all samples are added manually to the sample cell by a syringe. The temperature of the cell is controlled by a water circuit in the cell holder. The neutron beam penetrated through the silicon block and is reflected at the inner side of the silicon on which the bilayer-peptide system is formed.

To normalize the reflected intensity to the incoming intensity, the direct beam through the silicon block was measured. Before exposed to any sample, the silicon block was characterized in two contrasts to make sure it was properly cleaned, as well as to determine the properties of the tentatively occurring oxide layer on top of the silicon. This was followed by characterizing the pristine lipid bilayer, formed by vesicle fusion, in two contrasts, H_2O buffer and D_2O buffer. The peptides were then injected in the sample cell, and after incubation for 60 minutes the system was measured using three or four different contrasts to employ contrast variation. Again, the H_2O and D_2O buffers were used together with two mixtures equal to the SLD of silicon (62% H_2O, 38% D_2O, referred to as SiMB) and an SLD of 4 (34% H_2O, 66% D_2O, referred to as 4MB). Counting times were adjusted for each contrast according to the sample and to the intensity of the direct beam, to allow for satisfactory statistics for the NR curves.

4.6.3 Data analysis and modelling

The experimental data obtained at D17 and FIGARO was converted from RAW format reflectivity curves ($R(q)$) using the COSMOS routine, which is accessible from the LAMP software.[103] The background is subtracted. To combine the data measured at different angles to one reflectivity curve, the region where the two data sets overlap is determined, and a scaling factor is obtained and applied to ensure that the data sets from the different incident angles overlap. Data obtained from the SuperADAM reflectometer was reduced using the pySAred software, which was provided at the instrument.

The data was analyzed using the Aurore software application,[104] and the analysis was performed in a stepwise manner. A crude overview of the steps is outlined below, which are all described more in detail in the text following.

NR analysis overview

1. Fit data for bare substrate.
2. Transfer the parameters to the following steps as constants.
3. Fit data for pristine bilayer before injection of peptide using lipid plugin.
4. Transfer bilayer parameters as starting point for next step.
5. Evaluate different scenarios to find the most appropriate model for the system after peptide injection and incubation.

All systems were modeled as a series of homogeneous layers. Each layer is described in terms of SLD, thickness, t, buffer volume fraction, f_w, and interfacial roughness, σ. The steps described in the NR analysis overview in detail are as follows. (1) First the bare silicon substrate was characterized. The data was fitted to a model consisting of an infinite layer with the SLD of silicon, by an oxide layer, and by an infinite bulk aqueous layer. (2) Secondly, the parameters obtained from this step were then transferred to the following analyses of the sample deposited on that particular substrate. (3) The third step comprises the analysis of the data obtained for the pristine bilayer before injection of peptide. This step was done using the *lipid plugin* provided by the Aurore software. The bilayer is then modeled by three layers, or 12 parameters, for a symmetric bilayer where the inner and outer leaflets of the bilayer is coupled, or by four layers, that is 16 parameters, for an asymmetric bilayer where the leaflets are instead decoupled.

The plugin requires input of the scattering length, b_m, of the head groups and the tails

respectively, as well as the dry molecular volume, M_v. With these inputs the SLD is calculated for each layer by dividing b_m by M_v. If the molecule contains exchangeable protons, the b_m value is affected and should be determined for each contrast measured during the data acquisition. For the lipids included in this work, it is only the head group of POPS that contains exchangeable protons. The scattering length values for each isotope were obtained from the NIST webpage "Neutron scattering lengths and cross sections"[105], and the split between head group and tails was done as shown in Figure 4.9 for both POPC and POPS. To calculate the M_v, volumes from component groups obtained from the paper by Nagle and Tristram-Nagle,[106] were used to calculate the the respective M_v's required for input.

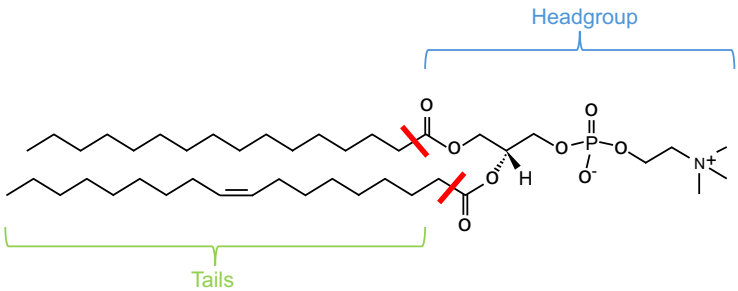


Figure 4.9: A schematic figure of a POPC molecule where the red lines marks the breakpoint between the part which is considered head group and the part considered as tails in the determination of the scattering length and the molecular volume. A similar corresponding breakpoint applies for POPS, which closely resembles POPC.

(4) For the fourth step, the parameters obtained from the third were used as the starting point in the process of analyzing the data obtained after peptide injection. (5) For this part of the analysis, different physically relevant scenarios were tested and evaluated to find the set of parameters best matching the experimental data set. The peptides included in this work do contain exchangeable hydrogen atoms, and in Paper I, an equation was derived describing the total SLD of a hydrated layer containing Hst5 using the slightly different SLD values in each solvent. In the Paper II, III and the ongoing study, the so-called SLD^0 value was used, which is the SLD value at SLD_{solv} = 0. The SLD values of all components used in this thesis is presented in Table 4.1.

Table 4.1: SLD values of the materials used in this thesis.

Material/Component	SLD [10^{-6} Å^{-2}]
Substrates	
Si	2.07
SiO_2	3.41
Solvents	
Hbuff (H_2O)	-0.56
Dbuff (D_2O)	6.35
SiMB	2.07
4MB	4.0
Lipids	
POPC HG	1.87
POPC tails	-0.28
PC_9PS_1 HG (H_2O)	1.98
PC_9PS_1 HG (D_2O)	2.07
PC_9PS_1 tails	-0.28
Peptides	
Hst5 and variants	2.40
KEIF	2.98

The Aurore software is using Parratt's recursive formalism[102] in the conversion from the modelled $\rho(z)$ profile to the corresponding reflectivity curve. The modelling of a SLD profile is done by concatenation of slices of finite size, where each slice is defined by three parameters: the thickness, t_i, the total SLD, ρ_i, and the interfacial roughness, $\sigma_{i,i+1}$, where i is the slice index. To find the best fit between the model and the experimental data set a least square function, χ^2 needs to be minimized.[104] This is performed in the Aurore software by implementation of the *MINUIT* software.[107]

Roses are red // Violets are blue // Unexpected '}' on line 32.
- Anonymous

5 | Computational methods

Contents

In this chapter, the models and methods used for the computational part of this thesis are described.

5.1 Simulation models

Simulation models are used as a representation of reality, and can contain different levels of detail. When a model is constructed or selected, enough detail has to be included, to accurately describe the system properties of interest. More details are not always preferable, since it can make the results harder to interpret and it also increases the computational cost, which can limit the system size and the time scales investigated. Therefore, one has to assess which model and which level of detail is needed for the

scientific question proposed. In this thesis, the interaction between an IDP and a lipid bilayer has been studied by two different models, a coarse-grained one, where several atoms have been grouped together into a spherical particle, and an atomistic one, where all atoms in the system are included, see Figure 5.1.

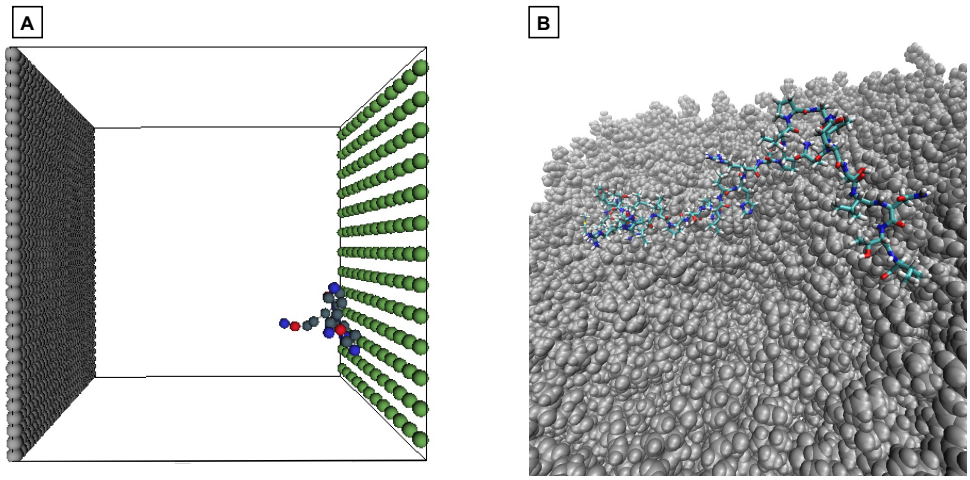


Figure 5.1: The two different models with different level of detail depicted. **(A)** The coarse-grained model of Hst5 within the cushion where the 990 gray particles makes up the silica surface, where each bead has a charge of -0.05 e, and the 156 green particles makes up the head groups of the bilayer, where each particle is given a charge of -0.5 e. In the bonded particles making up the peptide, negatively charged residues are represented by red spheres, positively charged one by blue spheres, and uncharged residues by the gray spheres. **(B)** The atomistic model of KEIF and a PC_9PS_1 bilayer where the bilayer for clarity is represented in gray for all atoms. The peptide atoms are colored so that carbons are represented in cyan, nitrogen in blue, oxygen in red, hydrogen in white, and sulfur in yellow.

5.1.1 The coarse-grained model

The coarse-grained model is a bead-necklace model where the amino acids are represented by hard spheres connected by harmonic bonds. Both termini are defined as additional residues to account for the extra charge they give rise to. The beads can be negative, positive, or uncharged, depending on the amino acid sequence at pH 7.4. The model was parameterized by Cragnell *et al.* for Hst5.[108] In the ongoing study presented in this thesis, the charge of the histidine residues are altered from 0 to 0.5 to 1 e, to investigate the effect of charge regulation as the peptide approaches a surface. This charge is constant over the entirety of the simulation. The total energy of the system has four contributions, where one only applies to the bonded beads in the peptide chain. It is

called the bond energy, denoted as U_{bond}, and is given by

$$U_{\mathrm{bond}} = \sum_{i=1}^{N_{seg}-1} \frac{k_{\mathrm{bond}}}{2}(r_{i,i+1} - r_0)^2, \tag{5.1}$$

where N_{seg} is the number of segments, referred to as beads, in the chain, $r_{i,i+1}$ is the centre-to-centre distance between two connected beads with the equilibrium distance $r_0 = 4.1$ Å, and $k_{\mathrm{bond}} = 0.4\ \mathrm{Nm}^{-1}$ is the force constant.

In addition to the peptide chain, the simulation model is set up to include either one, or two surfaces representing the head groups of a lipid bilayer or a solid silica surface. Both surfaces are represented by hard spheres distributed on a primitive cubic lattice, where the particles are frozen in their initial position, which is an approximation of the real system. The surface representing the head groups is built up of 156 particles, each given a charge of -0.5 e. 990 particles comprise the silica surface, where each particle has a charge of -0.05 e. This is the model which was set up for and used in Paper III. It was also used in the ongoing study included in this thesis.

The counterions are treated explicitly, whereas the salt is treated implicitly using the Debye–Hückel theory, in which the solvent, in this case water, is treated as a dielectric continuum. Each particle in the simulation has a radius of 2 Å. All non-bonded interactions are assumed to be pairwise additive. In addition to the bond energy, U_{bond}, there are three contributions to the total energy, the hard sphere potential, U_{hs}, the electrostatic potential, U_{el}, and a short ranged van der Waals interaction, denoted U_{short},

$$U_{\mathrm{tot}} = U_{\mathrm{hs}} + U_{\mathrm{el}} + U_{\mathrm{short}} + U_{\mathrm{bond}}. \tag{5.2}$$

The hard sphere potential, U_{hs}, is given by

$$U_{\mathrm{hs}} = \sum_{i<j} u_{ij}^{\mathrm{hs}}(r_{ij}), \tag{5.3}$$

in which all the particles in the system are included. r_{ij} is the center-to-center distance between particle i and particle j. The hard sphere potential between two particles in the model is given by

$$u_{ij}^{hs} = \begin{cases} 0, & r_{ij} \geq R_i + R_j, \\ \infty & r_{ij} < R_i + R_j, \end{cases}$$

where the radius of particle i and j is given by R_i and R_j, respectively.

The electrostatic potential, U_{el}, is given by an extended Debye–Hückel potential

$$U_{\mathrm{el}} = \sum_{ij} u_{ij}^{\mathrm{el}}(r_{ij}) = \sum_{i<j} \frac{Z_i Z_j e^2}{4\varepsilon_0 \varepsilon_r} \frac{\exp[-\kappa(r_{ij} - (R_i + R_j))]}{(1 + \kappa R_i)(1 + \kappa R_j)} \frac{1}{r_{ij}}, \tag{5.4}$$

where all the particles in the system are included in the sum. Z_i is the valency of particle i, e is the elementary charge, ε_0 is the permittivity of vacuum, ε_r is the dielectric permittivity for water, and κ denotes the inverse *Debye screening length*.

A short-ranged attractive interaction contributing to the total potential energy corresponds to the van der Waals interaction. It is given by

$$U_{\mathrm{short}} = -\sum_{i<j} \frac{\varepsilon}{r_{ij}^6}, \tag{5.5}$$

where ε determines the interaction strength. The attraction is defined to act between all beads in the chain, and in this study, an attractive potential of 0.6 kT at closest contact was used.

5.1.2 The atomistic model

For the atomistic model used in Paper IV, KEIF was first built as a linear chain in Avogadro, version 1.2.0, an open-source molecular builder and visualization tool.[109] The peptide was then placed 25 Å above a lipid bilayer, consisting of 306 POPC molecules and 34 POPS molecules per leaflet using CHARMM-GUI.[110–115] All atoms of both the peptide and the lipids were included. Solvent molecules and ions were then added to the model, which, were all treated explicitly. The CHARMM36m force field[116] was used and the model was adjusted to run in the GROMACS package.[117–120] The same model was also used for Hst5 and the random sequence variants presented as the ongoing study in this thesis.

The total potential energy depends both on bonded, that is *intramolecular* interactions, and non-bonded, *intermolecular* interactions, according to

$$U_{\mathrm{tot}} = \underbrace{U_{\mathrm{bond}} + U_{\mathrm{angle}} + U_{\mathrm{d}} + U_{\mathrm{id}}}_{\text{bonded}} + \underbrace{U_{\mathrm{LJ}} + U_{\mathrm{el}}}_{\text{non-bonded}}. \tag{5.6}$$

All of the interaction potentials are summed over all atoms involved in the interaction. The bonded potentials apply to covalently bonded atoms, and the first one is a harmonic potential for bond stretching,

$$U_{\text{bond}} = \sum_{b} \frac{1}{2} k_{ij}^{\text{b}} \left(r_{ij} - r_{ij}^{0} \right)^2 , \tag{5.7}$$

where k_{ij}^{b} is a force constant, r_{ij} is the distance between the bonded atoms i and j, which have an equilibrium bond length of r_{ij}^{0}. The bond angle vibration is a three-body harmonic potential between the atoms, which is the second term in the total potential.

$$U_{\text{angle}} = \sum_{\theta} \frac{1}{2} k_{ij}^{\theta} \left(\theta_{ijk} - \theta_{ijk}^{0} \right)^2 , \tag{5.8}$$

where k_{ij}^{θ} is a force constant, θ_{ijk} is the angle between the atoms $i - j - k$, with the equilibrium angle θ_{ijk}^{0}. The two remaining bonded potentials are torsion potentials with dihedral angles, which is an angle between two intersecting planes that controls the rotation of a bond around its longitudinal axis. The proper dihedral angle is defined according to the IUPAC/IUB convention,[121] as the angle ϕ_{ijkl} between the ijk and jkl planes. The proper dihedral potential displays periodic behavior with the periodicity n and the phase ϕ_s

$$U_{\text{d}} = \sum_{\phi} k_{\phi} \left[1 + \cos \left(n\phi_{ijkl} - \phi_s \right) \right] , \tag{5.9}$$

in which k_{ϕ} is a force constant. In defining an improper dihedral, unlike proper dihedrals, the atoms do not need to be linearly attached. It is based on an angle, ξ_{ijkl} between two planes, and it is harmonic in nature, according to

$$U_{\text{id}} = \sum_{\xi} \frac{1}{2} k_{\xi} \left(\xi_{ijkl} - \xi_l \right)^2 , \tag{5.10}$$

where k_{ξ} is the force constant. The bonded interactions are illustrated in Figure 5.2.

There are two non-bonded potentials contributing to the total potential, namely a Lennard-Jones interaction and a Coulomb interaction. They are both pairwise ad-

ditive, and the Lennard-Jones interaction contains both an attractive dispersion term, and a steric repulsion term, according to

$$U_{\mathrm{LJ}} = \sum_{i<j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right]. \tag{5.11}$$

The potential well depth is denoted by ε_{ij}, while σ_{ij} is the distance where the potential becomes zero. The electrostatic interactions are represented by the Coulomb potetial, which acts between two particles i and j, with charges q_i and q_j, respectively,

$$U_{\mathrm{el}} = \sum_{i<j} \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}}. \tag{5.12}$$

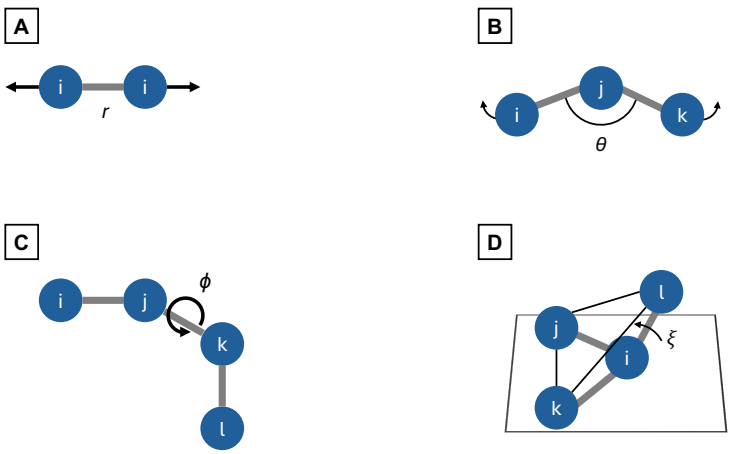


Figure 5.2: Schematic representation of the bonded interactions in the atomistic model, **(A)** bond stretching, **(B)** bond angle vibration, **(C)** proper dihedral torsion, **(D)** improper dihedral torsion.

5.2 Simulation methods

Computer simulations can act as a bridge between theory and experiments, and provides information on the microscopic level, while experiments in many cases gives information on the macroscopic level.

In the work for this thesis, two different simulation methods were used, *Monte Carlo* (MC), which was used to simulate the coarse-grained model, and *molecular dynamics* (MD) to simulate the atomistic model. The main difference between these two methods

is that MC calculates ensemble averages based on random sampling, while MD is based on Newton's equations of motion, which gives time averages. The first postulate of statistical mechanics, stated earlier in Chapter 3, tells us that the results of these two methods are the same, if the ensembles are large enough and the time samples are sufficiently long.

5.2.1 Metropolis Monte Carlo simulations

One numerical technique to evaluate Equation 3.8 is, as stated above, using the MC method. In the work performed for this thesis, the sampling scheme presented by Metropolis *et al.* [122] has been used. The MC method is based on random displacements, which is preluded by a trial move to evaluate if the move is viable. This is done by determining the energy of the new configuration and then compare it with the old one. If the new energy is lower than the old, the move is always accepted, however, configurations with a higher energy are also accepted, with a probability determined by a Boltzmann distribution. A basic Metropolis MC algorithm is outlined below. [123]

Metropolis Monte Carlo algoritm

1. Generate a random starting configuration.
2. Select a random particle and calculate its energy, U_{old}.
3. Perform a trial move to displace the particle.
4. Calculate the energy of the new configuration, U_{new}.
5. Compare the energy of the old and new configuration. The move is accepted with the probability

$$p_{acc} = \begin{cases} 1 & \text{if } U_{new} \leq U_{old} \\ \exp\left[-\frac{1}{kT}\left(U_{new} - U_{old}\right)\right] & \text{if } U_{new} > U_{old} \end{cases}.$$

6. If the new configuration is rejected, restore the old one.
7. Repeat from step 2.

It is possible to include a large variety of moves in the MC algorithm, where some of them are unphysical, as that will speed up the exploration of the configurational space. All the included moves are given a probability to occur. In this work, four different

moves were included, which will be described below.

Translational displacement of a single particle is, as the name suggests, a move of a single particle in the system. This can be either a bead connected in a chain (see Figure 5.3A), or a particle representing an ion, or, if explicit water molecules are included, translation of one such particle. The length of the displacement is limited by an input parameter that is defined in the simulation. This is the only move applied to single particles. What follows applies only to the bonded beads of the peptide chain.

Translational displacement of a chain is a movement where the entire chain is moved to a new position in the system, while the conformation of the chain is preserved, see Figure 5.3B.

Slithering move is when one bead is randomly displaced within a bond length, and the other beads are moved forward in the chain, as is shown in Figure 5.3C.

Pivot rotation is when one end of the chain is rotated around an axis, which is defined by a randomly selected bond. This move is displayed in Figure 5.3D.

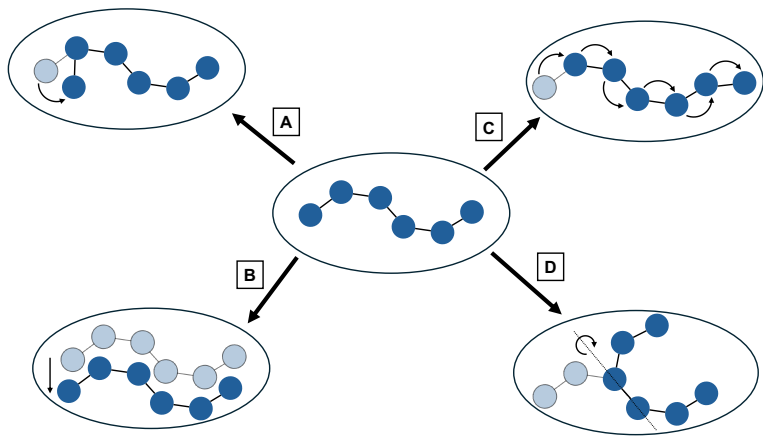


Figure 5.3: Schematic illustration of the trial moves included in the Monte Carlo simulations, **(A)** translational displacement of a single particle, **(B)** translational displacement of a chain, **(C)** slithering move, and **(D)** pivot rotation.

5.2.2 Molecular dynamics simulations

In MD simulations, contrary to MC, dynamical information about the system can be obtained, since *Newton's equations of motion* are used to move the particles. Newton's second law of motion states that the force, $\mathbf{F}_i$ of a particle i, with a mass, m_i, is proportional to the acceleration, $\mathbf{a}_i$. This can be expressed as the second derivative of the

position $\mathbf{r}_i$ with respect to time, t,

$$\mathbf{F}_i = m_i \cdot \mathbf{a}_i = m_i \cdot \frac{\mathrm{d}^2 \mathbf{r}_i}{\mathrm{d}t^2}. \tag{5.13}$$

By numerically solving Equation 5.13, the movements of the atoms are simulated. Starting velocities, positions, and interaction potentials of the atoms in the system are required to run a MD simulation. The forces are computed using

$$\mathbf{F}_i = -\frac{\mathrm{d}U(\mathbf{r}^N)}{\mathrm{d}\mathbf{r}_i}, \tag{5.14}$$

where $U(\mathbf{r}^N)$ is the potential, and $\mathbf{r}^N$ represents the complete set of atomic coordinates.

There are different integration methods available for MD simulations. In this work an integrator called the *leap-frog algorithm* is used,[124,125] where the positions, $\mathbf{r}_i$, and the velocities, $\mathbf{v}_i$, are calculated alternately in time. This is illustrated in Figure 5.4, using the following equations

$$\mathbf{v}_i\left(t + \frac{1}{2}\Delta t\right) = \mathbf{v}_i\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m_i}\mathbf{F}_i(t), \tag{5.15}$$

$$\mathbf{r}_i\left(t + \Delta t\right) = \mathbf{r}_i\left(t\right) + \Delta t \cdot \mathbf{v}_i\left(t + \frac{1}{2}\Delta t\right). \tag{5.16}$$

The energy conservation law will apply if the equations of motions are solved correctly in the simulations, and the equations are reversible in time.

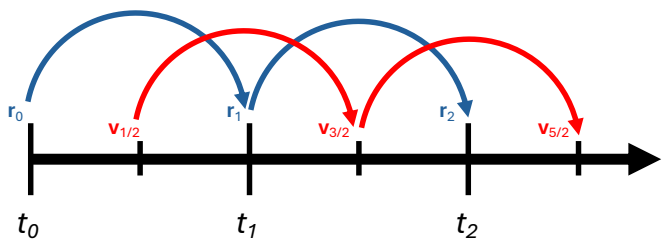


Figure 5.4: A schematic illustration of the leap-frog algorithm, where the calculations of the cooridnates (blue), and the velocities (red) are leaping past each other.

By repeatedly calculating the velocities, positions, and forces, a trajectory of the changes of the positions and velocities in time is created. A basic MD algorithm is summarized below.[125]

Molecular dynamics algoritm

1. Generate the starting configuration including the positions, velocities, and potential interactions of all atoms in the system.
2. Compute forces.
3. Update configuration by solving Newton's equations of motion.
4. Write output.
5. Repeat from step 2.

There are some additional aspects of MD simulations to consider in order to make sure that the simulations are effective and a good representation of a real system, which will be briefly discussed in the following sections.

Periodic boundary conditions. For a simulated system to be a relevant representation of a real system, a large amount of particles would have to be simulated. This would be extremely expensive (in the currency of computer resources) and time consuming. To overcome this problem, the required large system is approximated by a small model system, a unit cell, where *periodic boundary conditions* (PBC) are applied. PBC means that the unit cell is replicated in all directions, creating a periodic, infinite lattice throughout space,[126,127] see Figure 5.5. If a particle hits a wall of the simulation box it will "leave" the box and re-enter on the opposite side when applying PBC. In the simulations performed for this thesis, PBC were applied in all three directions (x, y, and z) in the MD simulations, and the MC simulations with no surface present. In the MC surface simulations PBC where only applied in x and y. z is the axis perpendicular to the surface.

The infinite lattice of replicates of the simulation box obtained with PBC would result in an infinite sum of interactions when all interactions within the system are considered, since the system would consist of an infinite number of particles. To avoid duplicate interactions, the *minimum image convention* is applied, which limits the interaction so that each particle only interacts with the closest image of the other particles.[128] To make simulations cost-effective, you usually only consider short-ranged interactions within a cutoff, usually spherical.

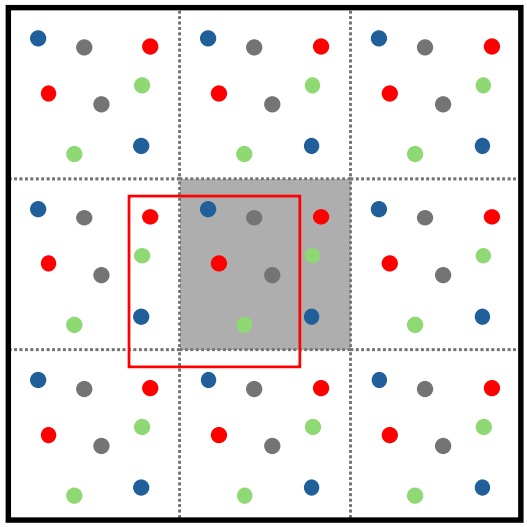


Figure 5.5: An illustration of a system with periodic boundary conditions in two dimensions, where the grey box is replicated in all directions. The minimum image convention is illustrated by the red square.

Long-range interaction corrections. As described in the previous part, cutoffs are implemented to truncate a large amount of interactions in the infinite PBC system, however, these cutoffs might also affect long-range interactions in the system. This is usually done by the *particle-mesh Ewald* (PME) method,[129] an improved version of Ewald summation, where the long-range interaction is separated into two rapidly converging parts, a short-ranged part treated as a direct sum, and a long-ranged part treated as a summation in reciprocal space. The Ewald summation on its own is unfortunately not suitable for large systems, as the computational cost scales with N^2. The PME method utilizes interpolation and fast Fourier transforms, which reduces the order to $N \cdot \ln N$, making it substantially faster and more suitable for larger systems.

Bond constraints. Another way to manage the computational cost of MD simulations is to control the length of the time step, where a longer one would decrease the cost. The time scale of the highest frequency motion in the system constrains the size of the time step, which limits it to around 1 fs. Using a longer time step might make the simulation unstable.[130] Simulations of biomolecules usually require simulation times in the order of μs−ms, which would have a very high computational cost if not treated in any way. Therefore, the bonds in the system can be constrained, by using, as in the work for this thesis, the *LINCS* algorithm,[131] which enables an increase in the length of the time step, in this thesis the time step used for MD has been 2 fs.

Temperature and pressure control. In order to relate the simulations to experiments, the MD simulations were performed in the isothermal-isobaric ensemble, where, like

for many experiments, the temperature and pressure is constant, through applying temperature and pressure couplings. There are many different options available for MD simulations. In this work the temperature was controlled using the *Nose-Hoover temperature coupling*,[132,133] and the pressure using the *Parrinello-Rahman pressure coupling*.[134] The temperature coupling changes the particles equation of motion as presented in Equation 5.13 to[125]

$$\frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2} = \frac{\mathbf{F}_i}{m_i} - \frac{p_\xi}{Q}\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t}, \tag{5.17}$$

where the equation of motion for the heat bath parameter ξ is

$$\frac{\mathrm{d}p_\xi}{\mathrm{d}t} = (T - T_0)\, N_f k. \tag{5.18}$$

T denotes the instantaneous temperature of the system, T_0 is the reference temperature, and N_f is the number of degrees of freedom. When using the Parrinello-Rahman barostat, the changes in pressure are represented by additional terms in the equations of motion containing the simulation box vectors.

5.3 Simulation analyses

To get characteristics from the simulated system, as well as information to compare with experimental results, different analyses were performed on the system, with the most important ones described below.

5.3.1 Size and shape

R_g is an average measure of the compactness of a structure, defined as the root mean square distance of the mass elements and the center of mass[135] given by

$$R_g = \sqrt{\frac{\sum_{i=1}^{n} m_i \|\mathbf{r}_i\|^2}{\sum_{i=1}^{n} m_i}}, \tag{5.19}$$

where m_i is the mass of each element, n is the total number of elements, and $\mathbf{r}_i$ is the position of the mass relative to the center of mass. In the atomistic simulations the elements are atoms, while in the coarse-grained simulations they are beads. In the latter,

each bead has equal mass. The *end-to-end distance*, R_{ee}, which is the distance between the N- and C-terminus. It is calculated as the root mean square of these positions according to

$$R_{ee} = \sqrt{\|\mathbf{r}_1 - \mathbf{r}_n\|^2}, \tag{5.20}$$

where $\mathbf{r}_1$ and $\mathbf{r}_n$ is the position of the first and last element, respectively.

5.3.2 Secondary structure

The coarse-grained model used for the work included in this thesis is too crude to obtain any secondary structure information of the IDPs. It is, however, accessible in the atomistic simulations, and has been determined using the DSSP algorithm implemented in GROMACS.[125,136] This algorithm determines the secondary structure elements present by detecting specific hydrogen bonding patterns between amino acid residues. There are ten structures recognized by the algorithm: (i) lack of structure, (ii) loops or irregular, (iii) turns, (iv) bends, (v)-(viii) four types of helices (including polyproline II helix, which is common in IDPs), and (ix)-(x) two types of β-structures. These results can be compared with those obtained from analysis of experimental CD results, and provide information on a more detailed level on the position of the different secondary structure elements.

5.3.3 Adsorption probability and distances

In paper III the adsorption of Hst5 (and variants thereof) was evaluated using, among other techniques, MC simulations. In the analysis of these simulations, the adsorption probability was calculated, where a bead was considered to be adsorbed if the center-to-center distance between the bead and the surface was less than a certain cutoff. From this, the adsorption probability of a peptide bead is obtained by dividing the number of steps where the bead is considered adsorbed with the total number of steps in the simulation.

In paper IV and the ongoing study the adsorption of peptides to lipid bilayers was investigated with MD simulations. This was done first by determining the minimum distance between each residue and the phosphorus atom in the lipid head group for each time step of the simulation. These minimum distances were then averaged over the whole simulation, to obtain the average minimum distance of each amino acid to the bilayer. The standard deviation was also obtained, and used as an indication of the fluctuation of the peptide chain in the vicinity of the bilayer.

5.3.4 Density profiles

Information about the density of different species was obtained from all the simulations in Paper III, Paper IV, and the ongoing study. Both in MC and MD simulations, the number density of the component of interest was calculated in the z-direction. For MC, this means that information about each particle type could be obtained, as well as for the entire peptide chain. For MD, similar information could be obtained, along with further details. For example, lipid tails and head groups could be examined separately, as could each amino acid residue of the peptide, and the distribution of water and ions could be determined.

5.3.5 Quality of sampling and sampling convergence

When performing molecular simulations, it is important to remember that the obtained results are never better than the statistical quality of the data sampling. There are two factors giving rise to errors, inaccuracy in the models and insufficient sampling.[137] While it is hard to ensure sufficient sampling, there are methods to detect this.[138] Additionally, there is no way to know if all regions of the configurational space have been visited during the simulation without previous knowledge about it. The focus is instead directed towards assessing the sampling quality of the visited regions of space.[137] An explanation of the methods used for the work in this thesis follows below. A more detailed guide can be found in e.g. the following references.[137,138]

As a first step, single observables as a function of time can be evaluated, to ensure that basic equilibration has been reached. Such observables are e.g. R_g and R_{ee}. These usually display large fluctuations for IDPs, due to the many interchanging conformations the peptides can attain. However, it is still usually possible to detect systematic changes. Calculating error estimates and observing correlations can be used to assess the quality of sampling of a single observable obtained from a simulation. The correlation time, τ_f, of a certain observable f can be explained as the time it takes for the simulation to forget about previous values of f. It is defined by

$$\tau_f = \int_0^\infty C(t)dt, \tag{5.21}$$

where $C(t)$ is the *autocorrelation function*, defined as[128]

$$C(t) = \langle f(t')f(t'+t)\rangle_{t'}. \tag{5.22}$$

The correlation time can be used to estimate the quality of sampling by $N = t_{tot}/\tau_f$, where t_{tot} is the total simulation time. A value of $N >> 1$ indicates good sampling.[137]

The block averaging method is based on dividing the simulated trajectory into M segments or blocks, with a length n. An average of the observable is calculated for each block, B_i, resulting in M values. The block size, n, is gradually increased, and the *block-averaged standard error* (BSE) is calculated for each block size, according to

$$\mathrm{BSE}(n) = \frac{\sum_{i=1}^{M}(B_i - \langle B \rangle)^2}{M(M-1)}, \tag{5.23}$$

where $\langle B \rangle$ is the total average for the given block size. The BSE can be used as an estimator of the true standard error, when the length of each block is significantly larger than the correlation time, meaning, the blocks are independent of each other. A converged BSE plot indicates that the error estimate of that specific observable is converged.

As the methods presented here only provide information about the sampling of single observables, and not about the global sampling quality, it is always beneficial to run several replicas with different initial simulation conditions to compare the results.

Science adjusts its views based on what's observed.
Faith is the denial of observation so that belief can be preserved.
- Tim Minchin

6 | Summary of main results

Contents

In this chapter, the main results of the papers are summarized, together with an ongoing project. The research has been focused on investigating the interaction between small IDPs and model membranes. The main focus has been directed towards Hst5 and variants thereof to understand how the primary structure affects the interaction behavior. Additionally, a study of KEIF was conducted to investigate the structure function paradigm commonly applied to IDPs and IDRs.

6.1 Interaction of Histatin 5 with phospholipid bilayers - cushion formation

From previous studies,[32–34] it is well known that Hst5 possesses antifungal properties, and that the peptide does not attack the cell membrane, but rather kills the fungi through intracellular processes. The mechanism behind this interaction is, however, unknown. The interaction of Hst5 with a model membrane was therefore investigated from a physiochemical perspective, where the effect of electrostatic interactions were investigated by altering different aspects of the system, each discussed separately below. The study was conducted using NR and QCM-D. The overall results of Paper I can be summarized as; upon exposure of Hst5 to a POPC:POPS in a 9:1 ratio (PC_9PS_1) SLB, deposited on a negatively charged silica surface, the peptide translocated the bilayer without disrupting the integrity of the bilayer. It then resided in the gap between the silica surface and the bilayer. The presence of the peptide close to the silica surface, most likely promoted counterion release, which induced an increase in osmotic pressure between the silica surface and the bilayer, which pushed the bilayer further away from the solid surface, forming a cushion. The negative charge of the silica mimics the potential of the mitochondrion in the cell, relating its required presence to the biological mechanism of Hst5.

6.1.1 Salt concentration

Four different salt concentrations were evaluated: 10, 80, 140, and 500 mM NaCl. The lowest and highest NaCl concentrations were studied with QCM-D for two different bilayer compositions, purely zwitterionic POPC and negatively charged PC_9PS_1. The data showed that at high salt concentration the adsorption was highly reduced, and completely hindered in the case of POPC, as is shown in Figure 6.1A and B. Data obtained from QCM-D does not allow to determine with certainty where in the system the peptide resides, only that the peptide adsorbed. Therefore, NR was utilized to investigate the interaction of Hst5 to a PC_9PS_1 bilayer. Measurements were performed

at a NaCl concentration of 10, 80, and 140 mM. From these measurements, no effect on the bilayer was detected at either 80 mM, or 140 mM NaCl. For the 10 mM sample, however, an increase in reflectivity was observed in the Dbuff contrast observed around q_z = 0.075× 10^{-6} Å^2, as well as a deepening of the minimum around q_z = 0.03×10^{-6} Å^2, as is shown in Figure 6.1C. This indicated a thickening of the sample at the silicon-water interface, but that the bilayer was otherwise unaffected. The data was interpreted such as that the peptide had translocated the bilayer, and formed a peptide cushion between the solid surface and the bilayer, which is schematically illustrated in Figure 6.1D.

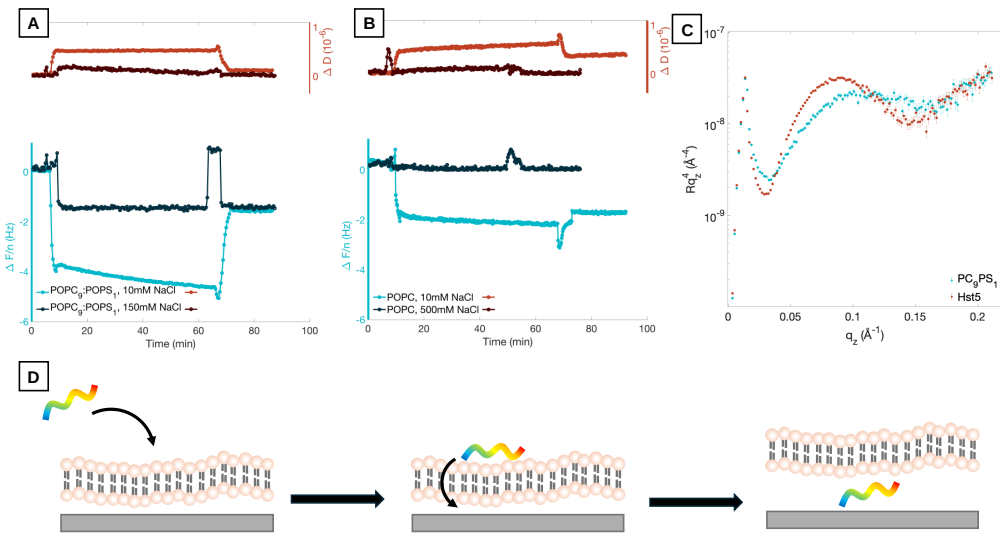


Figure 6.1: Normalized QCM-D for **(A)** a negatively charged bilayer, and **(B)** a zwitterionic bilayer upon interaction with Hst5. Measurements for both bilayers was performed at 10 mM and 500 mM NaCl. The baseline corresponding to the signal of an SLB (ΔF = -25 Hz) was set to 0. Data are shown for the 11th overtone, the most sensitive to thin films. **(C)** Displays reflectivity curves in Dbuff for the PC_9PS_1 SLB before and after Hst5 exposure. **(D)** A schematic representation of the Hst5 and bilayer interaction where the peptide translocates the bilayer without disrupting the integrity of the bilayer.

6.1.2 Solid surface

To investigate how the solid surface influences the Hst5-SLB interaction, NR measurements were conducted with the SLB on randomly oriented sapphire surfaces (Al_2O_3), in addition to the measurements performed on silicon surfaces. The sapphire surfaces were positively charged at the pH conditions used in this study. It was found that upon Hst5 exposure, the SLB desposited on sapphire was completely unaffected, hence, a negative charge below the bilayer was required for interaction to occur.

6.1.3 Bilayer composition

Three different lipid compositions were used to form bilayers; a purely zwitterionic bilayer composed of only POPC, and two compositions with a net negative charge; PC_9PS_1 and PC_6PS_4. As is shown in Figure 6.1A and B, when inveastigated with QCM-D, a higher frequency shift was observed when Hst5 was exposed to the PC_9PS_1 bilayer compared to the POPC one, for both NaCl concentrations. This indicated that a larger mass was adsorbed to the PC_9PS_1 bilayer, explained by the electrostatic attraction from the opposite charge of the peptide and the bilayer. In the NR experiments performed on the POPC bilayers, no structural nor compositional changes were observed at any NaCl concentration. The interaction of Hst5 with the PC_6PS_4 bilayer at 140 mM NaCl was investigated with NR, and showed, as opposed to measurements with the PC_9PS_1 bilayer at the same salt concentration, that Hst5 not only translocates the bilayer, forming a cushion, but also resides within its hydrophobic region. It is however important to note, that the surface coverage of the pristine bilayer was characterized to be roughly 85% for the lipid PC_6PS_4 sample, which means that the adsorption could be affected by the holes in the bilayer, exposing the underlying silica surface.

6.2 Interaction of Histatin 5 with phospholipid bilayers - the effect of histidines

In Paper I, the conditions under which the cushion was formed were determined. As the translocation across the membrane is an integral part of the antifungal mechanism of Hst5, the cushion formation can be used as a measure to indicate the effectiveness of different variants of Hst5 towards fungi. The sequence of Hst5 contains seven histidine residues, making up 29% of all the peptide's amino acids. Histidine is also able to charge regulate, meaning that it can change its charge when approaching other molecular matter, such as a lipid bilayer.[27] This has been hypothesized to be of importance for the ability to translocate the bilayer. In Paper II, we therefore investigated the interaction behavior of four variants of Hst5 with SLBs. The variants contained zero to four histidine, compared to the original seven. These variants were previously studied by Cragnell *et al.*[38] regarding their zinc induced oligomerization, and localization of zinc-binding motifs in the sequence of Hst5. Previous studies have investigated the antifungal effect of Hst5 in the presence of metal ions, such as copper,[39] iron,[139] and zinc,[140] all naturally present in saliva. Zinc was shown to increase the killing activity of Hst5.[140] Both the special characteristics of histidine, as well as its involvement in binding zinc to Hst5, encouraged this investigation, which was conducted without the presence of zinc to investigate the effect of altering the amino acid sequence alone. The

behavior of these peptides were investigated in three conditions; in bulk solution, interaction with a silica surface, and interaction with an SLB composed of PC_9PS_1. All experiments were performed at a salt concentration of 10 mM at pH 7.4. The results obtained from each experimental condition will be discussed separately below. The overall results of the investigation was that the number of histidines affects the penetration depth into the bilayer, as depicted in Figure 6.2, and thereby also the biological effect of Hst5. It also supports our proposed hypothesis that the possibility of histidine to charge regulate is important for the peptide to translocate the bilayer.

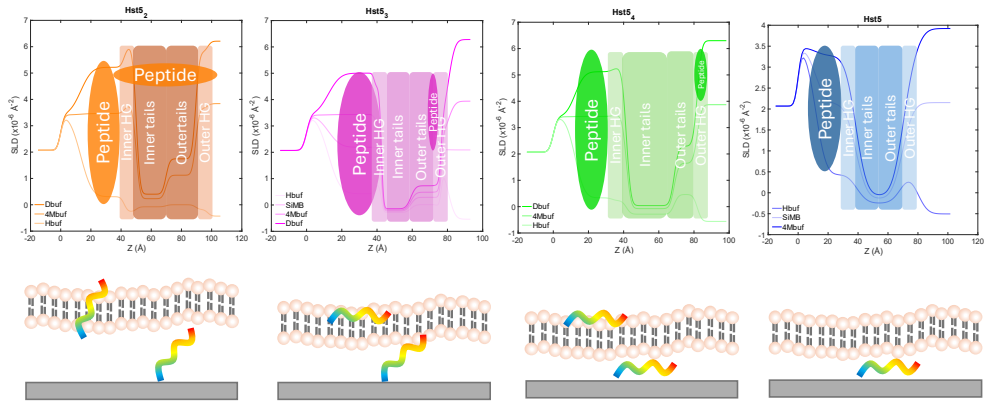


Figure 6.2: Schematic figure depicting the position of the peptide after rinsing from NR data, with increasing number of histidines in the peptide, left to right. The different parts of the SLD curves are highlighted for clarity.

6.2.1 Bulk properties of Histatin 5 and variants

To ensure that no major structural changes were induced upon varying the number of histidine residues in the sequence (sequences presented in Table 6.1), SAXS measurements were conducted, which, from the obtained normalized Kratky plots, suggested that all peptides were unfolded and had a flexible structure in solution. The Guinier approximation, Equation 4.5 was used to determine the R_g values for all peptides, which were all in the range between 11.6 to 12.0 Å. The secondary structure of the peptides were investigated with CD in both aqueous buffer supplemented with NaCl, and TFE. The measurements in TFE were conducted because Hst5 is known to adopt an α-helical structure in TFE,[13,29,30] which can be used to mimic the hydrophobic part of the membrane. This helical formation has been suggested to be important for the biological activity of Hst5.[26,42] In aqueous buffer, all peptides were mainly disordered, in agreement with the observations from SAXS, while in TFE, all peptides adapted a more α-helical structure, as expected. The conclusion from this part of the study is that the sequence alterations did not affect the shape or flexibility of the peptides, which is

important to know for the following parts of the study. Furthermore, any possible differences observed does not depend on structural differences present before interaction with the different surfaces.

Table 6.1: The amino acid sequences of the peptides included in this thesis. Positively charged residues are presented in blue, negatively charged ones in red, and histidine residues are presented in green.

Peptide	Sequence
Hst5	DSHAKRHHGYKRKFHEKHHSHRGY
$Hst5_0$	DSQAKRQQGYKRKFQEKQQSQRGY
$Hst5_2$	DSQAKRQHGYKRKFQEKQHSQRGY
$Hst5_3$	DSQAKRQQGYKRKFHEKHHSQRGY
$Hst5_4$	DSHAKRHHGYKRKFQEKQQSHRGY

6.2.2 Interaction with a silica surface

The adsorption behavior of all peptides to a negatively charged silica surface were investigated, using QCM-D, and analyzed using the Sauerbrey equation, Equation 4.10, to convert the frequency shifts into adsorbed mass, which was done for all peptides before rinsing (at $t \approx 65$ min) and after rinsing (at $t \approx 90$ min). The three peptides, namely $Hst5_2$, $Hst5_4$, and Hst5, displayed a similar amount of adsorbed peptide to the silica surface, around 145 ng cm^{-2}, and 105 ng cm^{-2}, before and after rinsing, respectively. The frequency shift of $Hst5_4$ and Hst5 were both stable upon incubation, while the frequency shift for the $Hst5_2$ slowly increased during this period. The increase in frequency for $Hst5_2$ could be explained by a change in conformation of the peptide on the surface, from a disordered to a more flattened state, which would release coupled water molecules, decreasing the adsorbed mass. The remaining two peptides, $Hst5_0$ and $Hst5_3$, both showed a significantly smaller frequency shift upon adsorption, corresponding to about 60 ng cm^{-2}. Upon rinsing, a significant amount of peptides were removed, and only 9% and 40% of the mass remained after rinsing, for $Hst5_0$ and $Hst5_3$, respectively. In both these peptides, the histidine residues in position 3, 7, and 8 were removed, which are all in the zinc motif suggested by Cragnell *et al.*[38]. Additionally, in previous computational studies,[28,141] upon interaction with a negatively charged surface, amino acids 5–13 were in closest contact with the surface. Hence, removing the histidine residues in this region, with their ability to charge regulate, probably diminished the attraction between the peptide and the surface, resulting in a smaller adsorbed amount.

6.2.3 Interaction with a supported lipid bilayer

To both get information about the position of the peptides in relation to the SLB, hence, the penetration depth, and how much peptide that was adsorbed to the SLB, NR and QCM-D measurements were conducted. Results showed that presence of at least two histidine residues in the peptide was required for bilayer translocation, and that the penetration depth increased with number of histidines. From the NR measurements it was found that the $Hst5_0$ variant behaved completely different compared to the other peptides, and will be discussed separately. While Hst5 was found only in the gap between the solid substrate and the bilayer, the other three variants were found in the gap, but also within the bilayer. Besides all three variants being found in the gap, their positions within the bilayer was as follows. $Hst5_4$ was in the head group region of the outer bilayer. $Hst5_3$ was in the head group region of the inner leaflet and in both the tail region and head group region of the outer leaflet. Lastly, $Hst5_2$ was found to reside within all layers of the bilayer. The bilayer exposed to $Hst5_2$ also showed signs of lipid removal. The thickness of the formed cushion also increased with decreasing number of histidines, possibly due to the peptides adopting a flatter structure with higher number of histidines, as suggested from interactions with soley a silica surface. These results support the hypothesis that the charge regulation of histidine plays a role in the translocation across the bilayer. In the case of $Hst5_0$, after incubation of the peptide and rinsing with pristine buffer, the obtained curves no longer resembled that of a bilayer, and no particular order or periodicity was observed. It was possible to determine a sample thickness of approximately 250 Å from the spacing between two subsequent fringes in the specular reflectivity data, using the relation $d = 2\pi/\Delta q_z$.

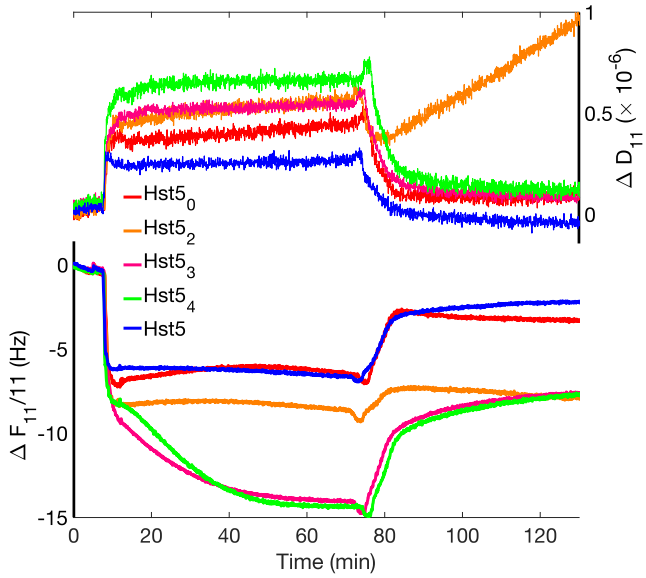


Figure 6.3: QCM-D data obtained for the 11th overtone for Hst5$_0$, Hst5$_2$, Hst5$_3$, Hst5$_4$, and Hst5.

The adsorption profiles obtained from QCM-D measurements for Hst5$_0$, Hst5$_2$ and Hst5 were similar to those obtained upon adsorption to a silica surface; a fast decrease in frequency shift upon adsorption, followed by a stable trace during incubation. This is shown in Figure 6.3. The frequency shift was then rapidly increased upon rinsing with pristine buffer, followed by a stable frequency trace upon continuous rinsing. In the case of Hst5$_3$ and Hst5$_4$, the frequency shift was further decreased during the incubation period. There are two possible explanations for the additional decrease in the frequency shift: (i) an additional, slower, adsorption process during which additional peptide molecules are adsorbed without affecting the dissipation values, or (ii) a slower translocation process, again, without increasing the dissipation of the system. In the latter, the bilayer is slowly lifted from the surface, increasing the dynamically coupled water in the cushion region, resulting in the observed decrease in frequency. With QCM-D, it was not possible to distinguish between these two scenarios.

6.3 Interaction of Histatin 5 with phospholipid bilayers - the effect of peptide chain length

It has previously been shown that the C-terminal part of Hst5 has a similar antimicrobial effect as the full length Hst5 peptide.[26,42–46] To investigate the importance of

chain length on the peptides ability to translocate a SLB, we conducted a study with a longer peptide chain, referred to as ^{48}Hst5, which is the tandem repeat of Hst5, and a shorter peptide ^{14}Hst5, corresponding to the last 14 amino acids of Hst5. Both of these peptides maintain the histidine ratio of 29% as in Hst5, which we showed in Paper II is important for the translocation across the bilayer, and their sequences are presented in Table 6.2. In this study, the three peptides were investigated in bulk to see how the sequence length affects the inherent structural properties. The interaction of the peptides with PC_9PS_1 SLBs were also investigated, and the main findings of Paper III was that at low NaCl concentration, both the shorter and longer peptide variant behaved similar to Hst5, while at high salt concentrations, both variants interacted with the SLB, while Hst5 did not. Each of these cases are discussed in more detail below.

Table 6.2: The amino acid sequences of the peptides included in this thesis. Positively charged residues are presented in blue, negatively charged ones in red, and histidine residues are presented in green.

Peptide	Sequence
Hst5	DSHAKRHHGYKRKFHEKHHSHRGY
^{14}Hst5	KRKFHEKHHSHRGY
^{48}Hst5	DSHAKRHHGYKRKFHEKHHSHRGYDSHAKRHHGYKRKFHEKHHSHRGY

6.3.1 Bulk properties of of the peptides in different conditions

As discussed previously, in a pH 7.4 buffer supplemented with 10 mM salt (NaCl/NaF), Hst5 is a rather flexible, disordered peptide, and in TFE the proportion of α-helical structure is significantly increased, making the peptide more ordered. In this study, it was also shown that Hst5 maintained the flexible and disordered structure in a buffer solution with pH 7.4 supplemented with 150 mM salt. In comparison with the 10 mM NaCl case, the only observed difference was a slight decrease in β-sheet structure, which was instead predicted to be disordered. The longer variant ^{48}Hst5 displayed a very similar behavior as Hst5, flexible and mainly disordered in aqueous buffer at both salt concentrations. It adopted mainly an α-helical structure in TFE, to a higher extent than Hst5. As for Hst5, the structure was largely unaffected by the salt concentration. The shorter ^{14}Hst5 did display a structural salt dependence, where the peptide, like Hst5 and ^{48}Hst5, was mainly flexible and disordered at a low salt concentration. However, at high salt concentration, the normalized Kratky plot obtained from the SAXS measurements, indicated a rather globular structure. This behavior was not captured in the CD measurements, which did not indicate much structural elements present in the peptide. In TFE, ^{14}Hst5 did not adopt a more α-helical structure like Hst5 and

^{48}Hst5, and the secondary structure elements predicted from CD data was the same in TFE as in aqueous buffer. It was previously observed by Raj *et al.*[42] that ^{14}Hst5 assumed α-helical structure to a lesser extent than Hst5, which was attributed to its shorter sequence length. However, in that study, the helical content was larger than we observed here.

6.3.2 Interaction with a supported lipid bilayer at low NaCl concentration

The interaction of the peptides with a PC_9PS_1 SLB at 10 mM NaCl, was shown to be very similar. As was previously observed with NR for Hst5, all peptides were found in the gap between the solid substrate and the lipid bilayer, forming a peptide cushion as shown by the volume fraction profiles (VFPs) in Figure 6.4. These VFPs are obtained from the analysis of NR data. The thickness of the formed cushion was very similar for Hst5 and ^{14}Hst5 at 20 $\pm$ 1 Å and 19 $\pm$ 1 Å, respectively, while for ^{48}Hst5, the thickness of the formed cushion was only 12 $\pm$ 2 Å. The thickness can be presented in terms of the thickness of the layer composed of only peptide molecules, given by $D_{pep} = \Phi_{pep} \times t_{pep}$, where Φ_{pep} is the volume fraction of peptide, and t_{pep} is the thickness of the entire layer. Values of D_{pep} = 9 $\pm$ 1, 6 $\pm$ 1, and 6 $\pm$ 1 Å, for Hst5, ^{14}Hst5, and ^{48}Hst5, respectively, were obtained, which did not differ significantly between the peptides. The similarity of these values indicated a different molecular organization and a different amount of hydration water of the peptides within the gap region.

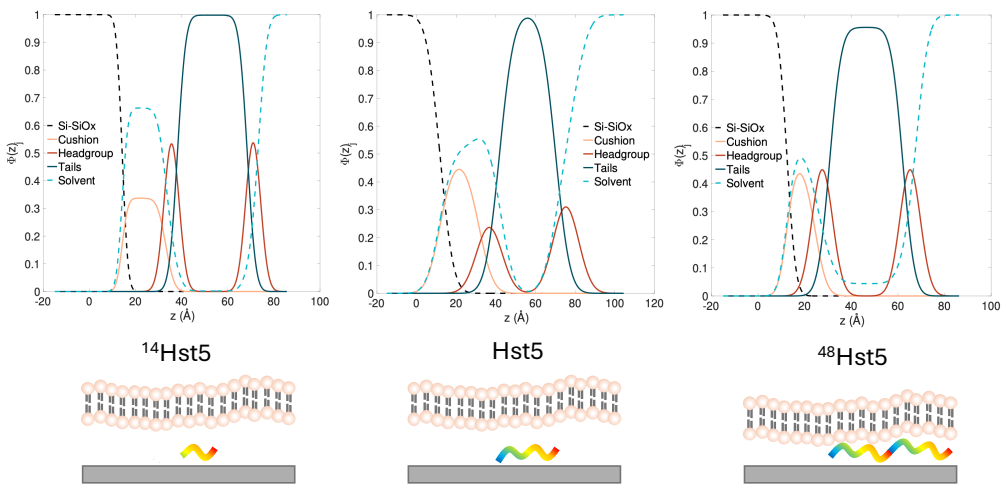


Figure 6.4: The VFPs obtained from the analysis of the NR data together with schematic figures depicting the position of the peptides after rinsing at 10 mM NaCl.

From the QCM-D measurements it was clear that upon injection of peptide sample,

the smallest frequency shift was observed for Hst5, while the largest one was found for ^{14}Hst5, and ^{48}Hst5 having an in intermediate shift, indicating that more peptide molecules and hydration solvent were adsorbed of the longer and shorter peptides, compared to Hst5. The dissipation value was also smallest for Hst5, indicating that the other two peptides adsorbed less rigidly upon injection. In all cases, when the system was rinsed with pristine buffer, the frequency shift became smaller, indicative of desorption of molecules. The most interesting part of the rinsing step was however that the dissipation shift of Hst5 and ^{14}Hst5 returned to zero, i.e. to the same rigidity as the pristine bilayer before exposure to the peptide solution, while the dissipation value for ^{48}Hst5 converged to 0.2×10^{-6}, hence, ^{48}Hst5 gave rise to a less rigid layer compared to the other two peptides. A similar conversion as the one performed on the NR data to obtain the D_{pep} value was done to the adsorbed mass obtained from the frequency shift, which was explained in Chapter 4. These conversions were performed to obtain values with similar properties to be able to more quantitatively compare the results from these two methods. The t^{QCM} values were determined after rinsing each system with pristine buffer, and the values were determined to be $t^{QCM} = 3.0 \pm 0.3$, 11.6 ± 0.4, and 5.7 ± 0.3 Å for Hst5, ^{14}Hst5, and ^{48}Hst5, respectively. It is important to note that the comparison between these values is not trivial, and that while they are in the same order of magnitude, the exact values might still differ due to the intrinsic difference between the NR and QCM-D measuring principles. However, the conclusions from the comparison were that the low values for both D_{pep} and t^{QCM} indicated that only a small amount of peptides interacted with the SLB. Additionally, the D_{pep} and t^{QCM} values obtained for ^{48}Hst5 are almost identical, which, together with the observed lower hydration of the cushion compared to the other two peptides, possibly indicated that ^{48}Hst5 has a flatter conformation within the cushion.

6.3.3 Interaction with a supported lipid bilayer at high NaCl concentration

The analysis of the interaction became less straight forward when the NaCl concentration was increased to 150 mM. As previously discussed, neither cushion formation, nor interaction of any kind was observed for Hst5, when studied with NR, as is shown in the VFP in Figure 6.5. However, with QCM-D, it was possible to detect some interaction. Upon injection of peptide to the PC_9PS_1 SLB, a small, almost negligible peptide adsorption was observed. During incubation, the frequency shift increased and became positive, implying that a mass greater than that of the peptide was removed, which was further observed upon rinsing. The overall shift in frequency, either negative or positive, was very minor, but to speculate, it could originate from minor reorganizations within the system when the peptide interacts with the bilayer, as depicted in Figure 6.5. As opposed to Hst5, both ^{14}Hst5 and ^{48}Hst5 did display significant interaction with the PC_9PS_1 bilayer, and will be discussed separately below.

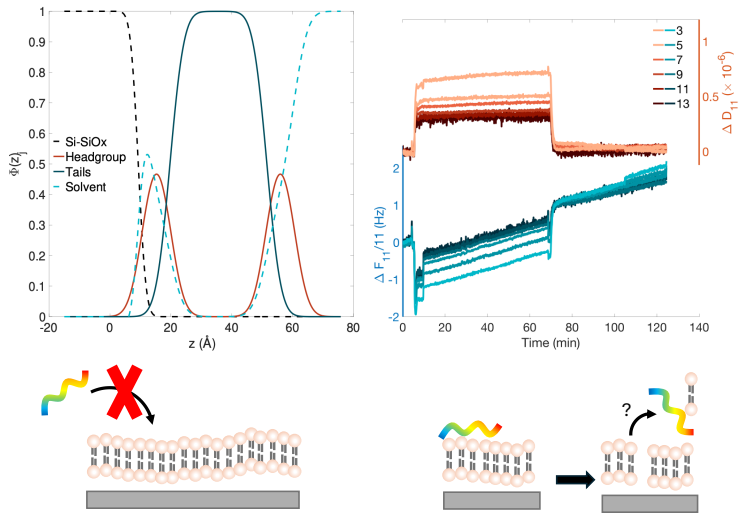


Figure 6.5: The VFP obtained from the analysis of the NR data, and QCM-D traces, together with schematic figures depicting the interaction of Hst5 at 150 mM NaCl.

^{14}Hst5

Through analysis of the NR data, the ^{14}Hst5 peptide was found to reside, not only in the gap between the solid surface and the bilayer forming a small cushion, but also in the whole bilayer, as indicated in the VFP in Figure 6.6. This interaction between the peptide and the bilayer affected the structure of the bilayer, where both the thickness and the volume fraction of solvent increased. The increased hydration of the lipid bilayer was not visible in the QCM-D measurements, which indicated that the sample remained rigid even after incubation of ^{14}Hst5, evident by the low dissipation values as well as the overlapping normalized frequency shifts.

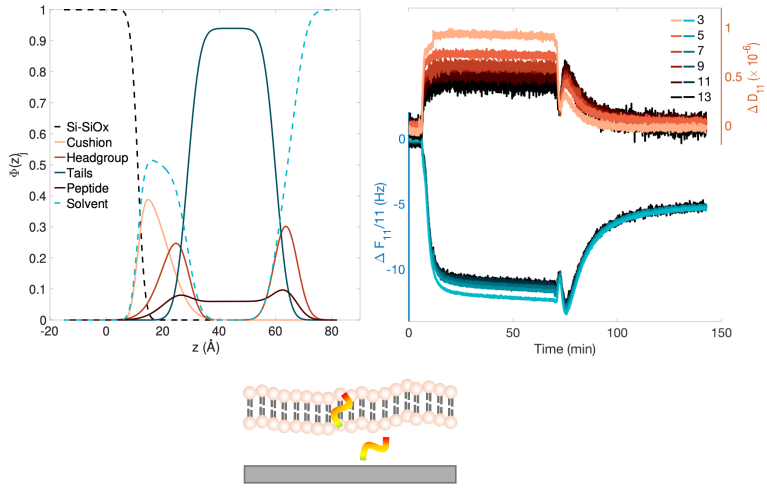


Figure 6.6: The VFP obtained from the analysis of the NR data, and QCM-D traces, together with schematic figures depicting the interaction of ^{14}Hst5 at 150 mM NaCl.

^{48}Hst5

The analysis of the NR data obtained for the ^{48}Hst5 sample differed from the other samples in that it was not possible to fit the data with the same number of layers as the pristine bilayer; for this sample an additional layer on top of the bilayer had to be added. The peptide was found to reside in the gap below the bilayer, forming a cushion, as well as on top of the bilayer in a very thick, highly hydrated layer composed of a mixture of peptides and lipid molecules, which is shown in the VFP, and small schematic figure in Figure 6.7. This finding was also confirmed with the QCM-D data, which showed the same behavior as already described for Hst5 and ^{14}Hst5 up until the rinsing step. During rinsing, an initial increase in frequency was observed, however, it was followed by a slight, but marked, decrease upon continuous rinsing. The dissipation values mirrored this behavior to some extent, but with faster stabilization and reaching values on the borderline between those characteristic of a rigid and a viscoelastic regime. Since no additional material was added upon the rinsing step, the observed behavior indicated a dynamic process where some material was initially removed from the adsorbed layer, without completely detaching, followed by a second adsorption, potentially in a mixture of peptides and lipid molecules in solution on top of the bilayer. The system does not return to a dissipation of zero, as the other systems, indicating that the adsorbed layer is more viscoelastic, which is coherent with the suggested structure of the sample.

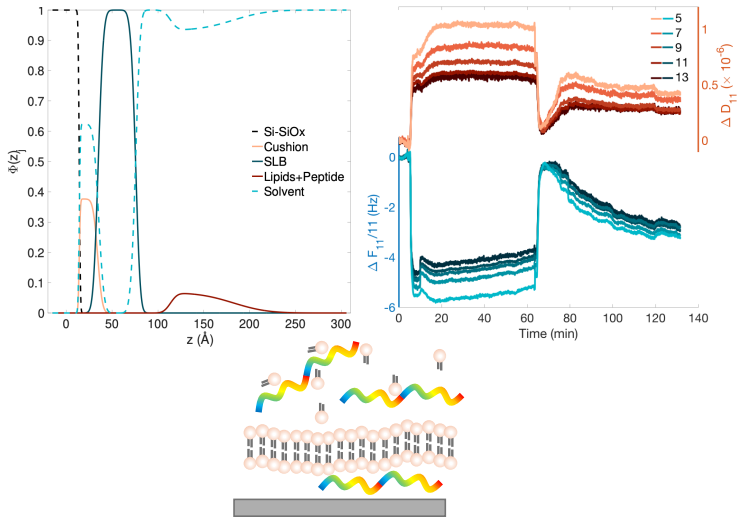


Figure 6.7: The VFP obtained from the analysis of the NR data, and QCM-D traces, together with schematic figures depicting the interaction of ^{48}Hst5 at 150 mM NaCl.

6.3.4 Simulations

Coarse-grained MC simulations were performed of a system designed to mimic the peptide residing within the cushion. For this, the peptide was enclosed in a box, with two parallel surfaces, one representing the silica surface and the other the lipid head groups of the bilayer. Three separations were selected, 20, 40, and 100 Å, to investigate if, and how the peptides were affected by this. All three peptides preferred to adsorb to the surface, mimicking the bilayer over the one mimicking the silica surface. This behavior is explained by that the bilayer surface has both a higher negative charge per particle, and also a higher overall charge, compared to the silica surface. To elucidate the adsorption behavior, both number density profiles (in the z-direction), displayed in Figure 6.8, and adsorption probability profiles, Figure 6.9, were obtained for all peptides in each scenario. In these figures, the results obtained in a salt concentration of 10 mM are displayed, while simulation with 150 mM salt were also performed, they are not presented here. The behavior was similar in 150 mM salt, and was therefore omitted for clarity. The number density and the adsorption probability both showed that all three peptides adsorbed to both surfaces when they were 20 Å apart, hence, the peptides were forming a bridge between the two surfaces. When the distance between the surfaces increased, as mentioned prior, all three peptides showed a preference for the bilayer surface.

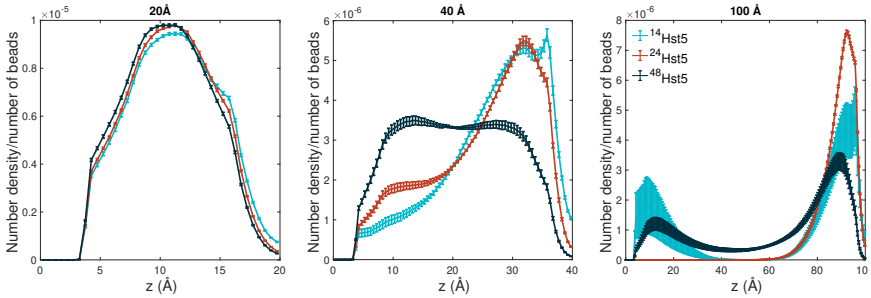


Figure 6.8: Number density summed over all the beads in a chain in z-direction divided by the number of beads in each chain. The surfaces are located at each end of the z-direction, where the surface mimicking the silica surface is placed at z = 0, and the one mimicking the bilayer is on the other end. Points were obtained every 0.5 Å and the salt concentration was set to 10 mM. The depleted region close to the silica surface is due to hard-sphere repulsion. This is not observed for the bilayer as that surface is not as densely packed with particles. Hence, the chain is allowed in between the surface particles.

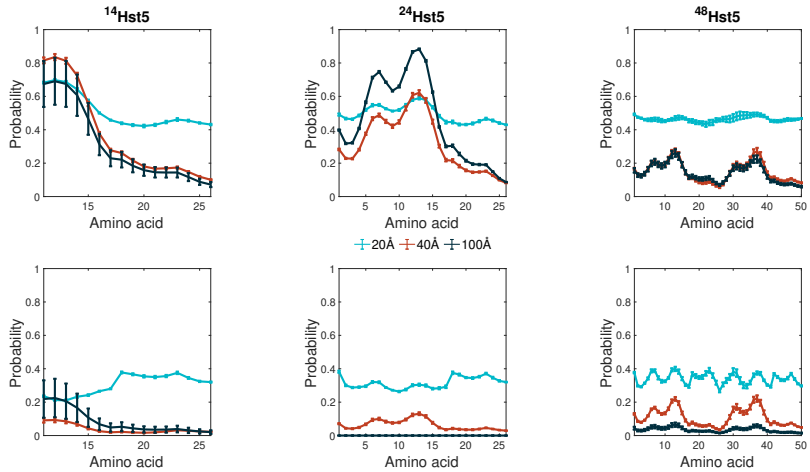


Figure 6.9: Adsorption profiles of the three different peptides to a surface mimicking a bilayer (top) with a total charge of -78e (-0.5e/point) and a surface mimicking a silica surface (bottom) with a total charge of -49.5e (-0.05e/point) at 10 mM salt. The distances are represented by turquoise for 20 Å, red for 40 Å, and dark blue for 100 Å.

6.4 Interaction of Histatin 5 with phospholipid bilayers - the effect of the peptide sequence order

In Paper II, the results suggested that it was not only the number of histidines in the sequence, but also their position, which was important for the translocation across the bilayer. We therefore decided to investigate how, not only the position of the histidines, but the position of all amino acids in the sequence, affected the ability of the peptide to

translocate the bilayer. For this, two variants with randomized sequences were investigated, which are presented in Table 6.3. This study was conducted using NR, QCM-D, and two computational methods, namely atomistic MD, and coarse-grained MC simulations. This work is still ongoing and has not yet resulted in a manuscript. The experiments and calculations presented here were performed in low salt concentration, thus, 10 mM. The conclusions of this study in its current state is that the order of the amino acid residues are in fact of importance, since neither of the two variants form a cushion without also affecting the bilayer. However, the explanation to why is not yet understood and further investigations are needed.

Table 6.3: The amino acid sequences of the peptides included in this thesis. Positively charged residues are presented in blue, negatively charged ones in red, and histidine residues are presented in green.

Peptide	Sequence
Hst5	DSHAKRHHGYKRKFHEKHHSHRGY
Hst5^{rand1}	HHYARKSDKHSFHYGRHKHERGKH
Hst5^{rand2}	HGHSYKKAGYKEHSKHHHHDRRRF

6.4.1 Experimental results

The results of Hst5 have been extensively discussed in Papers I–III, and will not be further discussed in this part. The results obtained for the two variants will be compared to the results of Hst5.

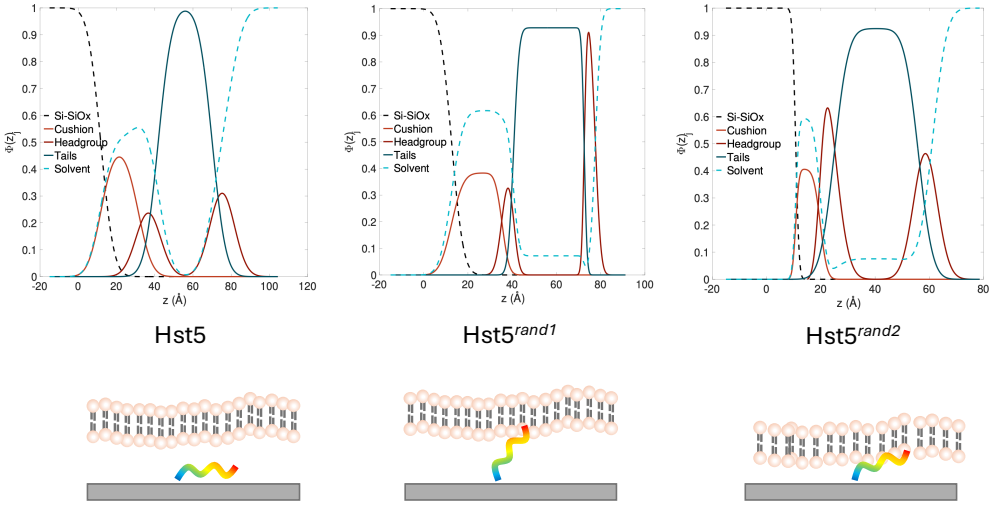


Figure 6.10: The VFPs obtained from the analysis of the NR data for Hst5 (left), Hst5^{rand1} (middle), and Hst5^{rand2} (right), together with schematic figures depicting the position of each peptide with respect to the SLB.

Hst5^{rand1}

The behavior observed for Hst5^{rand1} was similar, but not identical with what was found for Hst5, studied with NR. A peptide cushion, of thickness t_{gap} = 23 ± 2 Å, was formed under the bilayer, as illustrated in Figure 6.10 (middle). The Hst5^{rand1} peptide did not only reside in the gap between the solid substrate and the bilayer, but also in the head group region of the inner leaflet. The layers where the peptide was found were highly hydrated, almost 60%. The tail region and outer head group region were mostly unaffected by the presence of the peptide. In the QCM-D data, Hst5^{rand1} displayed a lower frequency shift, but a higher shift in dissipation upon injection and incubation of peptide, compared to Hst5, as is shown in Figure 6.11. This indicated a smaller adsorbed amount, together with a more viscoelastic adsorbed layer. Addition of adsorbed material during incubation without increase in viscoelasticity was showcased for Hst5^{rand1} as a slow decrease in frequency, while the dissipation value was constant, which is not observed for Hst5. Upon rinsing, the dissipation value decreased to a value close to zero, while the frequency was increased. The frequency shift after rinsing was larger for Hst5^{rand1} compared to Hst5, this could be an influence of more coupled water in the adsorbed layer, as was also observed from the NR data where the bilayer was more hydrated after interaction with Hst5^{rand1} compared to Hst5.

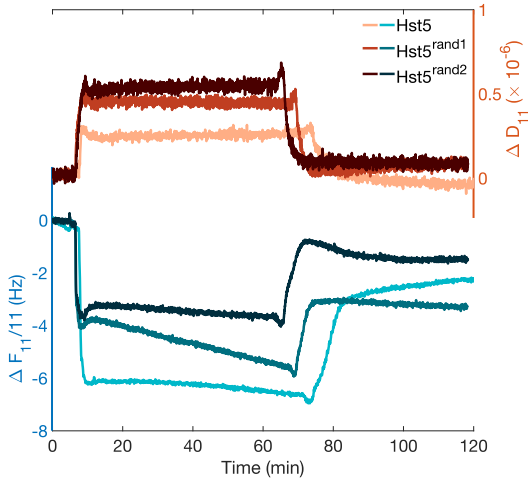


Figure 6.11: QCM-D data obtained for the 11th overtone for Hst5, Hst5^{rand1}, and Hst5^{rand2}.

Hst5^{rand2}

The results for Hst5^{rand2} indicated the formation of a cushion, however, the thickness of the peptide layer was significantly smaller, t_{gap} = 9 ± 1 Å, compared to both Hst5 and Hst5^{rand1}. Similarly to Hst5^{rand1}, the peptide molecules also penetrated into the region of the inner head group. The whole bilayer was affected by the peptide interaction, and the tail region of the bilayer contained roughly 10% buffer (compared to < 1% before peptide interaction), most probably as a result of imperfections in the bilayer structure due to peptide interaction, as appear in the VFP obtained from analysis of NR data, shown in Figure 6.10 (right). From the QCM-D data, Figure 6.11, a similar behavior of Hst5^{rand2}, compared to Hst5, was observed, where the frequency and dissipation values were constant upon incubation, indicating no change in either adsorbed mass or viscoelasticity during this period. Upon rinsing, a fast increase in frequency was observed, indicative of material removal, however, upon continued rinsing, the frequency shift for Hst5^{rand2} was slowly decreased, while the dissipation value remained stable. As previously discussed, this indicates addition of adsorbed material without affecting the rigidity of the adsorbed layer.

6.4.2 Computational results

Despite the overall similar interaction with the PC_9PS_1 bilayer, there were some differences of Hst5^{rand1} and Hst5^{rand2} compared to Hst5. To better understand these differ-

ences, molecular simulations, both coarse-grained MC and atomistic MD simulations were performed.

Coarse-grained Monte Carlo simulations

Coarse-grained MC simulations were performed for the three peptides adsorbed to a surface representing the outer head groups of the lipid bilayer at 10 mM salt and with three different protonation states: (i) the histidines depronated, that is, uncharged, (ii) the histidines partially protonated, with a charge of + 0.5 e, or (iii) the histidines fully protonated, with a charge of + 1 e. This is of importance for the system because Hst5, with its high content of histidines, can charge regulate, as it approaches the head groups of the lipid bilayer facing the bulk solution. This will slightly increase the charge of the peptide, as previously proven from computer simulations by our group, [28,141,142] which allows the peptide to diffuse through the bilayer. The adsorption profiles in Figure 6.12 show as expected, since these simulations primarily considered electrostatic interactions, that the highest adsorption probability was observed when the histidine residues were fully protonated, as that gives the peptides an additional positive charge of +7, resulting in a net charge of +12. For all peptides, the overall adsorption profile was the same when increasing the charge of the histidine residues from +0.5 to +1, however, the shape changed when the histidine residues were uncharged compared to when they carried a charge. What is interesting from these results is that there was a clear difference in the adsorption profiles between the different peptide chains. The two randomized variants, Hst5^{rand1} and Hst5^{rand2}, shared a similar adsorption profile, which differs from that of Hst5. From the sequence (presented in Table 6.3) one would expect Hst5^{rand2} to differ from the other two when the histidine charge is altered, due to the clustering of this amino acid in the C-terminus of the peptide.

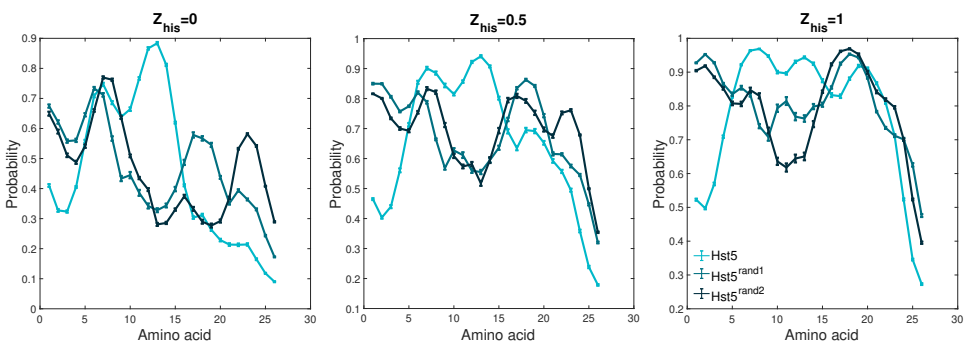


Figure 6.12: Adsorption probabilities of histidine charge 0 e (left), histidine charge + 0.5 e (middle), and histidine charge + 1 e, for Hst5, Hst5^{rand1}, and Hst5^{rand2}.

Atomistic molecular dynamics simulations

Atomistic MD simulations were performed for all three peptides in the vicinity of a PC_9:PS_1 bilayer to investigate differences in their behavior. In Figure 6.13A, the number density divided by the number of atoms of each component in the simulation and centered around the PC_9:PS_1 bilayer, is shown. The peptides were mainly outside the bilayer, however, the density of all peptides were somewhat overlapping with that of the lipid head groups, hence, indicative of interaction between the peptide and the bilayer. The Hst5^{rand1} showed the most overlap, albeit slight, with the bilayer head group, as seen from Figure 6.13B where the average minimum distance between each amino acid residue with a phosphorus atom in the bilayer is shown. Again, Hst5^{rand1} displayed the closest distance in comparison to the other peptides. For all peptides, an arginine residue, located at different sequence positions for the three peptides, was found at closest contact with the bilayer. For Hst5^{rand1} the interacting arginine was in position 16, while it was in position 12 for Hst5. The difference between Hst5 and Hst5^{rand1} was minor, where Hst5 was only slightly further away from the bilayer on average. In the density profile there was again an overlap of Hst5 density with the lipid head group. Lastly, Hst5^{rand2} was on average the furthest away from the bilayer, where the arginine residue at position 21 was at closest contact with the phosphorus atom in the bilayer. It was also clear from its density profile that the overlap between the head group region and the Hst5^{rand2} peptide was smaller compared to the other two peptides.

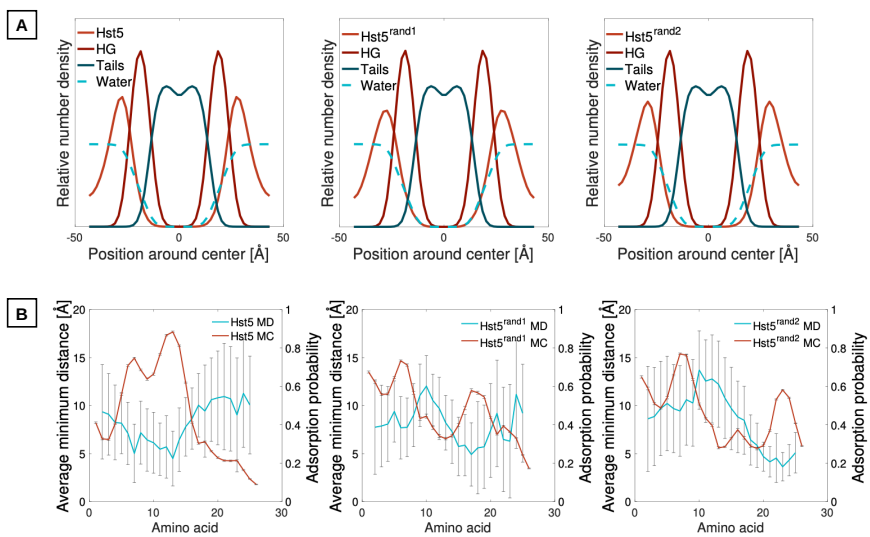


Figure 6.13: Molecular dynamics results obtained for the three peptides Hst5, Hst5^{rand1}, and Hst5^{rand2} displaying **(A)** number density profiles normalized by the number of atoms in each group. The density is centered around the bilayer. **(B)** Average distance from the bilayer for each residue, as well as its standard deviation. The calculation is performed from the change in minimum distance between each pair over time for the concatenated trajectory of the replicates.

To elucidate the origin of the interaction between the peptides and the bilayer the average minimum distance obtained from MD simulations and the adsorption probability per amino acid obtained from MC simulations are compared in Figure 6.13B. Since the MC simulations mainly captured the electrostatic interactions, this comparison indicated that the peptide interaction with the bilayer was mainly of electrostatic origin, as the adsorption profile follows the shape of the average minimum distance plot; the amino acids with short distance to the PC_9PS_1 bilayer observed by atomistic MD had a high adsorption probability observed by coarse-grained MC simulations. The error bars obtained for the MD data displayed in Figure 6.13B are very large and represents the change in minimum distance between each amino acid and the phosphorus atom of the bilayer, hence, it displays the movement of the amino acids around the average closest position with respect to the bilayer.

6.5 Interaction of KEIF with solid surfaces and phospholipid bilayers

The last study included in this thesis, is about the intrinsically disordered N-terminal region of MgtA, referred to as KEIF, and its interaction with both solid surfaces, as well as lipid bilayers. This peptide was used as a model system to investigate the structure-function relationship of surface active IDPs, meaning that they adopt a structure upon adsorption and thereby obtain their function. Therefore, both the adsorption properties of KEIF to a solid silica surface and lipid bilayers were investigated using NR, QCM-D, as well as atomistic MD simulations, while the secondary structure was evaluated in solution, adsorbed to vesicles, as well as adsorbed on a quartz surface using CD and OCD. This investigation resulted in Paper IV, with the main results that KEIF became more structured upon adsorption to surfaces. This change of structure allowed the peptide to enter a lipid bilayer, where it preferably resided within the hydrophobic tail region. This finding allowed us to to speculate that this part of the MgtA protein is submerged in the plasma membrane, where it possibly has a modulatory effect on the function of MgtA.

6.5.1 The interaction of KEIF with lipid bilayers

NR and QCM-D measurements were conducted to investigate the interaction of KEIF with SLBs of two different compositions, purely zwitterionic POPC and negatively charged PC_9PS_1. The NR measurements were performed at only 10 mM NaCl, while the QCM-D measurements were performed at both 10 and 150 mM NaCl. These experiments were complemented with atomistic MD simulations to obtain more details

on the interactions. Through analysis of the NR data, the peptide was found to reside in the hydrophobic tail region of the bilayers for both lipid compositions, as is shown in the SLD profiles in Figure 6.14. Slightly more KEIF and solvent were found in the PC_9PS_1 bilayer. The QCM-D data showed that for all samples there was a splitting between the normalized frequency overtones upon injection and incubation of KEIF, indicative of a non-rigid layer on top of the surface. Upon rinsing with pristine buffer, the harmonics were again overlapping, indicative of a more rigid adsorbed layer. In all systems, except the one where KEIF was injected to a PC_9PS_1 bilayer in 10 mM NaCl, the normalized frequency converged to positive values, indicative of lipid removal. The dissipation values for all samples, except KEIF-PC_9PS_1, returned to zero upon rinsing, meaning that they had the same rigidity as the pristine SLB.

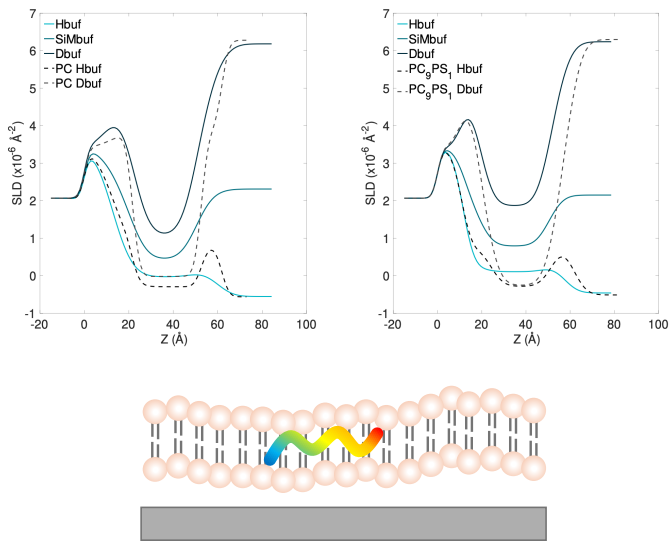


Figure 6.14: The SLD profiles obtained from the analysis of the NR data, together with a schematic figure depicting the position of KEIF with respect to the bilayer at 10 mM NaCl.

A clear difference between the bilayer compositions was observed during incubation of the peptide: for POPC in both 10 and 150 mM NaCl, the normalized frequencies and dissipation shift remained stable, while for the PC_9PS_1 SLBs in both 10 and 150 mM NaCl, the dissipation shifts were slowly decreasing, while the frequency shifts were stable. The decrease in dissipation while the frequency is constant indicates time-dependent rearrangements where the system became more rigid without adding or removing any material in the process, which likely stemmed from the peptide moving from the top of the bilayer towards the tail region of the bilayer. The lack of this transition in the POPC systems indicated that KEIF interacted with the negatively charged head groups for a longer time before it was incorporated within the bilayer, suggesting

a stronger electrostatic interaction between the peptide and the head groups.

The preferred adsorption to the hydrophobic region was further supported by QCM-D measurements where KEIF adsorbed to both hydrophilic and hydrophobic silica surfaces. These measurements initially showed more adsorption to the hydrophilic silica surface, while more peptide remained on the hydrophobic surface upon rinsing with pristine buffer. These findings together deduce that the electrostatic interactions play an initial role when KEIF adsorbs to the bilayer surface, followed by a stronger preference for the hydrophobic region of the bilayer, where it finally resides.

Using atomistic MD simulations, it was found that an arginine residue at position 8 (for POPC, 10 mM NaCl), or an arginine residue at position 16 (for the other systems), was on average, in closest contact with the bilayer. Analyzing both the minimum distance and the number density around the bilayer, for all systems studied, KEIF was found in closest contact with the POPC bilayer in 10 mM NaCl.

6.5.2 The structure of KEIF

The structure of KEIF was investigated in solution, both alone, and in the vicinity of lipid vesicles, as well as in a film, using CD and OCD. In solution, KEIF was found, as previously reported,[51,52] to be mainly disordered in solution with some of β-sheets. In the vicinity of charged lipid vesicles, PC_9PS_1 or PC_3PS_1, the peptide became more ordered, increasing the α-helical structure, as shown in Figure 6.15. It was observed that the peptide became more ordered both as the lipid charge increased and as the lipid to peptide ratio increased. These results indicated that the electrostatics were crucial for the interaction, as opposed to QCM-D and NR, where the effect was much smaller. This could be because CD data was measured directly after mixing, which showed only the initial interaction of the peptide with the lipids. It was demonstrated with QCM-D that the interaction with the charged bilayer was altered in a time-dependent manner, implying that the CD measurements were conducted while the peptide was still adsorbed to the LUV surface, rather than residing in the hydrophobic tail region.

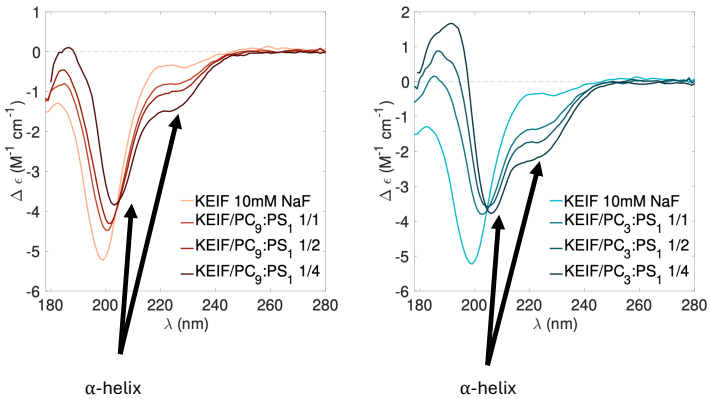


Figure 6.15: CD spectra of KEIF and PC_9PS_1 (left) at peptide:lipid ratios 1:1, 1:2, and 1:4, and PC_3PS_1 (right) at peptide:lipid ratios 1:1, 1:2, and 1:4. In both cases the positions indicating α-helical structure is indicated.

The increased proportion of structured elements upon adsorption supports the commonly adapted hypothesis that surface active IDPs gain secondary structure upon adsorption, as the increase in ordered secondary structure protects the hydrophilic amino acids in the sequence from unfavorable interactions with the hydrophobic part of the cell membrane.

OCD measurements were conducted to investigate the structure of KEIF in a film, which showed more structure compared to the samples measured in solution. It was also found that the peptide organized into β-structure rather than α-helical structure, and that it was partially inserted into the lipid membrane, according to the description presented in the paper by Bürck *et al.* [82]

The scientist is not a person who gives the right answers, he's one who asks the right questions.

- Claude Levi-Strauss

7 | Conclusions and outlook

In an effort to understand the mechanism behind the antimicrobial effect of Hst5, four studies were presented in this thesis, investigating different environmental conditions, as well as characteristics of the peptide sequence. The formation of a peptide cushion in the gap region between an SLB and the solid substrate is used as a measure of the peptides antimicrobal effect, as that mimics its known route of action. Electrostatic interactions have been shown to be of great importance in the mechanism, both for the initial adsorption of the peptide to the bilayer, but also for its ability to translocate across the bilayer, where a negative surface below the bilayer was required (mimicking the potential of the mitochondria in a cell). The ability to charge regulate, provided by the many histidine residues in the sequence, were proven to be important as decreasing the number of histidine residues in the sequence no longer allowed the peptide to translocate the bilayer without disrupting its integrity. In this study, the postition of the histidine residues was speculated to be of importance, however, the results obtained for Hst5 variants with randomized sequence could not verify this hypothesis. Small differences were observed compared to the Hst5 peptide: during cushion formation with Hst5, the bilayers integrity was maintained. For both of the randomized variants, the peptide was found to reside in the head groups of the inner leaflet, in addition to forming a cushion. Further studies are required to get insight into why these differences are observed. Altering the length of the peptide did not display large differences at low salt concentration, except that a thinner cushion was formed for the longer variant. At high salt concentration however, both the shorter and longer variant were able to interact with the bilayer, while Hst5 was did not display any interaction at all. This is hypothesized to depend on electrostatic interactions where the overall charge of the peptides are important. In the length study, charged patches containing lysine and arginine were identified as important for interaction. NR measurements of variants where the arginine residues have been replaced by lysine residues have been performed, but not yet analyzed. These measurements could give important answers if it is the specific interaction of arginine, or the general positive charge that is of importance. Therefore, analyzing this data is the obvious next step in the investigations regarding

the interaction of Hst5 with a model membrane system.

To continue the investigations of the antimicrobial effect of Hst5, studies with zinc have already been conducted,[140,143] and it would be interesting to conduct the investigations focusing more on the mechanism behind the interaction of Hst5 with bilayers in the presence of zinc. Furthermore, it would also be interesting to investigate the concentration dependence on the interaction mechanism, since it was found by Mochon and Liu,[37] that the antifungal effect is dependent on concentration. It could possibly change from targeting intracellular processes to rupturing the membrane.

The investigation of the N-terminal region of MgtA, referred to as KEIF confirmed the suggested hypothesis that this disordered, surface active peptide does obtain structure upon adsorption to a surface, allowing it to enter the bilayer and reside in the hydrophobic tail region. This is possibly where KEIF resides in the plasma membrane, and could have a modulatory effect on the MgtA protein. It would be beneficial to conduct an additional study to investigate the interaction of KEIF with bilayers composed of lipids commonly found in the bacterial membrane, rather than in eukaryotic, as is the case for the lipids used here. Additionally, to better understand the function of KEIF, conducting a study of MgtA in a SLB with (as the innate protein) or without KEIF (the N-terminal region) could give more information of its function.

References

[1] Federici, B. A. Chapter 193 - pathogens of insects. In Resh, V. H. & Cardé, R. T. (eds.) *Encyclopedia of Insects (Second Edition)*, 757–765 (Academic Press, 2009), second edition edn.

[2] National Institutes of Health (U.S.) and National Institute of Allergy and Infectious Diseases (U.S.). *Understanding Microbes in Sickness and in Health*. NIH publication (U.S. Department of Health and Human Services, National Institutes of Health and National Institute of Allergy and infectious Diseases, 2006).

[3] Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of b. influenzae. *British journal of experimental pathology* **10**, 226–236 (1929).

[4] Abraham, E., Chain, E., Fletcher, C., Gardner, A., Heatley, N., Jennings, M. & Florey, H. Further observations on penicillin. *The Lancet* **238**, 177–189 (1941). Originally published as Volume 2, Issue 6155.

[5] Hutchings, M. I., Truman, A. W. & Wilkinson, B. Antibiotics: past, present and future. *Current Opinion in Microbiology* **51**, 72–80 (2019).

[6] WHO, Antimicrobial Resistance. `https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance`. Accessed: 2024-08-26.

[7] World Health Organization. Antimicrobial resistance: global report on surveillance. `https://www.who.int/publications/i/item/9789241564748` (2014). Accessed: 2024-08-26.

[8] World Health Organization. Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2022. `https://www.who.int/publications/i/item/9789240062702` (2022). Accessed: 2024-08-26.

[9] World Health Organization. WHO fungal priority pathogens list to guide research, development and public health action. `https://www.who.int/publications/i/item/9789240060241` (2022). Accessed: 2024-08-26.

[10] Talapko, J., Juzbašić, M., Matijević, T., Pustijanac, E., Bekić, S., Kotris, I. & Škrlec, I. Candida albicans—the virulence factors and clinical manifestations of infection. *Journal of Fungi* 7 (2021).

[11] Oppenheim, F. G., Xu, T., McMillian, F. M., Levitz, S. M., Diamond, R. D., Offner, G. D. & Troxler, R. F. Histatins, a novel family of histidine-rich proteins in human

parotid secretion. isolation, characterization, primary structure, and fungistatic effects on candida albicans. *Journal of Biological Chemistry* **263**, 7472–7477 (1988).

[12] Troxler, R., Offner, G., Xu, T., Vanderspek, J. & Oppenheim, F. Structural relationship between human salivary histatins. *Journal of Dental Research* **69**, 2–6 (1990).

[13] Tsai, H. & Bobek, L. Human salivary histatins: Promising anti-fungal therapeutic agents. *Critical Reviews in Oral Biology & Medicine* **9**, 480–497 (1998).

[14] Edgerton, M., Koshlukova, S. E., Lo, T. E., Chrzan, B. G., Straubinger, R. M. & Raj, P. A. Candidacidal activity of salivary histatins. identification of a histatin 5-binding protein on candida albicans. *The Journal of biological chemistry* **273**, 20438–20447 (1998).

[15] Subramani, S., Perdreau-Dahl, H. & Morth, J. P. The magnesium transporter a is activated by cardiolipin and is highly sensitive to free magnesium in vitro. *eLife* **5**, e11407 (2016).

[16] Berg, J. M., Tymoczko, J. L., Gatto, G. J. & Stryer, L. *Biochemistry* (Freeman, 2015).

[17] van der Lee, R. *et al.* Classification of intrinsically disordered regions and proteins. *Chemical Reviews* **114**, 6589–6631 (2014).

[18] Uversky, V. N., Oldfield, C. J., Midic, U., Xie, H., Xue, B., Vucetic, S., Iakoucheva, L. M., Obradovic, Z. & Dunker, A. K. Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics* **10** (2009).

[19] Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annual Review of Biochemistry* **83**, 553–584 (2014).

[20] Romero, P., Obradovic, Z., Li, X., Garner, E., Brown, C. & Dunker, A. Sequence complexity of disordered protein. *Proteins: Structure, Function and Genetics* **42**, 38–48 (2001).

[21] Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Structure, Function, and Bioinformatics* **41**, 415–427 (2000).

[22] Tompa, P. Intrinsically unstructured proteins. *Trends in Biochemical Sciences* **27**, 527–533 (2002).

[23] Axel, H., Melina, M., Patricia, M., Liliana C., S. & Paulo C., M. Antimicrobial peptides: Interaction with model and biological membranes and synergism with chemical antibiotics. *Frontiers in Chemistry* **6** (2018).

[24] Andersson, D., Hughes, D. & Kubicek-Sutherland, J. Mechanisms and consequences of bacterial resistance to antimicrobial peptides. *Drug Resistance Updates* **26**, 43–57 (2016).

[25] Huan, Y., Kong, Q., Mou, H. & Yi, H. Antimicrobial peptides: Classification, design, application and research progress in multiple fields. *Frontiers in Microbiology* **11** (2020).

[26] Helmerhorst, E. J., van't Hof, W., Breeuwer, P., Veerman, E. C., Abee, T., Troxler, R. F., Amerongen, A. V. & Oppenheim, F. G. Characterization of histatin 5 with respect to amphipathicity, hydrophobicity, and effects on cell and mitochondrial membrane integrity excludes a candidacidal mechanism of pore formation. *The Journal of biological chemistry* **276**, 5643–5649 (2001).

[27] Lund, M. & Jönsson, B. Charge regulation in biomolecular solution. *Quarterly Reviews of Biophysics* **46**, 265—281 (2013).

[28] Kurut, A., Henriques, J., Forsman, J., Skepö, M. & Lund, M. Role of histidine for charge regulation of unstructured peptides at interfaces and in bulk. *Proteins: Structure, Function, and Bioinformatics* **82**, 657–667 (2014).

[29] Tsai, H., Raj, P. A. & Bobek, L. A. Candidacidal activity of recombinant human salivary histatin-5 and variants. *Infection and Immunity* **64**, 5000–5007 (1996).

[30] Tsai, H. & Bobek, L. A. Studies of the mechanism of human salivary histatin-5 candidacidal activity with histatin-5 variants and azole-sensitive and -resistant candida species. *Antimicrobial Agents and Chemotherapy* **41**, 2224–2228 (1997).

[31] Koshlukova, S. E., Lloyd, T. L., Araujo, M. W. & Edgerton, M. Salivary histatin 5 induces non-lytic release of atp from candida albicans leading to cell death. *The Journal of biological chemistry* **274**, 18872–18879 (1999).

[32] den Hertog, A., van Marle, J., van Veen, H., van't Hof, W., Bolscher, J., Veerman, E. & Nieuw Amerongen, A. Candidacidal effects of two antimicrobial peptides: histatin 5 causes small membrane defects, but LL-37 causes massive disruption of the cell membrane. *Biochemical Journal* **388**, 689–695 (2005).

[33] Helmerhorst, E. J., Breeuwer, P., van 't Hof, W., Walgreen-Weterings, E., Oomen, L. C., Veerman, E. C., Amerongen, A. V. N. & Abee, T. The cellular target of histatin 5 on candida albicans is the energized mitochondrion. *Journal of Biological Chemistry* **274**, 7286–7291 (1999).

[34] Ruissen, A., Groenink, J., Van 't Hof, W., Walgreen-Weterings, E., van Marle, J., van Veen, H., Voorhout, W., Veerman, E. & Nieuw Amerongen, A. Histatin 5 and derivatives: Their localization and effects on the ultra-structural level. *Peptides* **23**, 1391–1399 (2002).

[35] Li, X. S., Reddy, M. S., Baev, D. & Edgerton, M. Candida albicans ssa1/2p is the cell envelope binding protein for human salivary histatin 5. *The Journal of biological chemistry* **278**, 28553–28561 (2003).

[36] den Hertog, A. L., Sang, H. W. W. F., Kraayenhof, R., Bolscher, J. G. M., Hof, W. V., Veerman, E. C. I. & Amerongen, A. V. N. Interactions of histatin 5 and histatin 5-derived peptides with liposome membranes: surface effects, translocation and permeabilization. *Biochemical Journal* **379**, 665–672 (2004).

[37] Mochon, A. B. & Liu, H. The antimicrobial peptide histatin-5 causes a spatially restricted disruption on the candida albicans surface, allowing rapid entry of the peptide into the cytoplasm. *PLOS Pathogens* **4**, 1–12 (2008).

[38] Cragnell, C., Staby, L., Lenton, S., Kragelund, B. B. & Skepö, M. Dynamical oligomerisation of histidine rich intrinsically disordered proteins is regulated through zinc-histidine interactions. *Biomolecules* **9** (2019).

[39] Conklin, S. E., Bridgman, E. C., Su, Q., Riggs-Gelasco, P., Haas, K. L. & Franz, K. J. Specific histidine residues confer histatin peptides with copper-dependent activity against candida albicans. *Biochemistry* **56**, 4244–4255 (2017).

[40] Grogan, J., McKnight, C., Troxler, R. F. & Oppenheim, F. G. Zinc and copper bind to unique sites of histatin 5. *FEBS Letters* **491**, 76–80 (2001).

[41] Fagerberg, E., Månsson, L. K., Lenton, S. & Skepö, M. The effects of chain length on the structural properties of intrinsically disordered proteins in concentrated solutions.

The Journal of Physical Chemistry B **124**, 11843–11853 (2020).

[42] Raj, P. A., Edgerton, M. & Levine, M. J. Salivary histatin 5: dependence of sequence, chain length, and helical conformation for candidacidal activity. *Journal of Biological Chemistry* **265**, 3898–3905 (1990).

[43] Ruissen, A., Groenink, J., Krijtenberg, P., Walgreen-Weterings, E., van't Hof, W., Veerman, E. & Amerongen, A. N. Internalisation and degradation of histatin 5 by candida albicans. *Biological Chemistry* **384**, 183–190 (2003).

[44] Veerman, E. C., Valentijn-Benz, M., Nazmi, K., Ruissen, A. L., Walgreen-Weterings, E., van Marle, J., Doust, A. B., van't Hof, W., Bolscher, J. G. & Amerongen, A. V. N. Energy depletion protects candida albicans against antimicrobial peptides by rigidifying its cell membrane. *Journal of Biological Chemistry* **282**, 18831–18841 (2007).

[45] Welling, M. M., Brouwer, C. P. J. M., van ☒t Hof, W., Veerman, E. C. I. & Amerongen, A. V. N. Histatin-derived monomeric and dimeric synthetic peptides show strong bactericidal activity towards multidrug-resistant staphylococcus aureus in vivo. *Antimicrobial Agents and Chemotherapy* **51**, 3416–3419 (2007).

[46] Richardson, C., Johnsson, M., Raj, P., Levine, M. & Nancollas, G. The influence of histatin-5 fragments on the mineralization of hydroxyapatite. *Archives of Oral Biology* **38**, 997–1002 (1993).

[47] Lenton, S., Hervø-Hansen, S., Popov, A. M., Tully, M. D., Lund, M. & Skepö, M. Impact of arginine–phosphate interactions on the reentrant condensation of disordered proteins. *Biomacromolecules* **22**, 1532–1544 (2021).

[48] Maguire, M. E. Magnesium transporters: Properties, regulation and structure. *FBL* **11**, 3149–3163 (2006).

[49] Kehres, D. & Maguire, M. Structure, properties and regulation of magnesium transport proteins. *BioMetals* **15**, 261–270 (2002).

[50] Håkansson, K. O. The structure of mg-atpase nucleotide-binding domain at 1.6åresolution reveals a unique atp-binding motif. *Acta Crystallographica Section D* **65**, 1181–1186 (2009).

[51] Jephthah, S., Månsson, L. K., Belić, D., Morth, J. P. & Skepö, M. Physicochemical characterisation of keif—the intrinsically disordered n-terminal region of magnesium transporter a. *Biomolecules* **10**, 623 (2020).

[52] Koder Hamid, M., Månsson, L. K., Meklesh, V., Persson, P. & Skepö, M. Molecular dynamics simulations of the adsorption of an intrinsically disordered protein: Force field and water model evaluation in comparison with experiments. *Frontiers in Molecular Biosciences* **9** (2022).

[53] Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society, Faraday Transactions 2* **72**, 1525–1568 (1976).

[54] Evans, D. F. & Wennerström, H. *The colloidal domain : where physics, chemistry, biology, and technology meet* (Wiley-VCH, 1999).

[55] Kučerka, N., Nieh, M.-P. & Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1808**, 2761–2771 (2011).

[56] Cooper, G. *The Cell: A Molecular Approach* (ASM Press, 2000).

[57] Akbarzadeh, A., Rezaei-Sadabady, R., Davaran, S., Joo, S. W., Zarghami, N., Hanifehpour, Y., Samiei, M., Kouhi, M. & Nejati-Koshki, K. Liposome: classification, preparation, and applications. *Nanoscale Research Letters* **8**, 1–9 (2013).

[58] Rideau, E., Dimova, R., Schwille, P., Wurm, F. R. & Landfester, K. Liposomes and polymersomes: a comparative review towards cell mimicking. *Chemical Society Reviews* **47**, 8572–8610 (2018).

[59] Clifton, L. A., Campbell, R. A., Sebastiani, F., Campos-Terán, J., Gonzalez-Martinez, J. F., Björklund, S., Sotres, J. & Cárdenas, M. Design and use of model membranes to study biomolecular interactions using complementary surface-sensitive techniques. *Advances in Colloid and Interface Science* **277**, 102118 (2020).

[60] Tanaka, M. & Sackmann, E. Polymer-supported membranes as models of the cell surface. *Nature* **437**, 656–663 (2005).

[61] Majewski, J., Wong, J., Park, C., Seitz, M., Israelachvili, J. & Smith, G. Structural studies of polymer-cushioned lipid bilayers. *Biophysical Journal* **75**, 2363–2367 (1998).

[62] Wong, J., Majewski, J., Seitz, M., Park, C., Israelachvili, J. & Smith, G. Polymer-cushioned bilayers. i. a structural study of various preparation methods using neutron reflectometry. *Biophysical Journal* **77**, 1445–1457 (1999).

[63] Hill, T. L. *An introduction to statistical thermodynamics* (Dover Publications, 1986).

[64] Israelachvili, J. *Intermolecular and Surface Forces*. Academic Press (Academic Press, 2011).

[65] Debye, P. & Hückel, E. Zur theorie der elektrolyte. i. gefrierpunktserniedrigung und verwandte erscheinungen. *Physikalische Zeitschrift* **24**, 305 (1923).

[66] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. *Protein Identification and Analysis Tools on the ExPASy Server*, 571–607 (Humana Press, 2005).

[67] Anthis, N. J. & Clore, G. M. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Science* **22**, 851–858 (2013).

[68] Avanti Research. `https://avantiresearch.com`. Accessed: 2024-10-16.

[69] Kalb, E., Frey, S. & Tamm, L. K. Formation of supported planar bilayers by fusion of vesicles to supported phospholipid monolayers. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1103**, 307–316 (1992).

[70] Cremer, P. S. & Boxer, S. G. Formation and spreading of lipid bilayers on planar glass supports. *The Journal of Physical Chemistry B* **103**, 2554–2559 (1999).

[71] Wacklin, H. P. Composition and asymmetry in supported membranes formed by vesicle fusion. *Langmuir* **27**, 7698–7707 (2011).

[72] Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1751**, 119–139 (2005).

[73] Miles, A. & Wallace, B. Chapter 6 - circular dichroism spectroscopy for protein characterization: Biopharmaceutical applications. In Houde, D. J. & Berkowitz, S. A. (eds.) *Biophysical Characterization of Proteins in Developing Biopharmaceuticals*, 109–137 (Elsevier, 2015).

[74] Whitmore, L., Woollett, B., Miles, A., Wallace, B., Klose, D. & Janes, R. Pcddb: The

protein circular dichroism data bank, a repository for circular dichroism spectral and metadata. *Nucleic Acids Research* **39**, D480–D486 (2011).

[75] Abdul-Gader, A., Miles, A. J. & Wallace, B. A. A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics* **27**, 1630–1636 (2011).

[76] Tolchard, J., Walpole, S., Eaglen, L., Blumenschein, T., Miles, A., Wallace, B., Maytum, R. & Hackstadt, T. The intrinsically disordered tarp protein from chlamydia binds actin with a partially preformed helix. *Scientific Reports* **8** (2018).

[77] Sreerama, N. & Woody, R. W. Computation and analysis of protein circular dichroism spectra. In *Numerical Computer Methods, Part D*, vol. 383 of *Methods in Enzymology*, 318–351 (Academic Press, 2004).

[78] Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M. & Kardos, J. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E3095–E3103 (2015).

[79] Micsonai, A. *et al.* Bestsel: webserver for secondary structure and fold prediction for protein cd spectroscopy. *Nucleic Acids Research* **50**, W90–W98 (2022).

[80] Miles, A. J., Ramalli, S. G. & Wallace, B. A. Dichroweb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Langmuir* **31**, 37 (2022).

[81] Moffitt, W., Fitts, D. D. & Kirkwood, J. G. Critique of the theory of optical activity of helical polymers. *Proceedings of the National Academy of Sciences of the United States of America* **43**, 723–730 (1957).

[82] Bürck, J., Wadhwani, P., Fanghänel, S. & Ulrich, A. Oriented circular dichroism: A method to characterize membrane-active peptides in oriented lipid bilayers. *Accounts of Chemical Research* **49**, 184–192 (2016).

[83] Rodger, A. *Oriented Circular Dichroism Spectroscopy*, 1808–1809 (Springer Berlin Heidelberg, 2013).

[84] Svergun, D. I. & Koch, M. H. J. Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics* **66**, 1735–1782 (2003).

[85] Svergun, D. I., Koch, M. H. J., Timmins, P. A. & May, R. P. *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules.* (Oxford University Press, 2013).

[86] Receveur-Brechot, V. & Durand, D. How random are intrinsically disordered proteins? a small angle scattering perspective. *Current Protein & Peptide Science* **13**, 55–75 (2012).

[87] Jacques, D. A. & Trewhella, J. Small-angle scattering for structural biology - expanding the frontier while avoiding the pitfalls. *Protein Science* **19**, 642–657 (2010).

[88] Schnablegger, H. & Singh, Y. *The SAXS Guide: Getting acquainted with the principles* (Austria: Anton Paar GmbH, 2013).

[89] Manalastas-Cantos, K. *et al.* ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *Journal of Applied Crystallography* **54**, 343–355 (2021).

[90] Easley, A. D., Ma, T., Eneh, C. I., Yun, J., Thakur, R. M. & Lutkenhaus, J. L. A practical guide to quartz crystal microbalance with dissipation monitoring of thin polymer films.

Journal of Polymer Science **60**, 1090–1107 (2022).

[91] Laatikainen, M. & Lindström, M. Determination of adsorption isotherms with quartz crystal microbalance in liquid phase. *Journal of Colloid and Interface Science* **125**, 610–614 (1988).

[92] Höök, F., Rodahl, M., Brzezinski, P. & Kasemo, B. Energy dissipation kinetics for protein and antibody–antigen adsorption under shear oscillation on a quartz crystal microbalance. *Langmuir* **14**, 729–734 (1998).

[93] Richter, R., Mukhopadhyay, A. & Brisson, A. Pathways of lipid vesicle deposition on solid surfaces: A combined qcm-d and afm study. *Biophysical Journal* **85**, 3035–3047 (2003).

[94] Höök, F. & Kasemo, B. *The QCM-D Technique for Probing Biomacromolecular Recognition Reactions*, 425–447 (Springer Berlin Heidelberg, 2007).

[95] Sauerbrey, G. Verwendung von schwingquarzen zur wägung dünner schichten und zur mikrowägung. *Zeitschrift für Physik* **155**, 206–222 (1959).

[96] Höök, F. *Development of a novel QCM technique for protein adsorption studies.* Ph.D. thesis (1997).

[97] Dixon, M. Quartz crystal microbalance with dissipation monitoring: Enabling real-time characterization of biological materials and their interactions. *Journal of Biomolecular Techniques* **19**, 151–158 (2008).

[98] Cubitt, R. & Fragneto, G. Chapter 2.8.3 - neutron reflection: Principles and examples of applications. In Pike, R. & Sabatier, P. (eds.) *Scattering*, 1198–1208 (Academic Press, 2002).

[99] Fermi, E. & Zinn, W. Reflection of neutrons on mirrors. *Physical Review* **70**, 103 (1946).

[100] Fragneto-Cusani, G. Neutron reflectivity at the solid/liquid interface: examples of applications in biophysics. *Journal of Physics: Condensed Matter* **13**, 4973–4989 (2001).

[101] Crowley, T., Lee, E., Simister, E. & Thomas, R. The use of contrast variation in the specular reflection of neutrons from interfaces. *Physica B: Condensed Matter* **173**, 143–156 (1991).

[102] Parratt, L. G. Surface studies of solids by total reflection of x-rays. *Physical Review* **95**, 359–369 (1954).

[103] Gutfreund, P., Saerbeck, T., Gonzalez, M. A., Pellegrini, E., Laver, M., Dewhurst, C. & Cubitt, R. Towards generalized data reduction on a chopper-based time-of-flight neutron reflectometer. *Journal of Applied Crystallography* **51**, 606–615 (2018).

[104] Gerelli, Y. Aurore: new software for neutron reflectivity data analysis. *Journal of Applied Crystallography* **49**, 330–339 (2016).

[105] Neutron scattering lengths and cross sections. (2021). `https://www.ncnr.nist.gov/resources/n-lengths/`.

[106] Nagle, J. F. & Tristram-Nagle, S. Structure of lipid bilayers. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1469**, 159–195 (2000).

[107] James, F. Minuit function minimization and error analysis: Reference manual version 94.1. (1994).

[108] Cragnell, C., Durand, D., Cabane, B. & Skepö, M. Coarse-grained modeling of the

intrinsically disordered protein histatin 5 in solution: Monte carlo simulations in combination with saxs. *Proteins: Structure, Function, and Bioinformatics* **84**, 777–791 (2016).

[109] Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. & Hutchison, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* **4**, 1–17 (2012).

[110] Jo, S., Kim, T., Iyer, V. G. & Im, W. Charmm-gui: A web-based graphical user interface for charmm. *Journal of Computational Chemistry* **29**, 1859–1865 (2008).

[111] Lee, J. *et al.* Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Journal of Chemical Theory and Computation* **12**, 405–413 (2016).

[112] Jo, S., Lim, J. B., Klauda, J. B. & Im, W. Charmm-gui membrane builder for mixed bilayers and its application to yeast membranes. *Biophysical Journal* **97**, 50–58 (2009).

[113] Jo, S., Kim, T. & Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLOS ONE* **2**, 1–9 (2007).

[114] Wu, E. L. *et al.* Charmm-gui membrane builder toward realistic biological membrane simulations. *Journal of Computational Chemistry* **35**, 1997–2004 (2014).

[115] Brooks, B. R. *et al.* Charmm: The biomolecular simulation program. *Journal of Computational Chemistry* **30**, 1545–1614 (2009).

[116] Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmüller, H. & MacKerell, J., Alexander D. Charmm36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **14**, 71–73 (2017).

[117] Berendsen, H., van der Spoel, D. & van Drunen, R. Gromacs: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43–56 (1995).

[118] Pronk, S. *et al.* Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Swedish e-Science Research Centre eSSENCE – An eScience Collaboration Bioinformatics* **29**, 845–854 (2013).

[119] Páll, S., Abraham, M. J., Kutzner, C., Hess, B. & Lindahl, E. Tackling exascale software challenges in molecular dynamics simulations with gromacs. *Solving Software Challenges for Exascale* 3–27 (2015).

[120] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B. & Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).

[121] Moss, G. P. Basic terminology of stereochemistry (iupac recommendations 1996). *Pure and Applied Chemistry* **68**, 2193–2222 (1996).

[122] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087–1092 (1953).

[123] Frenkel, D. & Smit, B. *Understanding Molecular Simulation. From algorithms to applications.* (Academic Press, 2002).

[124] Hockney, R., Goel, S. & Eastwood, J. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics* **14**, 148–158 (1974).

[125] Abraham, M. *et al.* Gromacs 2024.1 manual. (2024).

[126] Lindahl, E. *Molecular Dynamics Simulations*, 3–26 (Springer New York, 2015).

[127] Alder, B. J. & Wainwright, T. E. Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics* **31**, 459–466 (1959).

[128] Allen, M. P. & Tildesley, D. J. 95 Molecular dynamics. In *Computer Simulation of Liquids* (Oxford University Press, 2017).

[129] Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089–10092 (1993).

[130] Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *Journal of Chemical Theory and Computation* **11**, 1864–1874 (2015).

[131] Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. Lincs: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**, 1463–1472 (1997).

[132] Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics* **52**, 255–268 (1984).

[133] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **31**, 1695–1697 (1985).

[134] Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182–7190 (1981).

[135] Stepto, R., Chang, T., Kratochvíl, P., Hess, M., Horie, K., Sato, T. & Vohlídal, J. Definitions of terms relating to individual macromolecules, macromolecular assemblies, polymer solutions, and amorphous bulk polymers (iupac recommendations 2014). *Pure and Applied Chemistry* **87**, 71–120 (2015).

[136] Gorelov, S., Titov, A., Tolicheva, O., Konevega, A. & Shvetsov, A. Dssp in gromacs: Tool for defining secondary structures of proteins in trajectories. *Journal of Chemical Information and Modeling* **64**, 3593–3598 (2024).

[137] Grossfield, A. & Zuckerman, D. M. Chapter 2 quantifying uncertainty and sampling quality in biomolecular simulations. vol. 5 of *Annual Reports in Computational Chemistry*, 23–48 (Elsevier, 2009).

[138] Grossfield, A., Patrone, P. N., Roe, D. R., Schultz, A. J., Siderius, D. & Zuckerman, D. M. Best practices for quantification of uncertainty and sampling quality in molecular simulations [article v1.0]. *Living Journal of Computational Molecular Science* **1**, 5067 (2018).

[139] Puri, S., Li, R., Ruszaj, D., Tati, S. & Edgerton, M. Iron binding modulates candidacidal properties of salivary histatin 5. *Journal of Dental Research* **94**, 201–208 (2015).

[140] Norris, H. L., Kumar, R., Ong, C. Y., Xu, D. & Edgerton, M. Zinc binding by histatin 5 promotes fungicidal membrane disruption in c. albicans and c. glabrata. *Journal of Fungi* **6** (2020).

[141] Hyltegren, K., Nylander, T., Lund, M. & Skepö, M. Adsorption of the intrinsically disordered saliva protein histatin 5 to silica surfaces. a monte carlo simulation and ellipsometry study. *Journal of Colloid and Interface Science* **467**, 280–290 (2016).

[142] Hyltegren, K. & Skepö, M. Adsorption of polyelectrolyte-like proteins to silica surfaces and the impact of ph on the response to ionic strength. a monte carlo simulation and ellipsometry study. *Journal of Colloid and Interface Science* **494**, 266–273 (2017).

[143] Norris, H. L., Kumar, R. & Edgerton, M. A novel role for histatin 5 in combination with zinc to promote commensalism in c. albicans survivor cells. *Pathogens* **10** (2021).

Part B

Scientific publications

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903

ISBN: 978-91-8096-070-0

Computational Chemistry
Physical and Theoretical Chemistry
Department of Chemistry
Faculty of Science