

Molecular Imaging of Biopolymers

Enabled by

Electrospray-Ion Beam Deposition

THIS IS A TEMPORARY TITLE PAGE
It will be replaced for the final print by a version
provided by the service academique.

Thèse n. 7581 2017
présenté le 24 janvier 2017
à la Faculté des Sciences de Base
Laboratoire de Science à l'Échelle Nanométrique
Programme Doctoral en Science et Génie des Matériaux
École Polytechnique Fédérale de Lausanne

pour l'obtention du grade de Docteur ès Sciences
par

Sabine Abb



acceptée sur proposition du jury:

Prof. Dirk Grundler, président du jury
Prof. Klaus Kern, directeur de thèse
Prof. Michele Ceriotti, rapporteur
Prof. Giovanni Costantini, rapporteur
Prof. Florian Klappenberger, rapporteur

Lausanne, EPFL, 2017

Voll Blüten steht der Pfirsichbaum,
nicht jede wird zur Frucht,
sie schimmern hell wie Rosenschaum
durch Blau und Wolkenflucht.

Wie Blüten gehen Gedanken auf,
Hundert an jedem Tag
Lass blühen! Lass
dem Ding den Lauf!
Frag nicht nach dem Ertrag!

Es muss auch Spiel und Unschuld sein
Und Blütenüberfluss,
sonst wär die Welt uns viel zu klein
und Leben kein Genuss.

—Hermann Hesse

Abstract

Biomolecular functional complexes often serve as inspiration for the development of artificial molecular nanostructures. The functionality of the biomolecular structures is based on complex and adaptable binding motifs that have evolved over billions of years. The detailed understanding of the underlying interactions is pivotal for the rational design and development of new materials, but often difficult to achieve as it requires submolecular resolution. In this thesis, I show how the molecular imaging of biopolymers by scanning tunneling microscopy (STM) and low-energy electron holography (LEEH) contributes to deciphering the structures and the respective binding motifs. These high-resolution imaging methods only perform well on surfaces in ultra-high vacuum (UHV) and for chemically identified species. Thus, electrospray ion beam deposition (ES-IBD) is essential for the controlled deposition for non-volatile species in UHV.

The first part of the thesis focuses on the characterization of the disaccharides as smallest representatives of polysaccharides. Deposited by ES-IBD, the first high-resolution STM characterization of this molecular class on a surface in UHV is performed. The disaccharides sucrose and trehalose can be distinguished as they exhibit specific self-assembly motifs and are imaged differently. For sucrose, the two monosaccharide building blocks are resolved, highlighting the impact that high-resolution imaging by STM might have in the future for the clarification of unknown sequences by identifying branching points.

In the second part of the thesis, peptides are employed as building blocks for novel structures on the surface. A combined approach of STM imaging with modeling by molecular dynamics and DFT is applied to identify the binding motifs and the underlying specific interactions. Understanding of the binding motifs of the decapeptide angiotensin-I on Au(111) enables the design of a long-ranged ordered honeycomb network by manipulating the sequence to the octapeptide angiotensin-II.

Representing a further binding motif that is often employed in catalytically active peptides, peptide-metal complexes are investigated on the surface showing a signifi-

Abstract

cant influence on the peptide structure. Based on the long-range ordered network for the pure angiotensin-II, we find that the introduction of a metal coordination center makes the molecular structure more compact and thus reduces the intermolecular interaction sites yielding an assembly of reduced dimensionality.

In the last part, the proof of principle is presented that LEEH, empowered by the ES-IBD sample preparation, is able to image single folded proteins on free standing ultra-clean graphene and might therefore provide protein structures without averaging. To this end, native mass spectrometry is adopted to ES-IBD for the deposition of folded proteins (cytochrome C, bovine serum albumin) and protein complexes (hemoglobin).

In conclusion, the versatile deposition of biopolymers by ES-IBD enables UHV based single molecule imaging techniques for a range of different samples, starting from imaging the sequence and characterizing 2D nanostructures of biomolecules by STM to 3D folded protein complexes imaged by LEEH. High-resolution imaging opens new perspectives for structure determination, such as the branching points of polysaccharides and the 3D structure of single proteins. Moreover, a detailed understanding of binding motifs is crucial for the rational design of new nanostructures.

Key words: Electrospray, Electrospray-Ion Beam Deposition (ES-IBD), Soft-landing, Ultra-high vacuum, Mass spectrometry, Scanning Tunneling Microscopy (STM), Biopolymers, Peptides, Proteins, Disaccharides, Nanostructures, Self-assembly, Molecular imaging, free standing ultra-clean graphene, Low energy electron holography (LEEh).

Zusammenfassung

Biomolekulare Komplexe dienen häufig als Inspiration für synthetische Nanostrukturen. Ihre Funktionalität, die sich über Milliarden von Jahren entwickelt hat, verfügt über hochkomplexe und anpassungsfähige Bindungsmotive. Das Verständnis der zugrundeliegenden Wechselwirkungen ist ausschlaggebend um neue Strukturen gezielt zu entwickeln. Dies gestaltet sich jedoch häufig schwierig, da submolekulare Auflösung nötig ist. In dieser Arbeit zeige ich, wie hochauflösendes Abbilden von Biopolymeren mittels Rastertunnelmikroskopie (engl: scanning tunneling microscopy: STM) und niederenergetische Elektronenholographie (engl: low energy electron holography: LEEH) dazu beitragen, die Struktur und ihre Bindungsmotive zu ermitteln. Da diese Methoden für chemisch identifizierte Moleküle geeignet sind und vor allem im Ultrahochvakuum (UHV) hohe Auflösung erzielen, verwenden wir Elektrospray-Ionenstrahldeposition (engl: electrospray-ion beam deposition, ES-IBD). Diese Depositionsmethode ermöglicht es, nicht flüchtige Verbindungen ins UHV zu leiten und massenselektiert auf einer Oberfläche zu deponieren.

Der erste Teil dieser Arbeit beschäftigt sich mit der Charakterisierung von Disacchariden, die exemplarisch für Polysaccharide stehen. Die ersten hochauflösenden STM Untersuchungen dieser Molekülklasse, die durch Deposition negativer Ionen möglich sind, werden gezeigt. Die Disaccharide Saccharose und Trehalose können anhand verschiedener Assemblierungsmotive und ihrer Abbildung unterschieden werden. So können die zwei Einzelzuckerbausteine der Saccharose aufgelöst abgebildet werden. Dieses Ergebnis deutet an, welchen Einfluss die hochaufgelösten STM-Messungen in Zukunft durch die Identifizierung von Verzweigungspunkten auf die Sequenzierung von Polysacchariden haben könnten.

Der zweite Teil der Arbeit zeigt auf, wie Peptide als Bausteine für neuartige Strukturen auf der Oberfläche verwendet werden können. Um die Bindungsmotive zu identifizieren ist eine Vergleich der STM Daten mit simulierten Modellen nötig, die auf einer Kombination aus Molekulardynamik (MD) und Dichtefunktionaltheorie (DFT) bestehen. Die Kenntnis der Bindungsmotive in den auf Angiotensin-I (10 Aminosäuren)

Zusammenfassung

basierenden Strukturen auf Au(111) ermöglicht es, durch die Sequenzveränderung hin zu dem Octapeptid Angiotensin-II ein langreichweitiges wabenförmiges Netzwerk zu realisieren.

Zusätzlich wird die Koordinierung zu Metallzentren untersucht, die ein weiteres Bindungsmotiv darstellen, das oft bei katalytisch aktiven Peptiden gefunden wird. Die Koordinierung von Peptiden zu Metallzentren hat einen starken Einfluss auf die Strukturen auf der Oberfläche: Ausgehend von dem vorgestellten Angiotensin-II Netzwerk finden wir, dass die Koordinierung an Metallzentren die 2D Struktur des Peptids auf der Oberfläche kompakter macht und dadurch die Stellen für Wechselwirkungen zwischen den Molekülen verringert. Dies führt zu einer Abnahme der Dimensionalität der Nanostrukturen.

Im letzten Teil der Arbeit zeige ich, dass die kontrollierte Deponierung durch ES-IBD es ermöglicht, einzelne gefaltete Proteine auf freistehendem, ultrasauberen Graphen mit LEEH abzubilden. Dadurch ist prinzipiell die Strukturaufklärung für Proteine ohne Mittelung möglich. Für die Deponierung von gefalteten Proteinen (Cytochrom C und Rinderalbumin BSA) werden die Methoden der nativen Massenspektrometrie auf unsere ES-IBD Anlage angewandt.

Die kontrollierte Abscheidung von vielfältigen Biopolymeren mithilfe ES-IBD ermöglicht es, eine Reihe von unterschiedlichen Proben mittels Einzelmolekül-Abbildungstechniken zu untersuchen, angefangen von dem Abbilden der Sequenz und der Charakterisierung von 2D Biomolekülstrukturen mithilfe von STM, bis hin zur Abbildung von 3D gefalteten Proteinkomplexen durch LEEH. Hochaufgelöstes Abbilden einzelner (Bio)moleküle eröffnet neue Möglichkeiten zur Strukturaufklärung, wie zum Beispiel die Identifizierung von Verzweigungspositionen in Polysacchariden. Desweiteren wird dadurch das Verständnis der Bindungsmotive vertieft und eine gezielte Entwicklung neuer (Nano)strukturen voran getrieben.

Stichwörter: Electrospray, Elctrospray-Ionenstrahl Deposition (ES-IBD), Ultrahochvakuum (UHV), Massenspektrometrie, Rastertunnelmikroskopie (STM), Biopolymere, Peptide, Proteine, Disaccharide, Nanostrukturen, Selbstassemblierung, freistehendes ultrasauberes Graphen, Niederenergetische Elektronenholografie (LEEh).

Contents

Abstract (English/Deutsch)	i
List of Figures	vii
List of Acronyms	ix
Introduction	1
1 Fundamentals	9
1.1 Theoretical Fundamentals	9
1.1.1 Electrospray ionization	9
1.1.2 Scanning tunneling microscopy	11
1.2 Experimental Set-up	14
1.2.1 Chamber set-up	14
1.2.2 Ion optics	16
1.2.3 Beam characterization	19
1.2.4 Software and data acquisition	20
1.3 The experimental workflow	21
1.4 Molecules and structure	24
1.4.1 Peptides/Proteins	24
1.4.2 Glycans	26
2 Saccharides: Self-assembly and characterization of isomers	31
2.1 Electrospray mass spectrometry of disaccharides	34
2.2 2D crystallization of disaccharides by self-assembly	36
2.2.1 Ag(111): Self-assembly into closed packed structures	36
2.2.2 Cu(100): Self-assembly into porous structures	37
2.3 Single molecules	45
2.4 Summary and Outlook	46

Contents

3	Peptides as sequence-controlled building blocks for self-assembly	49
3.1	Peptide assembly and nanostructures	49
3.2	Steric aspect to stabilize peptide tecton	52
3.3	Sequence-controlled manipulation	57
3.4	Characterization of the honeycomb network	58
3.5	Model of the honeycomb network	60
3.6	Defects	62
3.7	The role of the substrate	64
3.8	Conclusion and Outlook	66
4	Peptides: Metal-coordination to steer the conformation on the surface	69
4.1	The angiotensin-II-metal complex in the gas phase	70
4.1.1	Ion beams of AT-II-metal complexes	71
4.2	Peptide-metal coordination on the surface	72
4.2.1	1D nanostructures by coordination to Ni and Cu on Au(111) . .	73
4.2.2	Comparison of structural motifs between the pure peptide and the metal coordination complexes on Au(111)	81
4.2.3	Close packed nanostructures on Ag(111)	82
4.2.4	Isolated nanostructures of [AT-II + Cu] on Cu(111)	84
4.3	Conclusion and Outlook	89
5	Protein imaging by low energy electron holography	91
5.1	Low energy electron holography	94
5.2	Sample preparation and transfer	97
5.3	Native mass spectrometry with the ES-IBD set-up	98
5.4	Imaging folded protein structures by holography	102
5.5	Outlook and future plans	107
	Conclusion and Outlook	111
	Bibliography	115
	Acknowledgements	133
	Curriculum Vitae	135

List of Figures

1	Comparison of natural and artificial molecular machine	2
1.1	Schematic of the ESI	10
1.2	Schematic of the STM Principle	12
1.3	The combined set-up of STM and ES-IBD	15
1.4	Suitcase for UHV sample transport	16
1.5	First stability region on a-q space of Mathieu's equation	18
1.6	Amino acids and their properties	25
1.7	Ramachandran plot of amino acids	26
1.8	Structure of monosaccharides	27
1.9	Disaccharides and Polysaccharides	28
2.1	Structure of sucrose and trehalose	32
2.2	Mass spectra of disaccharides	35
2.3	STM image of trehalose on Ag(111)	36
2.4	Time series of assembly of trehalose on Cu(100)	38
2.5	Structural motifs of trehalose on Cu(100)	39
2.6	Assemblies of trehalose on Cu(100)	41
2.7	Assemblies of sucrose on Cu(100).	43
2.8	Single trehalose molecules on Cu(100)	46
2.9	First data for mannose pentamer on Cu(100)	48
3.1	The molecule angiotensin-I.	53
3.2	Overviews STM image of AT-I on Au(111).	54
3.3	Detail AT-I structures.	55
3.4	Sequence manipulation	58
3.5	Overview images of AT-II on Au(111)	59
3.6	The honeycomb network of AT-II on Au(111)	60
3.7	Atomistic model of AT-II on Au(111)	61
3.8	Defects in the honeycomb network.	63

List of Figures

3.9	AT-II assembly on Ag(111)	66
3.10	Possible sequences to test	68
4.1	Mass spectra of metal peptide coordination complexes before mass selection.	72
4.2	STM image of AT-II + Cu coordination complex on Au(111)	73
4.3	STM image of (AT-II + Cu) coordination complex on Au(111) after annealing to 360 K.	75
4.4	Model of (AT-II + Cu) coordination complex on Au(111)	76
4.5	STM images of (AT-II + Ni) coordination complex on Au(111)	78
4.6	Model of (AT-II + Ni) coordination complex on Au(111)	80
4.7	Comparison of the different models on Au(111)	81
4.8	Overviews STM image of (AT-II + Cu) coordination complex on Ag(111).	83
4.9	STM images of AT-II on Cu(111).	85
4.10	Assembly of AT-II trimer structures on Cu(111).	86
4.11	Overviews STM image of AT-I on Au(111).	87
4.12	Atomistic models of structures formed by (AT-II + Cu) coordination complex on Cu(111)	89
5.1	Holography set-up	95
5.2	Image reconstruction	96
5.3	Schematic work-flow of the experiment	97
5.4	Transfer standard	98
5.5	Mass spectra of CytC and BSA	100
5.6	Mass spectrum of Hb	101
5.7	Complete data set for CytC	103
5.8	Time evolution of CytC orientations	103
5.9	BSA micrographs	104
5.10	Comparison of low-energy electron micrographs of BSA with simulated electron density isosurface	105
5.11	Micrographs of Hb	106
5.12	STM image of Hb on Cu(100)	109
5.13	Comparison of mass spectra of BSA	110
5.14	Native mass spectrometry of AmtB by micelle protection	114

List of Acronyms

<i>AC</i>	alternating current
<i>AFM</i>	atomic force microscopy
<i>AT-I</i>	angiotensin-I
<i>AT-II</i>	angiotensin-II
<i>BSA</i>	bovine serum albumin
<i>cryo-EM</i>	cryo electron microscopy
<i>CytC</i>	cytochrome C
<i>DC</i>	direct current
<i>DFT</i>	density functional theory
<i>DNA</i>	deoxyribonucleic acid
<i>ESI</i>	electrospray ionization
<i>ES-MS</i>	electrospray-mass spectrometry
<i>ES-IBD</i>	electrospray-ion beam deposition
<i>EtOH</i>	ethanol
<i>Hb</i>	hemoglobin
<i>IMP</i>	integral membrane protein
<i>IMS</i>	ion mobility mass spectrometry
<i>LDOS</i>	local density of states
<i>LEEH</i>	low energy electron holography
<i>NMR</i>	nuclear magnetic resonance spectroscopy
<i>MD</i>	molecular dynamics
<i>MOCN</i>	metal organic coordination network

List of Acronyms

<i>MS</i>	mass spectrometry
<i>PCR</i>	polymerase chain reaction
<i>RF</i>	radio frequency
<i>RT</i>	room temperature
<i>SEM</i>	scanning electron microscopy
<i>SPM</i>	scanning probe microscopy
<i>STM</i>	scanning tunneling microscopy
<i>TEM</i>	transmission electron microscopy
<i>TOF</i>	time-of-flight
<i>TOF-MS</i>	time-of-flight mass spectrometry
<i>UHV</i>	ultra high vacuum
<i>XPS</i>	X-ray photoelectron spectroscopy

Introduction

It is very easy to answer many of these fundamental biological questions; you just look at the thing! –Feynman, Plenty of Room at the bottom

For their groundbreaking work in "the design and synthesis of [the first generation of] molecular machines", this years Nobel Prize in chemistry was awarded to the synthetic chemists Jean-Pierre Sauvage, Sir J. Fraser Stoddard and Bernard L. Feringa.¹ By exploring and generating molecular junctions that allow for mechanical movement and molecular rotors that exhibit restricted rotation in only one direction,² it has been eventually possible to create synthetic molecular machines that can perform directed movement, even in the macroscopic scale.³

The archetype of functional machines on a molecular basis is nature: The mammalian body is undoubtedly the best machine ever created. In every cell, there is an orchestra of molecular machines at work that all have very specific tasks such as transcribing DNA, converting energy to create the molecular fuel (ATP), regulating/transferring molecules, or sensing the environment.

A comparison between the artificial molecular machine, such as the nanocar,⁴ and the biological highly functional assembly, represented by the ATP Synthase,⁵ is depicted in Fig. 1, and highlights the challenges we face today: On the one hand, artificially designed molecular machines ensure the complete understanding of structure and synthesis, yet they lack the functionality and adaptability of their natural counterparts. On the other hand, the biological archetype of molecular machines is highly sophisticated and functional as well as optimized for its specific environment. However, they are very complex and often not yet completely understood, not to mention the working mechanism and their generation.^{6,7}

The prominence that natural structures still have over artificial machines synthesized by chemists is established over billions of years of evolution. Even the bare

Introduction

understanding of the final structure is sometimes very challenging and Feynmans exclamation "*you just look at the thing!*"⁸ is not easy to realize because the available tools, such as nuclear magnetic resonance spectroscopy (NMR), X-ray crystallography or cryo-electronmicroscopy (cryo-EM) have their drawbacks, including sample preparation, radiation damage or averaging. Nonetheless, it is motivation to develop new tools and engage in new approaches to deepen the knowledge and understanding of fundamental biologic questions.

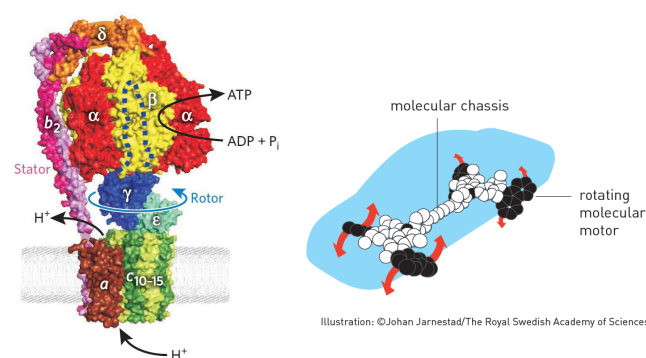


Figure 1: Comparison of natural and artificial machines. The natural machine (left) is represented by the protein complex ATP synthase, while the artificial molecular machines (right) is depicted by the nanocar. Images reproduced from [4] and [1] .

There are different approaches to bridge the gap between the artificial and the biological molecular machines. One way is to gain a better understanding on how the biological archetypes are built and work on an atomic scale. This is achieved by the structural determination and characterization on the molecular level (NMR, cryo-EM, X-ray crystallography, modeling).^{5,9,10} At the same time, being able to artificially synthesize these molecules provides a deeper understanding and enables the controlled manipulation to deviate from the natural archetype. Eventually, this also empowers the employment of the natural resources to create new materials with completely new properties.^{11,12}

Biopolymers: Modularity for highest functionality

Out of the four major classes of biomolecules (nucleic acid, lipids, carbohydrates and peptides/proteins) that constitute to the functional machinery, three, namely nucleic acids, carbohydrates and peptides/proteins, are built up as modular polymers. In them, a specific sequence is generated by a combination of relatively few building

blocks. This combination determines the shape or information content of the sequences and thus the function of the biopolymer.

DNA is a role model of how understanding binding motifs and accessing the artificial synthesis¹³ leads to new functional materials. First, DNA was identified as the carrier of the genetic code consisting of only 4 building blocks¹⁴ and later, in 1953, the structure was determined by X-ray crystallography as the known double helix.¹⁵ However, it took nearly 30 years to synthesize¹³ and amplify artificial sequences via polymerase chain reaction (PCR)¹⁶. The easy access and modularity of this biopolymer sparked material science and developed into the new field of DNA origami¹⁷ which enables the construction of nearly every geometrical shape based on the strong binding motif of complementary base pairs and its stability.¹⁸ Even functional and adaptable structures such as nano-tweezers, hinges and sensors are now developed based on the reversible binding motifs of this biopolymer.^{19,20} The verification of the created structures is done by imaging techniques such as atomic force microscopy (AFM), transmission electron microscopy (TEM) or scanning electron microscopy (SEM).²¹ Even before DNA, the automated synthesis of peptides was established in the 1960s by R. Merrifield and leaped forward the field.^{22,23} Built up by a combination of the 20 canonical amino acids and a sequence of up to 50 amino acids, the great versatility of peptides and their high affinities for specific targets make them the perfect species for therapeutic applications still being investigated nowadays.²⁴ Moreover, identifying specific structures and binding motifs, such as the α -helix or the β -sheet, has inspired material scientists to use these biopolymers for new materials.^{25,26} In contrast to nucleotides containing a combination of the 4 nucleic acids, biopolymers made of amino acids are more versatile and possess different functional groups due to the 20 different building blocks they can consist of. Thus, peptide-based materials can range from gels²⁷ to complex nanostructures with diverse properties.^{26,28} Proteins, consisting of more than 50 amino acids, are much larger than peptides. As in peptides, the amino acid sequence of proteins is linear. However, due to the immense length of the sequence, the automated synthesis does not work due to the high error rate during the synthesis. Thus, the production of proteins is achieved by manipulating the gene expression in an organism, such as *e. coli*, to express large amount of a recombinant gene to produce the protein which can be eventually purified.²⁹ Industrially, this process is for example used for the production of insulin for medication.³⁰

Proteins are highly functional and build up most of the molecular machines in the cell, either as single molecules, such as the electron transfer protein cytochrome C, or

as complex assembly of several proteins, such as myosin which is involved in muscle contraction. They are an essential part of virtually every process in the cell.³¹ The structure and the nanoscale dynamic of proteins and protein complexes is far more complex than any molecular machine that has been artificially constructed so far. Nonetheless, a protein can only function correctly if it is folded in the right shape, called the native conformation.³¹ While the sequence already contains the complete information of function and shape, factors such as chaperon molecules, post-translational functionalization and the environment (salt, solvent, temperature) also play a key role.³² However, the structure determination of proteins in their folded state is still challenging³³ and thus, there are plenty of natural proteins that are not yet characterized because they are not accessible by the standard methods of X-ray crystallography or cryo-EM for different reasons including sample preparation and radiation damage.^{34,35}

In contrast to the linear biopolymers DNA and peptides/proteins, carbohydrates can have a highly complex sequence of monosaccharides that exhibits branching. While carbohydrates are found in cell walls to provide structure (cellulose)³⁶ or in nutrition providing energy (glucose), the importance of their role in influencing protein folding³⁷ and intercellular recognition and adsorption is currently reevaluated, opening the field of glycomics.^{38,39} In general, the structure of carbohydrates is very diverse.⁴⁰ There are hundreds of different monosaccharides known to serve as building blocks in contrast to 4 (DNA) or 20 (peptide/protein) building blocks that make up a linear sequence. This gives rise to an enormous challenge for the identification and characterization of glycans,⁴¹ especially as several of the building blocks are isomers and only differ slightly in the regio- and stereochemistry. Moreover, the access to pure glycans has been very challenging, if not impossible, in contrast to the well-established synthesis and amplification methods of bacterial expression and PCR for proteins and oligonucleotides, respectively. This has drastically hampered the development of appropriate analytical tools.⁴² Only recently was the automated glycan assembly introduced, providing access to pure carbohydrates as well as glycoconjugates in large amounts (μg to mg).^{41,43} With these samples at hand, the field is certainly gaining momentum, with respect to analytics, structure determination and distinguishing structure-function relationships.⁴⁴

Characterization and imaging of biomolecules

The identification of the biopolymers' sequence is routinely done with mass spectrometry.⁴⁵ The development of electrospray ionization (ESI) and other ionization techniques, such as MALDI, leaped the analysis forward as large, fragile molecules became available in the gas phase.^{46,47} While most of the structure determination for biopolymers in their native state is performed by NMR, X-ray crystallography and cryo-electron microscopy (cryo-EM), new techniques in mass spectrometry, such as ion mobility⁴⁸ and new forms of MS/MS, such as collision induced dissociation (CID)⁴⁹ and surface induced dissociation (SID)⁵⁰ pave the way for mass spectrometry as structure characterization tool and enhance the importance of ESI-MS.

Moreover, cryo-EM improved the structure determination of native proteins,⁵¹ because it does not rely on crystalline samples as X-ray crystallography, nor does it need large amounts of sample, as is necessary for NMR.³⁵ However, it is only suitable for large structures and averages over large ensembles due to nearly instant radiation damage.^{34,52}

In contrast, scanning probe microscopy, such as atomic force microscopy (AFM) and scanning tunneling microscopy (STM), are powerful tools to image molecules on surfaces nondestructively. In the last decade, the adsorption behavior and intermolecular interactions of small molecules to self-assemble has been thoroughly investigated by STM, demonstrating that the principles of supramolecular chemistry are also valid on the surface.^{53,54} Recently, by functionalizing the tip, non-contact AFM in UHV was employed to unravel the molecular structure of large aromatic hydrocarbons that was not accessible otherwise.⁵⁵ However, the sample preparation is key for SPM to achieve high-resolution, thus the samples are prepared in UHV. The limited accessibility in UHV due to their non-volatile character has long hampered molecular imaging of biopolymers by SPM.

Combining ion soft-landing, as demonstrated by Cooks in 1977⁵⁶, with soft ionization techniques grants access to complex macromolecules on surfaces in vacuum^{57,58}, such as nanographene⁵⁹, dyes⁶⁰ or biomolecules^{61,62} and even viruses⁶³. Electrospray ionization (ESI) is most suitable to transfer biomolecules from liquid into gas phase because this generation of gas phase ions is not limited by the thermal stability. Moreover, by controlling the solution parameters, manipulation in the gas phase conformation⁶⁴ as well as the formation of metal complexes is possible.⁶⁵ The gas phase ion can subsequently be transferred into UHV and deposited on a sample.^{57,58} The deceleration of the molecular ions for soft landing avoids fragmentation and ensures a deposition of intact species.

It has also been shown that deposited proteins can maintain their biological activity.⁶¹ By IR and Raman, it could be demonstrated that by slight activation in landing energy, covalent attachment to functional self-assembled monolayer (SAM) surfaces is possible.⁶⁶ As the deposited species are ions, charge retention has been confirmed for deposition on insulating surfaces such as fluorinated SAM (FSAM) surfaces.⁶⁷ However, the properties of the deposited biomolecules, such as conformation, charge state and interaction with the substrate strongly depend on the substrate.⁶⁸

In addition to spectroscopy methods, such as Raman, IR, and MS, imaging techniques, like TEM, AFM and STM proved to be powerful to investigate the properties of deposited biopolymers. The influence of the landing energy has for example been investigated by TEM images of protein complexes GroEL and ferritin.⁶⁹ Here, the preservation of the protein structure in general could be shown. Increasing the landing energy, apoferritin was created on the surface, whereas for soft-landing, the holoferritin with an iron core was preserved. Similarly, STM investigation could show the influence of the landing energy and demonstrate that the gas phase conformation of the unfolded protein cytochrome C being controlled by charge state also influences the conformation on the surface.⁷⁰ While for TEM investigations, staining has been used, STM can be employed directly after landing, achieving sub-molecular resolution, even for complex biopolymers.^{70,71}

Thus, combining the capabilities of ESI-MS and high-resolution imaging techniques is a potential approach to investigate structural motifs and new materials. The transfer into vacuum of large complex biomolecules and the subsequent chemically and physically controlled deposition is fundamental for meaningful molecular imaging of complex biomolecules.

Structure of this thesis

In this thesis, I show how the molecular imaging of glycans, peptides and native proteins by STM and low energy electron holography (LEEH) contributes to deciphering the structures and the respective binding motifs. These high-resolution imaging methods only perform well on surfaces in UHV and for chemically identified species. Thus, ES-IBD is essential for the selected deposition for non-volatile species in UHV.

To this end, the working principles of ES-IBD and STM are explained in **Chapter 1**, along with the set-up and the experimental procedure. The understanding of characteristic motifs of peptides/proteins as well as carbohydrates is important for the following experiments and are therefore shortly introduced in this chapter.

Chapter 2 deals with the characterization of carbohydrates and highlights the impact

that STM might have in the future for the clarification of a unknown sequence by imaging branching.

In **Chapter 3**, peptides are explored as functional building blocks for the self-assembly of new nanostructures. In combination with molecular dynamics and density functional theory, atomistic models are developed, giving insight into the binding motifs at work and help to gain an atomistic understanding of the new structures. Based on the fundamental understanding of the binding motifs in the previous chapter, I demonstrate the conformation change by coordination of the peptide to a metal center in **Chapter 4**.

While the high resolution imaging of small biopolymers have been presented in the previous chapters, I demonstrate in **Chapter 5** that ES-IBD is a powerful tool that also enables other UHV imaging techniques. Expanding the deposition to natively folded proteins, STM is not a suitable imaging technique as it is not capable of imaging large three-dimensional structures of folded proteins with high resolution. With the new approach of combining ES-IBD with LEEH, a proof of principle is performed to show that high-resolution imaging of proteins is possible without averaging.

The last chapter gives a short summary and an outlook how the vacuum fabrication and analysis of adsorbed biopolymers by ES-IBD can be employed for future research. Promising instrumental approaches as well as unsolved biological questions are sketched demonstrating the rich variety of samples that ES-IBD enables to investigate.

1 Fundamentals

1.1 Theoretical Fundamentals

1.1.1 Electrospray ionization

Electrospray ionization (ESI) is a powerful technique that has opened new routes for the analysis of large, non-volatile molecules by transferring charged molecules from liquid into gas phase in a very gentle way so that no fragmentation occurs and even non-covalent bonds are preserved.^{72,73} For his development of ESI, John B. Fenn was awarded the Nobel Prize in chemistry 2002, together with Koichi Tanaka (soft laser desorption) and Kurt Wüthrich (NMR spectroscopy). In order to “make elephant molecules fly” by electrospray,⁷⁴ the molecules have to be dissolvable. Small amounts of acid or base is added for positive or negative mode respectively to enhance the ionization. The analyte solution is pumped through an emitter⁷⁵ which is held at a high voltage with respect to the orifice. Due to the strong resulting field, charge separation occurs in the solution at the tip of the emitter and deforms the meniscus into a cone.⁷⁶ This phenomenon – the so-called Taylor cone – was first observed by Zeleny⁷⁷ and theoretically described by Taylor⁷⁸. The sharp curvature of the cone leads to a locally enhanced electric field due to which a fine jet of liquid is emitted at the Taylor cone. In application, this regime is the most stable mode known as the cone-jet mode.^{79–81} Since the jet emerges from the point of highest charge density, it carries an excess of ions. In the electric field, it gets further accelerated and breaks into an aerosol plume of highly charged droplets, the electrospray. At this point, the droplets are of micrometer size and carry some 10^4 charges.⁸² Subsequent solvent evaporation decreases the size of the droplet and increases the

charge density. The charge density increases until the Coulomb repulsion is equal to the surface tension γ . This limit of charge carrier in a droplet is called Rayleigh limit,⁸³ defining the maximal number of charges as:

$$z_R = \frac{8\pi}{e} \sqrt{\epsilon_0 \cdot \gamma \cdot R^3} \quad (1.1)$$

with R being the radius of the droplet, and ϵ_0 the vacuum permittivity. As soon as the Coulomb repulsion exceeds the surface tension by further shrinkage, the droplets disintegrate into smaller droplets. Similar to the formation of the Taylor cone, the droplets are deformed in the electric field and thus a droplet jet fission happens. This process repeats several times until nanodroplets or ions emerge. The whole process from the microdroplet to the isolated ion takes place in atmosphere and takes less than one millisecond.⁸⁴

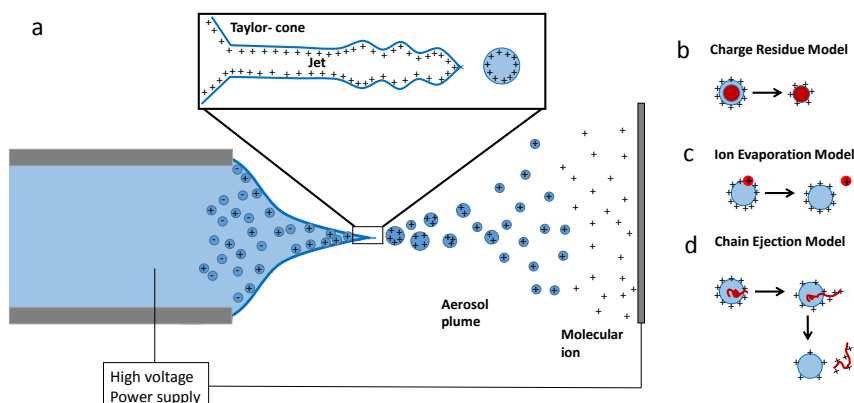


Figure 1.1: Schematic of the ESI. **a:** The deformation of the meniscus and the creation of a Taylor cone (inset) is depicted that leads to charged droplets in the aerosol plume. The different models for the creation of charged analyte ions are presented: **b:** The charge residue model is accepted for the creation of large globular species. **c:** Ion evaporation model for the ejection of small ions. **d:** chain ejection model, which describes the ejection of expanded and long analyte species.

However, the mechanisms of ion formation are still under debate. Three different mechanisms are stated for the final creation of the isolated molecular ion depending on the properties of the analyte molecule.^{85,86} In the oldest model, the charged-residue model (CRM)⁴⁶, the ion is formed by complete evaporation of the solvent shell. This model is widely accepted for large globular species, such as native proteins.

The ion evaporation model (IEM)⁸⁷ is assumed for low weight species (inorganic ion) and describes the ejection of the ion from the droplet surface. Recently, the chain ejection model (CEM)⁸⁸ has been introduced for long polymer chains assuming a consecutive ejection of the beads of a linear chain from the droplet.

Once a beam of non-solvated ions is created, it can be guided into vacuum and manipulated by ion optics to be analyzed by mass spectrometry⁸⁰ or for deposition on sample targets^{59,89,90}.

1.1.2 Scanning tunneling microscopy

Since its invention in 1981 by Rohrer and Binnig,⁹¹ STM has revolutionized the field of surface analysis and granted them the Nobel prize in 1986.⁹² This technique enables the imaging of conducting surface topographies in real space at the atomic scale by a sharp tip scanned over the surface while the tunneling current is monitored, as depicted in Fig. 1.2a. Moreover, STM measurements can be performed at a large range of temperatures (mK to ~1300 K) and at different interfaces, such as the liquid-solid, air-solid or UHV-solid interface. Nonetheless, due to the highly controlled environment and the clean surfaces, the best resolution is achieved in UHV.

For measurements, STM can be conducted in two different modes: in the constant height mode, the tip is approached to the surface and the z-position of the tip is kept constant during scanning while the tunneling current varies and thus carries the information. This mode does not require any feed-back, and thus, the scan-speed can be higher. However it is only suitable for very flat surfaces as the tip can otherwise crash into the sample.

In the constant current mode, which is used for all measurements shown in this thesis, a feed-back loop controls the z-position of the tip in order to keep the tunneling current constant. Thus, by plotting the z-piezo movement of the tip, the image is obtained. This mode allows to measure rougher surfaces or large areas that contain several terraces without crashing into the surface.

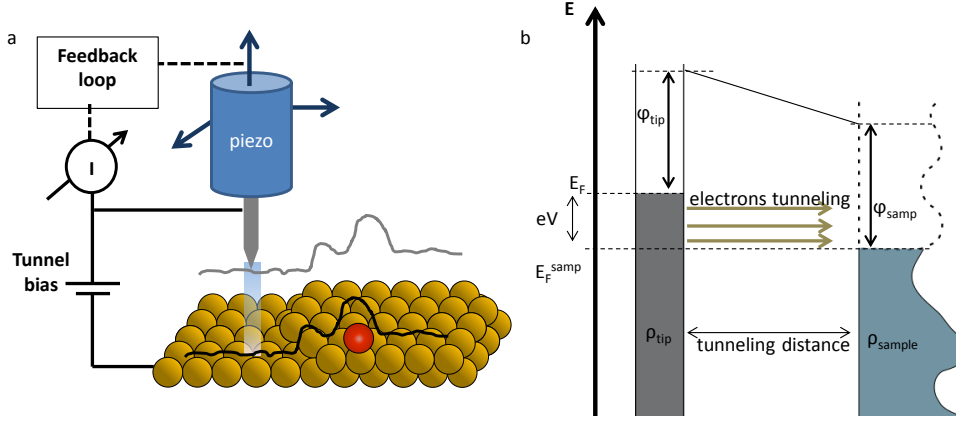


Figure 1.2: Schematic of the STM principle. **a:** A sharp metallic tip is scanned over a surface. Depending on the mode, the feedback is turned on (constant current) or off (constant height). **b:** Schematic of the energy alignment. Due to the applied bias the Fermi energies E_F are shifted against each other by eV and thus allow for a measurable current from the tip to the sample.

In order to know how to interpret the obtained STM data, a concise derivation of the tunneling current is given in the following. A more extensive approach can be found in literature.^{93,94} The working principle is based on the tunneling effect between two metal electrodes.⁹⁵ The electrons tunnel through a small insulating gap - the tunnel junction - due to their quantum mechanical properties: electrons can be described by wave functions, which decay exponentially into the barrier of the tunneling junction. For small gaps, the wave function still exists behind the gap and therefore the probability of finding an electron in or beyond the barrier in the other electrode is finite. By applying a bias voltage to one electrode, the Fermi levels of this electrode is shifted against the other which drives the electrons in one direction, as shown in Fig.1.2b, resulting in a constant tunneling current that can be measured.

Using first order perturbation theory, Bardeen describes the tunneling current between two electrodes as:^{96,97}

$$I = \frac{2\pi e}{\hbar} \sum_{\mu\nu} f(E_\mu) [1 - f(E_\nu + eV)] \times |M_{\mu\nu}|^2 \delta(E_\mu - E_\nu) \quad (1.2)$$

with the Fermi function $f(E)$, the applied voltage V and the energies E_μ and E_ν given with respect to the Fermi level E_F . The summation occurs over all states μ and ν of the unperturbed electrodes, respectively with $M_{\mu\nu}$ being the tunneling matrix element between the states ψ_μ and ψ_ν at the electrodes. In the limits of small voltages and low temperature (Tk_B being much smaller than the required energy resolution), the Fermi

function can be seen as step function, simplifying the expression to

$$I = \frac{2\pi e}{\hbar} \sum_{\mu\nu} |M_{\mu\nu}|^2 \delta(E_\nu - E_F) \delta(E_\mu - E_F) \quad (1.3)$$

In order to be able to compute the current, the tunneling matrix $M_{\mu\nu}$ has to be calculated. Bardeen could show that this matrix is given by

$$M_{\mu\nu} = -\frac{\hbar^2}{2m_e} \int dS (\psi_\mu^* \nabla \psi_\nu - \psi_\nu \nabla \psi_\mu^*) \quad (1.4)$$

where any surface must be evaluated that lies within the tunneling gap. This evaluation proved to be very difficult for extended electrodes.

However, in a UHV-STM, the two electrodes are represented by a sharp metallic tip on one side and the (extended) sample on the other, while the tunnel junction is a precisely controlled vacuum gap.⁹⁸ The tip is brought close to the surface of the sample, resulting in a tunneling gap of distance z (in the order of Å). The tunneling current between the tip and the sample is estimated by the Tersoff-Hamann model which applies Bardeen's quantitative theory of tunneling to STM by assuming a spherical tip possessing only s-states.^{97,99} For small voltages $V \ll \Phi_s, \Phi_t$ the tunneling matrix could then be approximated to

$$M_{s,t}(z) \propto e^{-2z\sqrt{\frac{m_e}{\hbar^2}(\Phi_s + \Phi_t)}} = e^{-2z\sqrt{\frac{2m_e}{\hbar^2}\Phi}} \quad (1.5)$$

Merging the workfunctions Φ_t and Φ_s of the tip and the sample to a general workfunction $\Phi = \frac{\Phi_t + \Phi_s}{2}$ accounts for the often unknown workfunction of the tip (and the sample) in the experimental set-up.

For small bias voltages V and tip and sample being in thermal equilibrium at low temperatures T (around 0 K), the tunneling current $I(V)$ can now be related to the local density of states (LDOS) of the tip (ρ_t) and the samples $\rho_s(x, y)$ by combining eqn. 1.5 with eqn. 1.3:

$$I(V, x, y, z) \propto e^{-2z\sqrt{\frac{2m_e}{\hbar^2}\Phi}} \int_0^{eV} dE \rho_s(E, x, y) \rho_t(E - eV) \quad (1.6)$$

The prefactor (tunneling matrix) describes the exponential dependence of the tunneling current from the distance implying that the tip apex dominates the tunneling process. For example, given an estimated workfunction $\Phi = 4$ eV, a displacement of 1 Å yields a ten-fold variation in current, explaining the atomic-scale capabilities of

STM. The integral factor represents the contribution from the LDOS from the sample and the tip to the tunneling current. Since both factors, the integral and the tunneling matrix, influence the tunneling current, the measurement with an STM represents always the convoluted information of the real surface contour and the electronic configuration of the tip and sample system. This is what makes STM data not straightforward to interpret.

1.2 Experimental Set-up

1.2.1 Chamber set-up

The ES-IBD set-up used in this thesis is a homebuilt set-up.⁸⁹ The complete set-up with the ES-IBD instrumentation as well as the preparation (Prep) and STM chamber is shown in Fig. 1.3 in top view. The described ionization process for ESI completely happens in atmospheric conditions. To analyze and further process the generated ions, they have to be transferred into vacuum. Originally, ESI was developed as an ionization technique for mass spectrometry to access large, non-volatile molecules. In contrast to mass spectrometry, which can be operated at low ion currents in the high vacuum range (10^{-6} mbar), the application of ESI for ion beam deposition requires high transmission rates into ultrahigh vacuum (UHV) while still maintaining the power to characterize the beam. In order to overcome the pressure difference between atmosphere for the ionization and UHV for deposition, a differentially pumped system is used consisting of six separately pumped chambers connected by 2 mm apertures. It consists of four sections: the atmospheric interface for the ion source (red)⁹⁰, the high pressure stage in which the beam is formed (yellow), the high vacuum section for beam characterization and high vacuum sample deposition (green) and finally the ultra-high vacuum section for deposition (blue). In the UHV deposition chamber (DEPO, 10^{-10} mbar), a sample stage for 7 different sampleholders (2 Omicron, 5 for external UHV samples) is implemented on the manipulator that can be cooled by liquid nitrogen to 90 K. The manipulator is used to position the employed sampleholder by x,y,z and ϕ alignment in front of the aperture for deposition and transfer the Omicron sample holder in the preparation chamber for further transfer into the STM. Moreover, a load lock in line of the beam is available to access the deposition chamber with external UHV samples via a suitcase.

In the preparation chamber (Prep chamber, 10^{-10} mbar), substrates are prepared in

vacuo and can be transferred into the deposition chamber or the STM. The adjacent STM chamber (STM, 10^{-10} mbar) is decoupled by a soft bellow from the preparation chamber to reduce mechanical noise because the pumps on the preparation chamber and the ES-IBD are not turned down for STM measurements. This chamber is actively damped by a piezoelectric active vibration isolation system (TMC stacis 2100). The variable temperature STM (VT-Omicron) in this chamber has been substituted by an Omicron Fermi SPM¹⁰⁰ during this thesis which allows for measurements between 10 K and 400 K. Samples are introduced by the load lock (LL) that is connected to the preparation chamber.

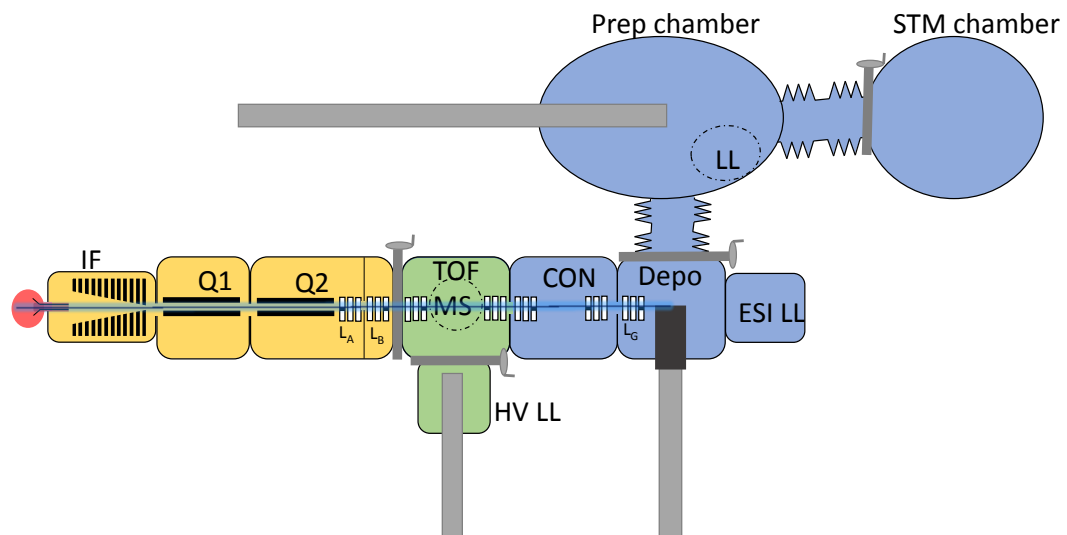


Figure 1.3: The combined set-up of STM and ES-IBD. In the ES-IBD apparatus, ions are transferred via the atmospheric interface (red) into the vacuum (yellow, high pressure region) where they are collimated by ion optics (e.g. ion funnel IF, quadrupole Q1, lenses and steering plates L_A, \dots) and mass selected (quadrupole Q2). In the high vacuum region (green), the beam is characterized by mass spectrometry and energy measurement before the beam is guided into the UHV (blue) for deposition.

UHV-suitcase

The versatile samples that can be produced by ES-IBD enable the investigation by a lot of different techniques, such as X-ray photo electron spectroscopy, transmission electron microscopy, low energy electron holography or low temperature STM and scanning tunneling spectroscopy. In earlier experiments, a prototype suitcase was used to transfer the samples from the deposition chamber into other UHV set-ups

without breaking the vacuum. This approach has now been optimized and realized with 7 suitcases. Each suitcase, depicted in Fig. 1.4 a, has been developed in collaboration with FerroVac and is constructed by FerroVac. It consists of a small chamber and a wobble stick that is explicitly adapted to suit the sample holder uptake from a specific experiment. The whole suitcase is pumped by a NexTorr pump combining a passive non-evaporable getter pump and an ion-pump.

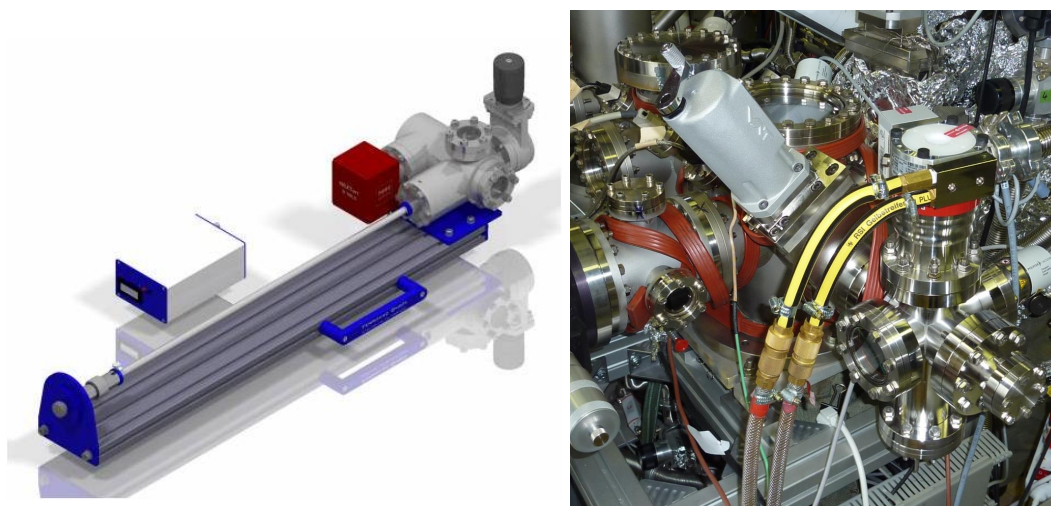


Figure 1.4: Suitcase for UHV sample transport. **a:** Schematic of the complete suitcase from FerroVac. **b:** Photograph of the ESI load lock with the cryo-pump at the bottom of the chamber.

A load lock is set up in-line of the beam at the deposition chamber, as shown in Fig. 1.4b, where the suitcase can be attached. For fast entry, a cryo pump is installed additionally to the turbo pump. It consists of a copper rod covered with charcoal as adsorption material and is cooled by liquid nitrogen. This reduces the pumping time and enables pressures of 10^{-9} mbar within 2 hours without heating out the UHV chamber. Thus, a fast transfer from the suitcase in the sample shuttle is possible and saves precious time on the experiment.

1.2.2 Ion optics

The ion optics in every chamber of the ES-IBD, which are based on standard ion optics as used in mass spectrometry experiments, are optimized to achieve high transmission. The ion beam enters the vacuum through a transfer capillary and expands into

vacuum. The capillary can be heated to ensure complete desolvation of the ions. The expanding ion cloud is collimated by an ion funnel (IF) in the first pumping stage and a quadrupole Q1 which also serves as ion guide. A second quadrupole Q2 in the third pumping stage is finally used for m/z filtering of the beam. Further down the beam, electrostatic lenses (LA-LG) and deflection plates are used to focus and steer the beam in order to get high transmission through the series of apertures. For higher control, the ion current can be monitored at every aperture.

Ion Funnel

Compared to a skimmer, which is often used for mass spectrometry instruments, the ion funnel provides effective pumping and collimation of the beam at the same time which increases the transmission enormously. The ion funnel is built of 82 disc shaped ring electrodes on which an RF-field as well as a DC gradient potential can be applied. While the gradient potential pulls the ions through the funnel, the RF-field creates a potential that collimates the ions. At the pressure of 1 mbar in the ion funnel, frequent collisions with the background gas thermalize the ions and yields an effective collimation.

Quadrupoles: Ion guide and m/z filter

Both quadrupoles in this set-up can in principle act as bandpass filter.^{101,102} Applying Newton's law of motion to the ions in the field of a quadrupole $E_u = \frac{\partial \Phi(x,y,z)}{\partial u} = -\frac{2u}{r_0^2}(U - V_{RF}\cos(\Omega t))$ with U as a DC voltage between the poles and ground and V_{RF} as the AC voltage between two opposite poles, the Mathieu's differential equation can be drawn as a general solution for the path of the ions through a quadrupole.

$$\frac{d^2 u}{d\xi^2} + (a_u - 2q_u \cos(2\xi))u = 0, \quad (1.7)$$

where, u denotes either x or y direction of the quadrupole and $\xi = 2t\Omega$. r_0 is the field radius that denotes the distance from the center to an electrode of the quadrupole. The parameters $a = \frac{8zeU}{mr_0^2\Omega^2}$ and $q = \frac{4zeV_{RF}}{mr_0^2\Omega^2}$ simplify the six-dimensional problem (involving z , Ω , r_0 , m , U and V_{RF}) to a two-dimensional problem of a and q .

In general, two types of solution to the Mathieu equation can be classified: the bounded solution relates to a case where the path of the ion is stable, meaning that the ion can pass through the quadrupole with finite displacement. In contrast, the

Chapter 1. Fundamentals

ion's trajectory is unstable for unbounded solutions. These particles get filtered out by collisions with the electrodes. There are several regions, in which the x and y are both stable. These regions are called stability regions. Fig. 1.5 shows the first stability region in a-q space for the Mathieu's equation.

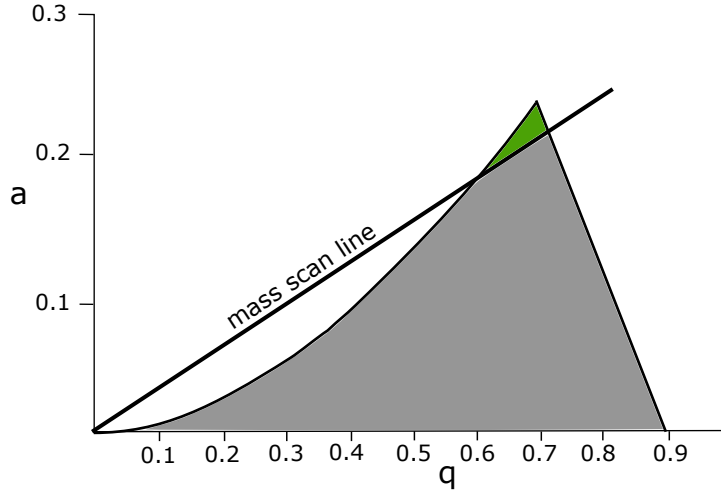


Figure 1.5: First stability region in a-q space of Mathieu's equation. The mass scan line is given by the slope $\frac{2U}{V_{RF}}$. Only the ions with m/z in the area of the stable region and above the mass scan line can pass the quadrupole (green area).

Here, the shaded region represents a collection of stable trajectories in both x and y direction in the a-q space. Since both parameters a and q are inversely proportional to m/z, ions with small m/z can be found in the upper right corner whereas large m/z (heavy, low charge state) ions are represented in the lower left corner.

In practice, both parameters, a and q, are kept related to each other by a ratio. This is ensured by keeping a constant ratio between the DC voltage U and the AC potential. This restricts the operation of the quadrupole to a straight line with a zero intercept and the slope of $\frac{2U}{V_{RF}}$, the so called "mass scan line" in Fig. 1.5. Only m/z above the mass scan line and within the stable region can pass the quadrupole (green area). By changing the ratio between U and V_{RF} , the slope of the mass scan line is changed, yielding a different width of the bandpass filter.

In the used set-up, both quadrupoles have the same geometry of 10 cm length and 5 mm diameter of the rods, giving a trapping volume of 6 mm minimal diameter each. The first quadrupole Q1 is run in RF-mode only at amplitudes of 50-200 V. This results in a large m/z window available. Since it works at a pressure of 10^{-2} mbar, it can serve as an ion guide due to the collimating effect of the collision with the background gas. The collimated ion beam leaves Q1 through a 2 mm aperture and enters the

quadrupole Q2 in the third pumping stage at high vacuum (10^{-4} mbar).

The second quadrupole Q2 is employed as bandpass mass filtering quadrupole with an additional differential DC voltage superimposed to open a narrow m/z window of 10-100 Th for transmission. With a kinetic energy that is defined by the electrostatic potential of the Q1-Q2 aperture vs ground, the mass selected ion beam is leaving the quadrupole Q2 entering the first of the electrostatic lenses L_A . At roughly this point, the pressure is low enough to assume no background collisions happening anymore.

1.2.3 Beam characterization

Flux of the ion beam

The flux of the ion beam can be measured as an electric current in which the analyte ions are the charge carrier. Since the kinetic energy of these ions is in the hyperthermal energy range (1-100 eV), no secondary electrons are created when the ions impinge on an electrode. Thus, if the charge state of the ions is known, the material flux can easily be obtained from the current measured with an electrometer (Keithley 616).

The current is monitored in multiple places: on the apertures L_A to L_G , as well as on the TOF aperture in order to monitor the stability of the beam and to help focusing and aligning the beam through these apertures. Moreover, the current is also monitored on the sample during deposition in order to control the coverage.

Chemical composition

The chemical characterization is achieved by a linear time-of-flight mass spectrometer (TOF-MS) of Wiley-McLaren geometry¹⁰³ in the fourth differentially pumped stage. For mass spectrometry analysis, the TOF-MS is positioned in the beam from the top. In this position, the ions enter a field-free extraction volume through a slit aperture and get extracted by push/pull pulses of ~300 V and 70 μs length into an acceleration volume with a repetition rate of 130 μs^{-1} . Within 100 mm, they are accelerated to 4950 eV/charge. After a drift tube of 400 mm the ions impinge on the fast dynode detector (ETP 14882). Both polarities, negative and positive ions, can be detected, depending on the mode of the ES-IBD. The signal is recorded by pulse counting and histogramming. The measured TOF spectra are converted into mass spectra by the calibration relation $m/z = A(t - t_0)^2$, in which the parameters A ($\sim 3.46 \cdot 10^{-6}$ Th ns⁻²) and t_0 (~ 372.5 ns) are obtained by fitting to a known spectrum.

A resolution of $M/\Delta M = 1000$ can be achieved with this set-up, depending on the spatial collimation and energy dispersion of the beam. A dynamic range of three orders of magnitude can be covered. By changing the amplification voltage of the dynode, the sensitivity can be adjusted.

This mass spectrometer is optimized for its purpose in the ES-IBD set-up and is mobile to be moved out of the beam during deposition while it spans a large dynamic range and can be adjusted in sensitivity. Thus, it is very suitable to monitor the composition and purity of an already known analyte species.

Kinetic energy

For measurement of the kinetic beam energy, the HV sample holder including the energy detector is moved into position in same chamber that is used as for mass spectrometry. The energy detector consists of an electrode and two metallic grids, one at the cover plate to avoid the deformation of the beam and one free-standing in front of the electrode where an independent voltage can be applied (retarding grid). By ramping a voltage at the retarding grid and measuring the transmitted current at the electrode behind, the beam energy distribution can be detected. Only by knowing the beam energy, the landing energy for a deposition can be adjusted by applying a voltage to the sample. The deposition energy $E_{\text{dep}} = zeU$ is defined by applying a bias voltage to the sample with respect to the beam energy: $E_{\text{dep}} = E_{\text{kin}} - E_{\text{applied}}$.

Accelerating the ions on the surface yields reactive landing. Due to the high landing energy, the landing ions can break, resulting in reactive fragments. This can be used for surface modification.¹⁰⁴

In order to image intact molecule as well as molecule complexes, soft landing is performed.^{58,90} For this, a retarding voltage is applied to the sample in order to decelerate the ions. In this thesis, only soft-landing experiments are described.

1.2.4 Software and data acquisition

Owing to the large dimension of the system and the almost 80 voltage channels that need to be served, the voltage output is controlled by a LabView program. This enables the operator to manipulate the different channels in a controlled way and simultaneously get feedback by the current monitoring. A computer hosts 5 digital-to-analog cards with 16 channels ± 10 V each. They control the voltage output channels with 16 bit precision ranging from 5 kV for the TOF and the emitter to 20 V for the

steering plates. The current readings of the Keithley 616 are digitalized by an ICP DAS card.

1.3 The experimental workflow

Creation of the ion beam

In order to generate an ion beam, the analyte molecule has to be dissolved. The most common solvent for biomolecules is water, as this is the natural environment. To facilitate the desolvation process and stabilize the spray, mixtures of water with ethanol, methanol or acetonitrile are commonly used when handling non-native ion beams. However, other organic solvents can also be used, e.g. if the analyte molecule only dissolves in nonpolar medium. Furthermore, small concentrations of a volatile acid (formic acid or acetic acid) or base (ammonium hydroxyl solution, ammonium acetate) are added to enhance the ionization. For ES-IBD, analyte concentrations of 10^{-3} to 10^{-6} M are used.

Table 1.1: Table of typical concentration and solutions used and the responding ion currents for all beams used in this thesis.

Molecule	Solution	Concentration	Ion Current on TOF
Sucrose	1:1 H ₂ O/EtOH + NH ₄ OH	$< 10^{-4}$	-100 - -400 pA (neg)
Trehalose	1:1 H ₂ O/EtOH + NH ₄ OH	$< 10^{-4}$	-100 - -400 pA (neg)
Man5mer	1:1 H ₂ O/EtOH + NH ₄ OH	10^{-4}	-30 - -50 pA (neg)
Angiotensin I	1:2 H ₂ O/EtOH +FA	10^{-4}	500 - 700 pA
Angiotensin II	1:2 H ₂ O/EtOH +FA	10^{-4}	600 - 800 pA
Angiotensin II + Cu	1:2 H ₂ O/EtOH + CuCl ₂	10^{-5}	200 - 300 pA
Angiotensin II + Ni	1:2 H ₂ O/EtOH + NiCl + NH ₄ OH	10^{-4}	150 - 200 pA
Cyt C	NH ₄ Ac 50mM	10^{-5}	~80 pA
BSA	NH ₄ Ac 50mM	10^{-5}	~50 pA
Hemoglobin	NH ₄ Ac 50mM	10^{-5}	~50 pA

The solution is pumped through the emitter – a fire-pulled silica column- at flow rates of 15-30 $\mu\text{L/hr}$. When the first macroscopic drop forms at the emitter, the emitter is set in the capillary inlet and a voltage between the emitter and the inlet is ramped up to 3-5 kV until a stable beam is formed. The ion beam is guided to the first current measurement at the aperture after the quadrupole Q2. With this current monitoring, the parameters of the flux, the quadrupoles, the ion funnel and the first lens can be optimized to get a well collimated beam with high transmission and stability. The stability can be judged by the appearance of the spraying emitter (no change should be visible) and the continuity of the ion current.

The ion yield strongly depends on the analyte species (solubility, ionization yield) and the solution parameters (solvent, concentration, pH, additives). Moreover, high losses in beam intensity occur due to mass selection and imperfect transmission from high vacuum to ultra-high vacuum. Thus, if the yield is too low, it might be helpful to adjust the solution parameters to improve the ionization and purity of the beam as well as the interface parameters (ion funnel, capillary temperature, curtain gas). For a guideline, typical solution parameters are noted in the Tab. 1.1 for the molecules used in this work.

Beam characterization

Once the beam is stable, it is guided into the TOF-MS to monitor the composition of the beam and control the proper mass selection by the quadrupole Q2. In order to ensure high sensitivity and a correct mass-filtering, the amplification of the TOF signal is adjusted by the detector voltage to yield a good signal. The comparison of the ion current before and after m/z selection gives a measurement on how pure the beam is. Even for a beam that looks reasonably pure in the mass spectrum, a rough m/z selection is advisable as this excludes large clusters that are not easily observable in the mass spectrum, but carry a lot of material.¹⁰⁵ Subsequently, the kinetic energy of the beam is measured.

UHV sample preparation and characterization

In this work, single crystals of noble metals (Au, Ag, Cu) were employed. The hat-shaped crystals (outer diameter 6.3 mm, 1 mm height, inner diameter 5 mm, 1.5 mm height) were commercially obtained (MaTeck GmbH and SPL). Prior to deposition, all samples are cleaned in the adjacent preparation chamber by several cycles of

Ar⁺-sputtering (10-30 min, 1 kV) and e-beam annealing to temperatures ensuring the surface reconstruction of the respective noble metals. This process results in clean, atomically flat surfaces with extended terraces. After the last annealing step, the samples are checked in the STM at RT and subsequently transferred in the deposition chamber for ES-IBD.

The detailed conditions for all samples employed are noted in Tab. 1.2. Two type of sample holders were used: a double-decker omicron plate was used for the VT-Omicron. It offered integrated e-beam heating and temperature reading directly at the single crystal. For the Omicron Fermi SPM, a single molybdenum omicron plate is used on which the sample is fixed by a molybdenum sheet hold tight with 4 screws. For this sample holder, the filament in the manipulator is used for e-beam heating and the temperature is read at the edge of the omicron plate.

Table 1.2: Table of single metal crystals, their annealing condition and the deposited molecules.

Crystal	Annealing temperature	Deposited Molecules
Au(111)	930 K	AT-I, AT-II, [AT-II + Cu], [AT-II + Ni]
Ag(111)	930 K	Trehalose, AT-II, [AT-II + Cu]
Cu(111)	930 K	Sucrose, Trehalose, AT-II, [AT-II + Cu]
Cu(100)	800 K	Sucrose, Trehalose, Man5mer

Deposition

After a thorough beam characterization, the beam is guided through all apertures by adjusting the lenses and steering plates. With the last steering plate (L_G), the beam can be scanned over the cover plate of the sample holder. Once the transmission is optimized, the sample is transferred from the preparation chamber into the deposition chamber while the valve in the TOF chamber is closed to ensure no uncontrolled deposition. Finally, by opening the TOF chamber valve, the deposition starts because the beam reaches the sample. During deposition, the current on the sample as well as on some apertures in line of the beam is monitored in order to survey the stability of the beam. Moreover, the coverage of the sample can be estimated because the charge state and molecular mass of deposited species is known. For this, the integration of the current signal is implemented in the software. The final coverage is given in units of pAh (picoAmpère hour) as this accounts for the accumulated charge Q . With an aperture diameter of 5 mm, the deposition area can be estimated to 20 mm². The coverage on the sample Θ is given by $\Theta = \frac{Q}{zA}$. Exemplarily, for 1 pAh and charge state

$z = 1$, the coverage accounts for $0.001 \text{ nm}^{-2} = \frac{1}{1000 \text{ nm}^2}$.

In general, the deposition rate of ES-IBD is rather small and only applicable for coverages up to a few monolayers. Nonetheless, taking into account the initial amount of material used, the efficiency of this method is high: the major fraction of the analyte molecule can be ionized with the nano-ESI emitter at low flow rates. In combination with the high control (chemical, energetic and spatial), this deposition method is highly suitable for UHV sample preparation.

STM

For characterization, the sample has to be transferred into the STM. Here, the sample can be scanned either at room temperature or it can be cooled down with liquid nitrogen or helium by a flow cryostat. In general, the sample is imaged in constant current mode. The STM is controlled by a Nanonis SPM control system (Nanonis 4.1). Subsequent image processing is performed with WSxM 5.3¹⁰⁶ including background subtraction, flattening and contrast adjustment.

1.4 Molecules and structure

1.4.1 Peptides/Proteins

Peptides and proteins are the functional and versatile biopolymers that nature is employing for structure, such as found in hair, fingernails, feathers and chitin (scaffold), as well as function (enzymes, hormones). In general, peptides are smaller species of up to 50 amino acids, whereas proteins can consist of several hundred amino acids. Nonetheless, both consist of a linear sequence of a combination of the 23 amino acids offering an enormous chemical space (20 canonical amino acids). The number of different sequences consisting of three amino acids already accounts for $20^3 = 8000$. As shown in Fig. 1.6, the moieties of the amino acids exhibit different properties: Threonine, serine and asparagine are for example polar amino acids which can interact by polar interactions and hydrogen bonds while phenylalanine, leucine and valine are hydrophobic, non-polar amino acids. Their interaction is mostly dispersive. Moreover, the moieties also differ in their steric aspect: alanine and glycine are very small and thus can serve as hinges within the sequence, whereas asparagine or phenylalanine possess large residues and therefore need more space, thus, making the whole se-

quence stiffer due to steric hindrance. It is this sequence – the so-called primary structure - that already specifies the final shape and function. Therefore, it serves as “natural blueprint” for the molecule.

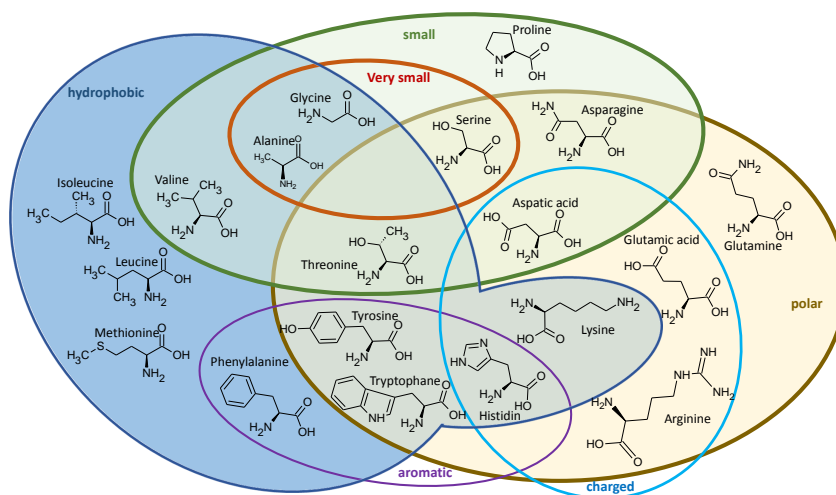


Figure 1.6: Chemical structure of the amino acids ordered in a Euler diagram according to their properties.

In the secondary structure, the shape of the peptide backbone is specified. For long sequences, this can also happen on parts and is not necessarily for the complete sequence. Common structures are, e.g., the α -helix or the β -sheet. These secondary structures are defined by hydrogen bonds between the peptide backbone. As the partial double bond character of the peptide bond keeps the peptide planar,¹⁰⁷ the conformation can be described by the two dihedral angles of an amino acid in a peptide/protein.

Developed by Ramachandran and Sasisekharan in 1963, the so called Ramachandran diagram of the dihedral angles Ψ and ϕ gives energetically stable regions that are characteristic for specific structures, such as the β -sheet or the α -helix.¹⁰⁸ In practice, these plots can illustrate which angles are possible for a residue in a protein (compare glycine with proline in Fig. 1.7) or validate a structure that is found.¹⁰⁹

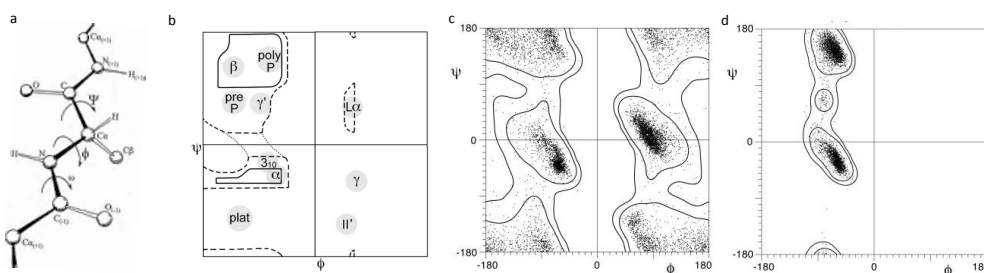


Figure 1.7: Ramachandran plot of amino acids reproduced from [109]. **a:** The two dihedral angles of the peptide backbone that are used as parameters. **b:** the general Ramachandran plot gives regions of allowed backbone conformations (α -helix, β -sheet,...) for different angles Φ, Ψ . **c:** Ramachandran plot for glycine demonstrating a huge variability in the backbone conformation. **d:** Ramachandran plot for proline showing the restricted conformations of this amino acid and its preference for the α -helix conformation.

In the tertiary structure, the conformation of the moieties are evaluated and the final 3D shape of the peptide can be distinguished. In the natural aqueous environment, globular structures arise driven by hydrophobic interactions as well as specific interactions such as disulfide bonds, salt bridges and hydrogen bonds.

The same interactions drive the formation of the quartary structure. This is the protein structure representing the fully functional assembly which can consist of more than one peptide or protein. For example, hemoglobin is a tetramer consisting of two α subunits and two β subunits.¹¹⁰ The process of finding the functional shape due to the linear sequence is called folding.¹¹¹

1.4.2 Glycans

The second class of biopolymers being considered in this work are glycans. The building blocks of glycans are monosaccharides. There are a huge number of monosaccharides, that can be used in a sequence.^{39,41} The most common structure is the hexose $C_6H_{12}O_6$, which exist as aldohexose or ketohexose, depending on the position of the carbonyl group. All hexoses are isobaric, as they have the same composition. The aldohexose exhibits an aldehyde group at the end. Thus, there are 4 chiral centers which results in $2^4 = 16$ stereoisomers. However, only the D isomers occur in nature. The aldohexose glucose is shown in both chiralities, D-glucose and L-glucose, in Fig. 1.8a. Fig. 1.8b depicts the ketohexose D-fructose.

The most stable form of an aldohexose and ketohexose is the cyclic form of a hemiacetal or hemiketal, respectively. The cyclization can form a six-membered ring (pyranose) or a five-membered ring, which is called the furanose form. By this cyclization, another center of chirality is formed, e.g. at the C1 carbon for aldohexoses. The two different configuration isomers are called anomers as they differ in the anomeric position at the C1 carbon and are denoted as α and β . The spontaneous ring opening and closing of the hemiacetal or hemiketal allows for equilibration between the α and β anomers and is called mutarotation.

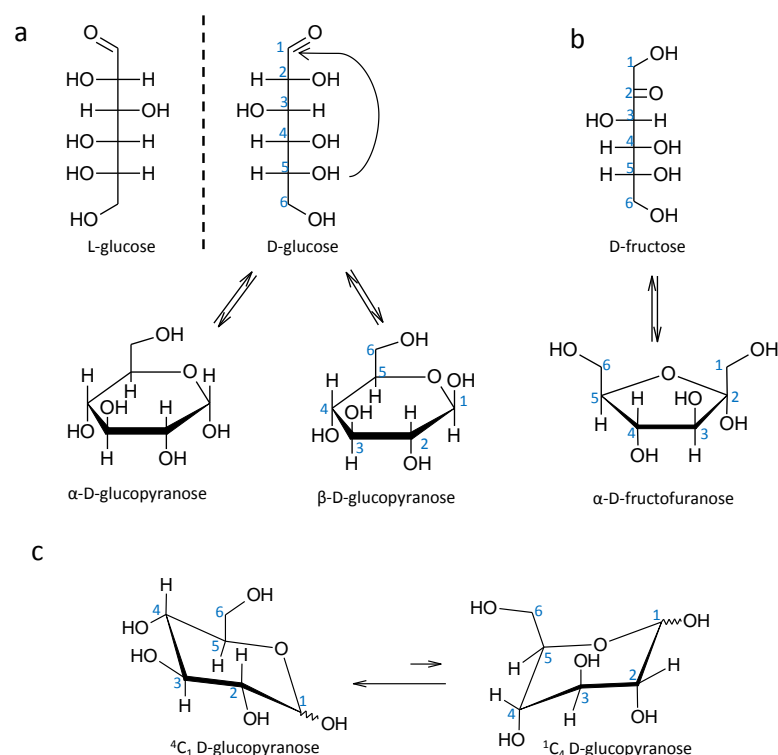


Figure 1.8: Structure of monosaccharides. **a:** The enantiomers L-glucose and D glucose are shown. By ring-closing, the anomers α and β are created, depending on the position at the C1. **b:** Linear and cyclic structure of D-fructose. **c:** different conformations of the cyclic D-glucopyranose. The 4C_1 conformation (left) is energetically preferred over the 1C_4 conformation.

In the cyclic form, there is also some conformation variation possible. I will exemplary discuss this for the pyranose form which is related to the conformation of cyclohexane and tetrahydropyran. Compared to the possible conformations of chair, envelope, twist and boat conformation, the most stable form for a six-membered ring is the

chair conformation. It exists in two conformations, as shown in Fig. 1.8. Depending on which C atom is above or below the main plane, it is labeled 1C_4 or 4C_1 , both shown in Fig. 1.8 c. As large residues are energetically preferred in equatorial position rather than in axial position, this determines which conformation is the most stable. For instance, D-glucose is mostly found in 4C_1 conformation.

In principle, every hydroxyl group can act as condensation site to connect with another monosaccharide. The new bond that is formed is named glycosidic bond. Due to the many possible sites for condensation, a huge variety of different di-, tri- or polysaccharides is possible that all exhibit different properties. Particularly to mention is the condensation at the anomeric position as this changes the hemiacetal to an acetal and thus prevents the ring opening. Saccharides that are completely prevented from ring-opening are called non-reducing, in contrast to the reducing saccharide that still have some open-ring structure. In Fig. 1.9 a-c, a small variety of disaccharides based on only glucose is given that exhibit different properties. The disaccharide trehalose (Fig. 1.9 a) is a non-reducing disaccharide created by a α, α -1-1 glycosidic linkage. Cellobiose (β -1-4 glycosidic linkage), which is a fragment of the polymer cellulose, and gentiobiose (β -1-6 glycosidic linkage), only differ in the condensation position. The second monosaccharide here is not specified as α or β due to the possible mutarotation.

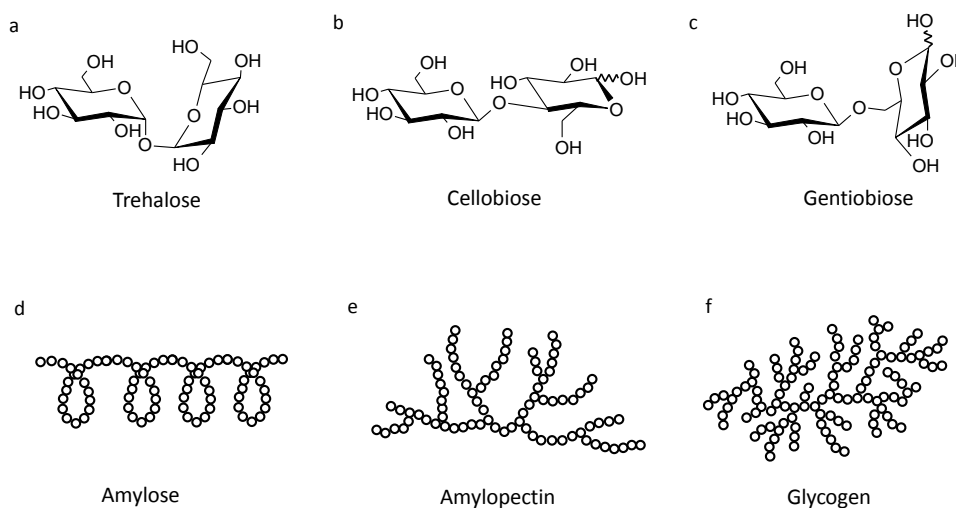


Figure 1.9: Structure of different disaccharides and polysaccharides based on D-glucose as building block. **a:** Trehalose, a non-reducing disaccharide with 1,1 glycosidic linkage. **b:** Cellobiose, linked by 1-4 glycosidic bond is a fragment of the polymer cellulose. **c:** Gentiobiose is linked by 1-6 connection. **d,e,f:** Different levels of branching can change the properties (amylose; amylopectin; glycogen).

With more than three monosaccharides connected, branching is possible if one monosaccharide is linked to more than two other monosaccharides. This can have drastic effects, as e.g. seen in the structural differences between amylose, a linear polysaccharide, amylopectin, exhibiting some branching, and glycogen, which is highly branched. All three polysaccharides have the same building block, which is α -D-glucose. Polymerization is realized by (1-4)linkage whereas branching in amylopectin and glycogen happens by (1-6) glycosidic bonds.

Coupled with the high variety of glycans, this class covers a huge range of function, from energy storage and structuring purposes, such as found in starch or cellulose, respectively, to the identification and adsorption of cells. Up to today, the identification of glycans extracted from an organism is very difficult. Often, the composition is investigated by antibody arrays as the isobaric nature of the building blocks impede a straight forward analysis by mass spectrometry.

2 Saccharides: Self-assembly and characterization of isomers

Most of the world's biomass is made up from carbohydrates. The most abundant macromolecule on earth is the carbohydrate cellulose, with more than 10^{12} tons synthesized on our planet each year.³⁶ Carbohydrates are found in plants and trees, and are part of every living creature, from the human species to bacteria and viruses.^{38,112} They are one of the most important energy sources, as found manifold in nutrition (sugar, starch,...) and are being discussed for biofuels (corn, sugar cane). Furthermore, polysaccharides such as cellulose are employed as structural component in cell walls.¹¹³ Due to their outstanding mechanical properties when assembled in fibers, they are used for textiles (cotton consists of 90 % of cellulose), paper or packaging material.¹¹⁴

Besides their macroscopic use, carbohydrates have lately gained a lot of attention due to the growing field of glycomics.³⁹ It has been shown that an immense variety of glycans exists and is employed for different functions within the cell. Glycosylation of proteins and peptides occurs very specifically and manipulates the folding and thus the function of proteins.³⁷ Moreover, glycans are involved in almost every biological process, including signal transducing, cell adhesion, cell localization and differentiation, and play a role in nearly every human disease.^{115–117} For most of these functions, glycans are located at the outer surface of cells where they can interact with other species.¹¹⁸ However, the structural and functional aspects of these glycans are only poorly understood. Their sequence is often difficult to identify due to the limited amount of material that can be extracted from biological samples, mostly without purification, as well as to the huge number of different species and their highly complex structures.⁴¹

In contrast to other biopolymers, such as oligonucleotides (DNA) and proteins, the

sequence of saccharides can be highly complex (see also chapter 1.4 "Molecules and Structure"): Not only are the basic building blocks for carbohydrates intriguingly similar and even isobaric as they are stereo- or regioisomers of the chemical sum formula $C_6H_{12}O_6$, but also can the sequence be branched and differ in linkage position. Fig. 2.1 shows the molecular structure of the two disaccharides, trehalose (α -D-gluco-pyranosyl-(1-1)- α -D-glucopyranoside) and sucrose (α -D-glucopyranosyl-(1-2)- β -D-fructofuranoside), that will be investigated in this chapter. Both molecules are non-reducing disaccharides, forming a stable aqueous solution where no isomerization reactions due to ring opening are to be expected. The two molecules only differ in the second monosaccharide building block: Trehalose is built by a condensation of two glucose rings at the 1, 1' position. It possesses C_2 symmetry around the glycosidic linkage in some conformations due to the usage of the same building block. In sucrose, the first glucose unit is the same as for trehalose, while the second monosaccharide is fructose, in its furanose form (5 membered ring). Both building blocks are structural isomers with very similar dimensions, accounting for 0.83 nm and 0.92 nm in length for sucrose and trehalose respectively.

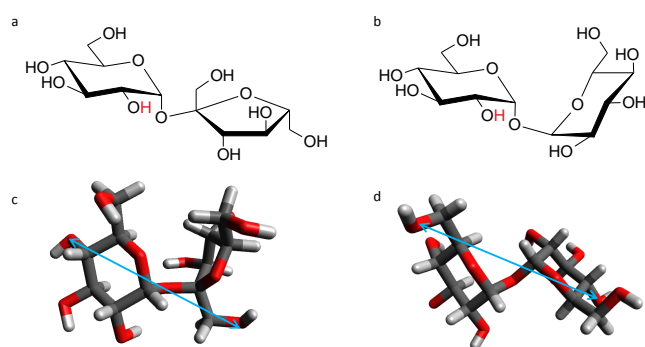


Figure 2.1: Structure of sucrose and trehalose. **a, b:** Haworth projection of sucrose (a) and trehalose (b). The red H marks the deprotonation site. Explicit hydrogens were removed for clarity. **c, d:** stereoscopic view on sucrose (c) and trehalose¹¹⁹ (d). The distance between arrow reflects the length dimension of the molecule to be 0.83 nm and 0.92 nm, respectively.

The vastly different behavior of carbohydrates (such as sweetness, melting point, etc.) arises from rather subtle differences in the stereochemistry of the monosaccharides as well as in different linkage between the single monosaccharide building blocks. While sucrose is mostly used as sweetener and energy source, trehalose is the main energy storage in insects and is also employed for its anhydrobiosic properties.^{119,120} In general, the structural characterization of carbohydrates is very challenging, due

to the subtle difference in structure and the limited amount of material.³⁸ So far, the characterization is mostly performed either on crystalline samples for X-ray diffraction or neutron scattering^{121–123} or by NMR for molecules in solution¹²⁴ or by gas chromatography and MS in the gas phase^{125,126}. While XRD as well as NMR need large amounts of sample material, MS is more suitable in this case. However, the mass of the molecule by itself is not sufficient information to identify a carbohydrate because most of constituting monosaccharide building blocks are isobaric. In mass spectrometry, the sequence as well as the regiochemistry is mostly accessible by fragmentation pattern in the gas phase.¹²⁶ Only recently has it been demonstrated that ion mobility mass spectrometry (IMS) is able to differentiate between sterically different structures by measuring the collision cross section (CCS).^{42,127}

In principle, scanning techniques are able to image the structure of a single molecule⁵⁵ and thus obtain structural information, such as connectivity or branching. So far, the structural alignment of polysaccharide crystals has been investigated by electron microscopy and AFM.^{128–130} Due to the periodicity in the crystal, it was possible to recognize single monosaccharide building units in the AFM measurements. By transferring molecules on a surface, a grown crystal would not be needed for imaging by scanning probe microscopy.

In many experiments, it has been demonstrated that STM is able to image single molecules as well as 2D assemblies.^{53,54} Moreover, employing STM, it could be demonstrated that this imaging technique is an appropriate analytical tool to distinguish between different stereoisomers for different adsorbed molecules,^{131–133} as well as obtaining sub-molecular resolution routinely. However, to achieve high resolution in STM, the environment and the sample has to be well-defined at the atomic level, thus working in UHV is necessary. As glycans are non-volatile and cannot be evaporated onto the sample, ES-IBD is employed to transfer the molecules into the gas phase and onto the sample in UHV.

In this chapter, I present the first STM investigations of saccharides on metal surfaces and explore the structural characterization capabilities of STM. To this end, two disaccharides, namely trehalose and sucrose, are chosen as samples with reduced complexity and are deposited on the surface by ES-IBD. I explore the adsorption and 2D crystallization of these disaccharides on clean metal surfaces by STM in UHV. Moreover, I look for characteristics of single molecules in order to differentiate between different isomers. For this, the comparison between sucrose and trehalose is also suitable as it is a comparison between a homosaccharide and a heterosaccharide

with different ring sizes.

2.1 Electrospray mass spectrometry of disaccharides

In the last decade, ESI-MS became quite popular and established as analytical tool for carbohydrates.^{125,126} The established MS of carbohydrates demonstrates that ESI is possible and gives good starting parameters. However, for deposition, the focus lies on intensity rather than on analytics. Thus, the solution and ESI-parameters have to be optimized for high flux and intensity of the wanted species.

To explore the ESMS of disaccharides, we chose maltose, sucrose and trehalose as disaccharides to investigate. Fig. 2.2 shows the mass spectra for all three molecules in positive (left) and in negative mode (right). The solutions were prepared in a 1:1 ratio of water and ethanol to a concentration of 10^{-4} M of the disaccharide species. Subsequently, formic acid (1 % v/v) or NH_4OH (2 % v/v) was added to enhance the ionization in positive or negative mode, respectively.

In positive mode, the molecular peak at 342 Th cannot be observed. Instead, the most prominent peak is at 281 Th, which is a fragment formed by the exclusion of a species with 60 Th. This might be due to the added formic acid, which promotes ring opening and thus enables the exclusion of a $C_2H_4O_2$ species. Due to its abundance and reproducibility, the positive mode is often used to study the $[M + Na]^+$ adduct and its fragmentation. In our experimental set-up, we also obtained this species at 364 Th. While the intensity of the adduct species is very low when only formic acid is added, it can be enhanced by adding NaCl instead of formic acid (MS not shown here). Moreover, cluster peaks at 534 Th ($[3M + 2Na]^{2+}$), 705 Th ($[2M + Na]^{1+}$) and 877 Th are also obtained for all beams in positive mode. It has been shown that these clusters carry most of the material in the beam.¹⁰⁵ Moreover, IMS measurements of the adduct species always resulted in smaller cross section evidencing that carbohydrates bind strongly to the metal adduct by coiling around it.⁴² However, for the first attempt to clarify the structure of carbohydrates by imaging, extended molecular species are favored. Thus, the positive mode for disaccharides is not chosen as deposition mode because it does not meet these requirements.

2.1. Electrospray mass spectrometry of disaccharides

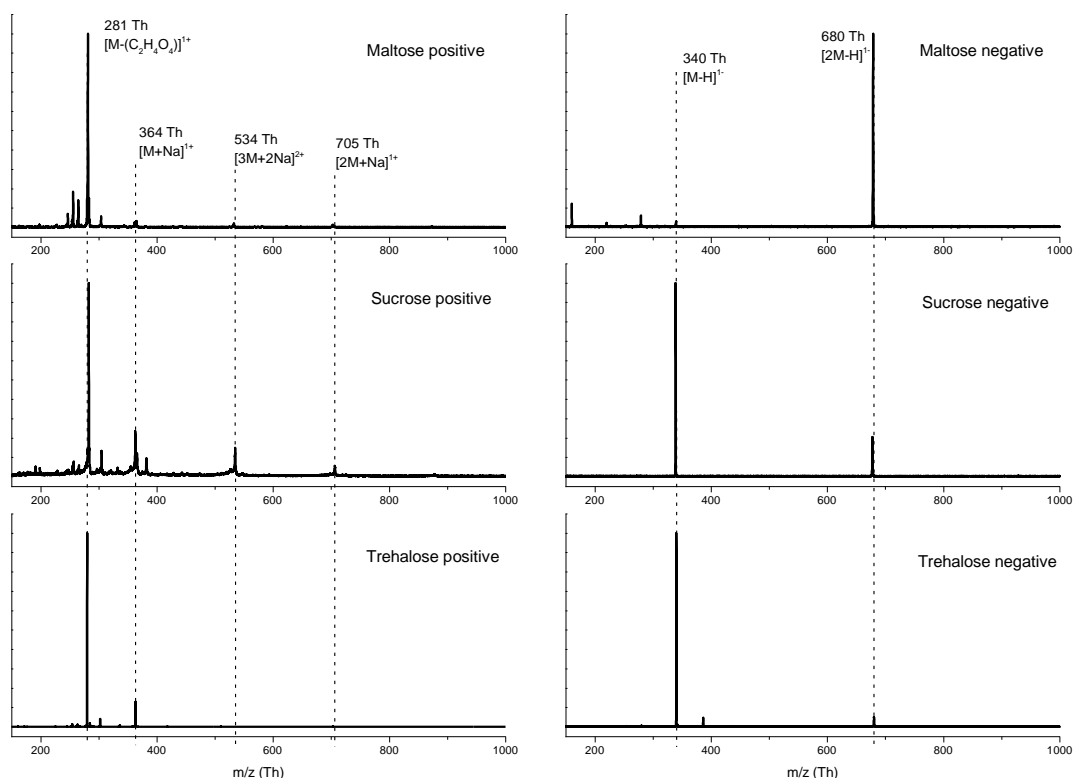


Figure 2.2: Mass spectra of disaccharides. On the left panel are the mass spectra for positive mode, on the right side the negative mode mass spectra can be found of maltose (top), sucrose (middle) and trehalose (bottom). Important species are labeled.

In negative mode (right panel in Fig. 2.2), the deprotonated molecular ion $[M - H]^{1-}$ can be found at 340 Th for all three molecules. In contrast to the adduct formation, the position of the deprotonation is controlled by the acidity of the hydroxyl groups and is thus very specific for each molecule.¹³⁴ For sucrose and trehalose, the deprotonation occurs at the C₂ hydroxyl group of the glucose unit (marked red in Fig. 2.1). This is advantageous for deposition because it ensures that only one ion species is created and subsequently deposited. For sucrose and trehalose, the $[M - H]^{1-}$ ion is the main peak, while it is only scarcely present for maltose. Instead, the main peak for maltose is the two-molecule cluster peak $[2M - H]^{-1}$ at 780 Th. However, the intensity of the molecular ion can be enhanced for maltose using an ESSI source¹³⁵ (not shown here). In order to keep the same spray conditions, sucrose and trehalose are chosen for deposition. With the beams established, 30% - 50% of the current before mass selection is preserved after mass selection accounting for ~100-400 pA on the TOF aperture.

However, we noticed that the stock solution can produce more clusters upon aging

in the time scale of months, thus, the obtained mass selected current is lower. For abundant molecules, such as the disaccharides, stock solutions were prepared freshly before each experiment to avoid clustering. For more complex carbohydrates, the clustering can be prevented by sonicating the stock solution before use.

2.2 2D crystallization of disaccharides by self-assembly

Here, we investigate the adsorption and self-assembly of trehalose and sucrose on an atomically clean metal surfaces by STM. These molecules exhibit a 3D geometry, as shown in Fig. 2.1c,d and are rather flexible. Thus we need a combined approach of experimental STM data and theoretical calculations to decipher the adsorption geometry of the molecule on the surface.

The molecular ion (mass 340 Th) is mass selected for deposition and a landing energy of 5 eV is ensured by applying a decelerating potential on the sample. For sub-monolayer coverages, a coverage of 75-180 pAh were deposited over 1-3 hours.

2.2.1 Ag(111): Self-assembly into closed packed structures

On Ag(111), deposited trehalose condensates into large islands of round shape. Fig. 2.3a shows that some molecules are still mobile at room temperature, resulting in streaks in the STM image and fuzzy borders of the islands. Within the island, the molecules are fixated and a substructure is observed, highlighted by an arrow in Fig. 2.3a.

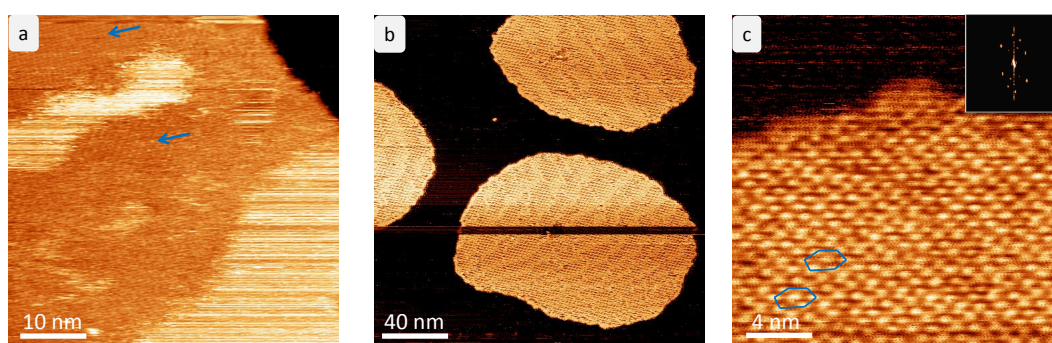


Figure 2.3: STM images of trehalose on Ag(111). **a:** At room temperature, islands can be observed with an internal periodic structure as highlighted by the blue arrows. **b:** At 40 K, isotropic growth of the islands is observed. Additionally, domain boundaries within the islands can be seen. **c:** Detailed STM image of a domain with molecular resolutions depicting a hexagonal assembly of the trehalose molecules.

Upon cooling to 40 K, all molecules are completely assembled into islands, as shown in Fig. 2.3b. The islands are of round shape and exhibit smooth edges. The isotropic growth suggests that the formation is not kinetically hindered but a thermodynamic minimum.¹³⁶

Within the island, the assembly is crystalline as can be seen by the periodic substructure in the islands in Fig. 2.3c. A hexagonal close packed structure is observed that is evident by the Fourier transformation of Fig. 2.3. Due to the drift in the measurement, the structure is distorted. Relating the size of the features (1 nm) with the molecular dimensions from the 3D structure (0.92 nm in length), we conclude that each feature corresponds to one single molecule. Thus, on Ag(111), trehalose molecules assemble in a hexagonal close packing. As no submolecular resolution is achieved, the adsorption geometry of the molecule can only be deciphered by theoretical input (as discussed later). Having hydroxyl groups at the periphery, an assembly stabilized by hydrogen bonding can be assumed and can also explain the close packing of the molecule because this would permit the highest the number of stabilizing intermolecular interactions.

2.2.2 Cu(100): Self-assembly into porous structures

In contrast to Ag(111), Cu(100) is a more reactive surface with a four-fold symmetry. Most of the investigations were performed on this surface.

Trehalose on Cu(100)

Already at room temperature, ordered anisotropic structures are observed for trehalose on Cu(100) by STM. In the time series in Fig. 2.4 a-c, some fluctuation of the assembly is observable on the time scale of a scan (2-6 min) demonstrating that the molecules are still mobile on the surface. The fluctuations and mobility of the adsorbates are also observed in the streaks in the STM images.

Within the assemblies, features are observed that are 0.89 nm in length and 0.56 nm in width. The dimensions fit very well to the expected dimensions of a single molecule, thus, each feature is attributed to a single molecule. In the STM images, three different motifs (A-C), in which the molecules can assemble, are identified and marked in Fig. 2.4b. A zoom-in into these motifs as well as the corresponding models is given in Fig. 2.5.

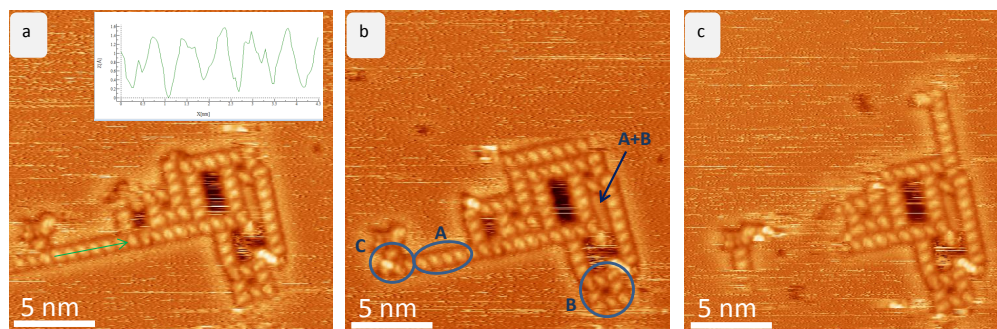


Figure 2.4: Time series of assembly of trehalose on Cu(100). Consecutive scans are taken after 6 minutes. **a:** The inset shows line scan as marked with a green arrow in the image. **b:** 3 different motifs (A-C) are marked in the image.

Motif A is an anisotropic assembly that is formed by parallel aggregation of molecules as shown in Fig. 2.5a. The in-line spacing of 7.6 Å is depicted in the line scan in Fig. 2.4a. Within the lines, the molecules are tilted by 39°–54° with respect to the orientation of the linear assembly. In agreement with the fourfold surface symmetry, the lines are formed in orientations which are rotated by 90° to each other. If two lines assemble next to each other, they are staggered with an interline spacing of 1 nm.

Motif B (Fig. 2.5b) consists of eight molecules in a squared arrangement. Four of the molecules, which possess a blue outline in the cartoon in Fig 2.5e, form a chiral windmill motif exhibiting a pore in the middle. The remaining four molecules are located in the corner of the square, where they align with the neighboring molecule. Due to the agreement in orientation between motif A and motif B, both motifs can transition into each other, as sketched by the light green molecule in the cartoon of motif B that starts a line of motif A.

A third motif, motif C, is observed in some assemblies. It is characterized by a bright inner feature and up to four surrounding molecules. At room temperature, the location of it appears irregularly, but is mostly found at the corners of motif B, terminating the assembly (best seen in Fig. 2.6a).

2.2. 2D crystallization of disaccharides by self-assembly

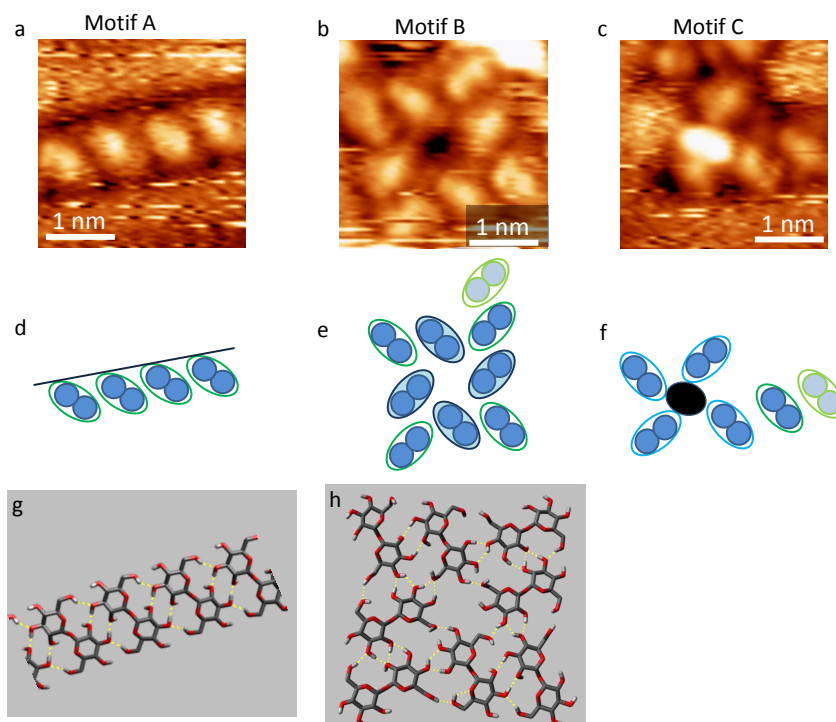


Figure 2.5: Structural motifs of trehalose on Cu(100). **a, b, c:** Zoom into the STM image depicting motifs A, B and C of Fig. 2.4. **d, e, f:** Cartoons illustrating the motifs. Each molecule is depicted as elongated lobe with two circles. **g, h:** MD optimized structures of the motif A and B, respectively.

One has to note that all motifs A-C are chiral and only one enantiomer is found in consequence of the deposition of the enantiopure molecule trehalose. For motif B and C, this is easily seen in the windmill formation which is only observed in one rotational direction. The single molecule is found in two different orientations which is in agreement with the 4fold symmetry of the surface. Due to the limited resolution and the tilt angle close to 45° , they can be mistaken as mirror images, but they are rotated by 90° . This is especially important in order to understand the chiral alignment in motif A. All horizontal lines consist of molecules that are tilted from top left to bottom right, whereas vertical lines of motif A consist of molecule rotated by 90° (top right to bottom left). The mirrored assembly does not exist because there is only one enantiomer on the surface.

Due to the variability in conformations of the trehalose molecule and the multitude of possible intermolecular interactions, theoretical support is necessary to decipher the molecular conformation within the self-assembled motifs. For this structure clarification, MD simulations were performed¹ using GROMACS¹³⁷ with the 53A5 force

¹in collaboration with Ludger Harnau

field¹³⁸. The 3D structure of trehalose in its anhydrous form, Treh(a), from the crystal structure¹¹⁹ was put on the surface and relaxed to yield a quasi-2D conformation. This quasi-2D conformation was subsequently used to model the assembly on the surface, starting with initial guesses based on the constraints from the STM images. During all simulations, the substrate Cu atoms were frozen in their crystallographic position.

Fig. 2.5g, h present the simulated assembly motifs A and B for trehalose, respectively. For all different motifs, the conformation of the molecule is the same and very similar to the employed starting conformation. The molecule still possesses a 4C_1 conformation in both rings and an approximated C_2 symmetry. In order to enhance the interaction with the surface, the amount of hydroxy groups pointing to the surface is maximized. To this end, the oxygen atoms in both rings point away from the surface, while 6 of the 8 hydroxy groups and the bridging oxygen can interact with the surface. In this conformation, several hydrogen bonds can be formed within the assemblies stabilizing their formation.

In motif A, the linear assembly is obtained by a close stacking, as shown in Fig. 2.5g. Due to the tilt of 59° , the molecules can interlock and thus enhance the hydrogen bonding. The in-line spacing of 7.7 \AA agrees well with the experimental value. In total, 8 intermolecular hydrogen bonds to the adjacent molecules are formed and stabilize this arrangement. However, the last molecule in the chain is only stabilized by half the number because it only has one neighboring molecule. This makes the last molecule prone to disassemble and can therefore result in the structural fluctuation that was observed experimentally in the assemblies at RT (see Fig. 2.4).

For motif B, the four inner molecules (blue outline in Fig 2.5e) of the windmill motif are stabilized by the most hydrogen bonds, as depicted in Fig. 2.5h. Within the windmill motif, they form 2-3 hydrogen bonds to the adjacent molecules via the hydroxyl groups of the ring closer to the center. Additionally, they interact via hydrogen bonds with the molecules in the corners, forming a pattern that resembles motif A. The corner molecules are only stabilized by 5-7 hydrogen bonds that are formed with the two adjacent molecules of the windmill motif, leaving 2 sides of the molecule vacant for other intermolecular interactions. The long side of the molecule that is vacant can be employed for the transition between motif A and motif B (compare to the cartoon). In general, the distribution of intermolecular interactions to several partners makes the motif less likely to form, but yields more stable structures because more interactions can be obtained.

Additionally, combinations of the different motifs are observed. A combination of

2.2. 2D crystallization of disaccharides by self-assembly

motif A and motif B results in hollow rectangles of different dimensions as marked with an arrow in Fig. 2.4b. In these formations, the length of motif A determines the dimensions whereas the corners are in the same manner as in motif B. Other combinations, such as motif A+B+C are possible, as well as two merged octamers of motif B, by sharing either the edge or the corner, as depicted in Fig. 2.6a .

Interestingly, the same motifs A-C are also observed at low temperatures (40 K). Upon cooling down, extended regular structures are formed consisting of the combined motifs A+B+C, as shown in Fig. 2.6b. While the pure motifs are barely found, the combination is frequently observed to form large islands with rectangular pores separating the lines of motif A. Within the pores, some mobile adsorbates are trapped. Moreover, unstructured agglomerations are found on the surface as well as small features (1 nm in size) with a repeating shape, as marked by yellow circles in Fig. 2.6b. Due to the repeating orientation and the consistency in dimensions, these features are assigned to single molecules on the surface. The adsorption of single molecules is discussed in a later section of this chapter.

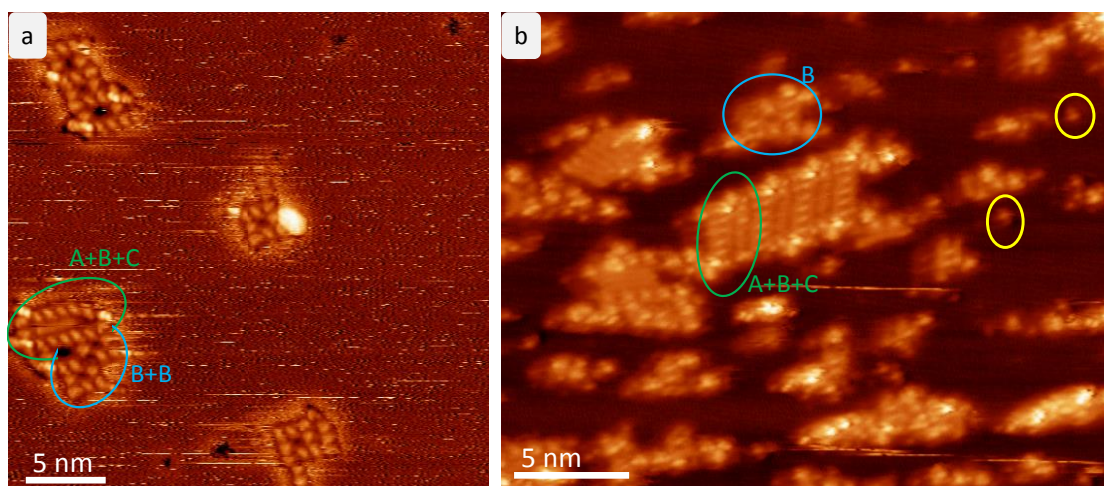


Figure 2.6: STM images of the nanostructures of trehalose on Cu(100). **a:** nanostructures at room temperature showing the combination of the different motifs A-C. **b:** STM image at 40 K depicting the same motif B and an often observed assembly consisting of a combination of motif A -C. Moreover, unstructured agglomeration as well as single molecules (yellow circle) are observed on the surface.

Sucrose on Cu(100)

In contrast to the homosaccharide trehalose, the heterosaccharide sucrose is built up by two different monosaccharide building units. It is formed by a 1,2-condensation of α -glucose and β -fructose, which differ by the size of the ring. Due to the two alkyl chains attached to the C₂ and C₄ in the fructofuranoside, this monosaccharide building block has a higher steric aspect.

At room temperature, the molecules are mobile on Cu(100) yielding no stable structures that could be observed by STM. After cooling to 40 K, we observe nearly exclusively 2D islands as shown in Fig. 2.7a. Most of the molecules are adsorbed at the step edges and grow islands from there. Some islands are also found in the middle of the terraces. As shown in Fig. 2.7b, these islands are made up from a porous network and exhibit frayed ends. Another phase, that will not be discussed here, is seen at the left side of the island as a linear strand exhibiting a zig-zag assembly. The porous assembly is formed by a windmill motif of elongated lobes. This motif becomes clearly visible in the magnified STM image in Fig. 2.7c. Due to the appropriate dimensions of the elongated lobe, that is 1 nm in length and 0.5 nm in width, we assign it to a monomer. Fig. 2.7c, the molecule is resolved as two separate round features that differ in height by ~ 0.3 Å. They correspond most likely to the single monosaccharide building units. In the cartoon that is overlaid to the STM image in the lower right, the molecule is sketched as green oval with two circles of different color, corresponding to the different intensity. Interestingly, in the 2D network, the molecule is aligned in the way that each node of the windmill is formed by only one kind of feature: all bright features come together to one node (blue box) or the node consists of only darker features (violet box). Thus, there are two different nodes, which are formed by exclusively either the glucose or the fructose building block. Moreover, these nodes also differ in their compactness: the low-feature node measures 0.9 nm in diagonal, whereas the bright feature node measures 1.1 nm diagonally across the node. In Fig. 2.7d, the two windmill motifs are presented separately. However, no assignment of the glucose or fructose unit is possible by the STM data alone. The non-symmetric appearance is due to some drift in the image.

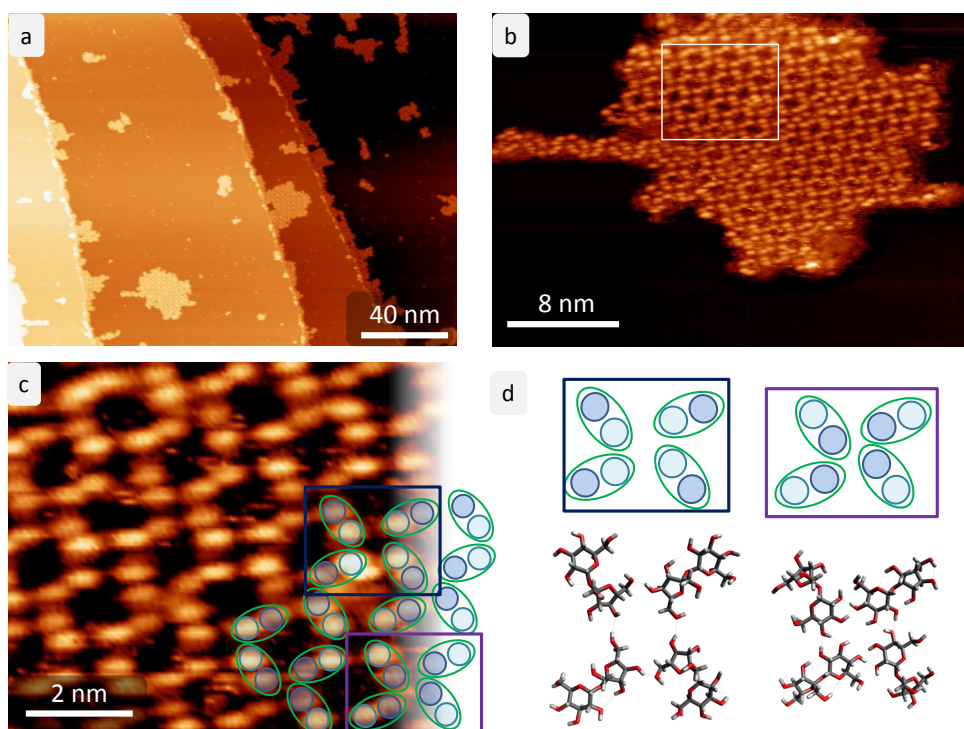


Figure 2.7: Self-assembly of sucrose on Cu(100). **a:** Islands are observed nearly exclusively on Cu(100). **b:** STM image of one island is presented. The island is mostly formed by a porous network. **c:** Zoom into the white rectangle in b, resolving the single monosaccharide building blocks in the molecule. A cartoon is overlaid in which the molecule is represented by a green oval. The two monosaccharide units are depicted as round circles with different color representing the different intensity. **d:** The two different nodes of the windmill motif are sketched (top). The corresponding MD optimized structures are shown below in top-view.

Employing the same structure clarification approach as described for trehalose, we find a preferred adsorption geometry for sucrose on Cu(100) that is also used in the structures of the windmill motifs in Fig. 2.7d. Similar to trehalose, the glucose unit possesses a 4C_1 conformation to enhance the interaction of the hydroxyl oxygens with the surface. Due to the steric demand of the two alkyl groups, the furanose ring of the fructose building block cannot adsorb flat on the surface but is tilted. In this geometry, the alkyl groups at the C_2 is pointing away from the surface.

This conformation might explain the different contrast we observe within the single molecules, thus enabling to distinguish the different monosaccharide units. Assuming that the electron density is equally distributed within the two rings of the molecule because the hydroxyl groups are everywhere in the periphery, we can neglect the electronic effect. Moreover, the topography has a stronger effect (exponential) on the

convoluted measurement that the STM represents than the electron density (linear) as shown in eqn. 1.6. Therefore, in this case, the high intensity in the STM image can approximately be attributed to the topography. In combination with the MD optimized molecular structure for sucrose adsorbed on Cu(100), we can thus deduce in a rough approximation that the fructose unit corresponds to the brighter feature. With this adsorption geometry, the structure of the porous network can also be clarified in the model shown in Fig. 2.7d. Here, the rather flat glucose rings come together to form a node which is stabilized by four hydrogen bonds per molecule (right structure in Fig. 2.7d). At the other end of the molecule, the fructose units also form a node, which is, however, only stabilized by two hydrogen bonds per molecule. In total, six hydrogen bonds are formed. Additionally, the assembly is most likely also stabilized by vdW interaction between the hydrocarbon ring of the sucrose molecule.

Influence of charge and substrate on the assembly

So far, the MD modeling was performed assuming neutral (not deprotonated) species, yielding reasonable structures for trehalose as well as for sucrose. This simplification is done because the estimate of the distribution and amount of charge on the molecule is very complex. However, due to the deprotonation occurring in the ESI process, the molecules are deposited on the surface as deprotonated, hence negatively charged species. While the charge of the molecule on the surface might be reduced after deposition by charge transfer to the surface, the molecule will stay deprotonated. The deprotonated state of the molecule might also have impact on the self-assembly due to an enhanced dipole moment of the molecule or a stronger binding to the surface. As observed on Ag(111), disaccharides can assemble in an isotropic close-packed structure, demonstrating that the possible charge on the molecule does not repel the molecules. Changing the substrate to Cu(100), results in anisotropic, porous assemblies with different motifs. This might be due to the influence of the substrate, including binding strength, symmetry as well as commensurability. As shown with flat, aromatic molecules, the influence of the substrate can have a large influence on the self-assembly.¹³⁹

Another possible explanation for the open network structure is that the deprotonated hydroxyl group (hydroxylate) is coordinating to a metal center that is incorporated from the surface adatoms. This process is well studied and abundant for carboxylate functional groups on Cu surfaces, as Cu possesses reactive surface adatoms.^{140,141} Metal organic coordination networks (MOCN) often possess a porous assembly be-

cause of the directional character of the coordination bond.

2.3 Single molecules

So far, I have presented the investigation of the 2D crystallization of disaccharides on metal surfaces. In order to characterize and differentiate a molecule, however, the study of single molecules is advantageous as there are no neighboring effects due to surrounding molecules. As a cold deposition and a cold transfer into the STM is not possible so far, the challenge is to separate the molecules or deposit them in a way that they stay separated. Trying out different substrates (Au(111), Cu), this approach was not successful for sucrose and no single molecule could be observed. Additionally, the manipulation with the tip to drag a single molecule out of the 2D assembly also did not yield the desired result.

For trehalose, small features with reoccurring size, shape and orientation are found on Cu(100) in addition to the 2D networks, as mentioned before and depicted in Fig. 2.8. We assign these species to single molecules due to their abundant occurrence and size (1 nm length, 0.5 nm width) that agrees with a single trehalose molecule. Interestingly, in contrast to the observed shape of the molecule within the assembly, the single adsorbate is chiral with both chiralities observed on the surface, as seen in Fig. 2.8b. Moreover, the appearance of the adsorbed species is different to the one in the assembly as it resembles a drop, with the highest intensity at the round feature of the drop. At the side where the feature tapers, a dark protrusion can be observed. As this protrusion appears at the same position of the adsorbate and for different orientations, this is not a tip effect but relates to the adsorption. This effect is known for carboxylate molecules and appears if a strong binding to the surface is present.^{142–144} While we observe this variation of intensity within the adsorbed species, we cannot resolve the different monosaccharide building units. The contrast to sucrose molecules within the network, for which the building units can be resolved and intuitively identified, suggests that the resolution within the disaccharide depends strongly on the adsorption geometry and the interaction with the surface.

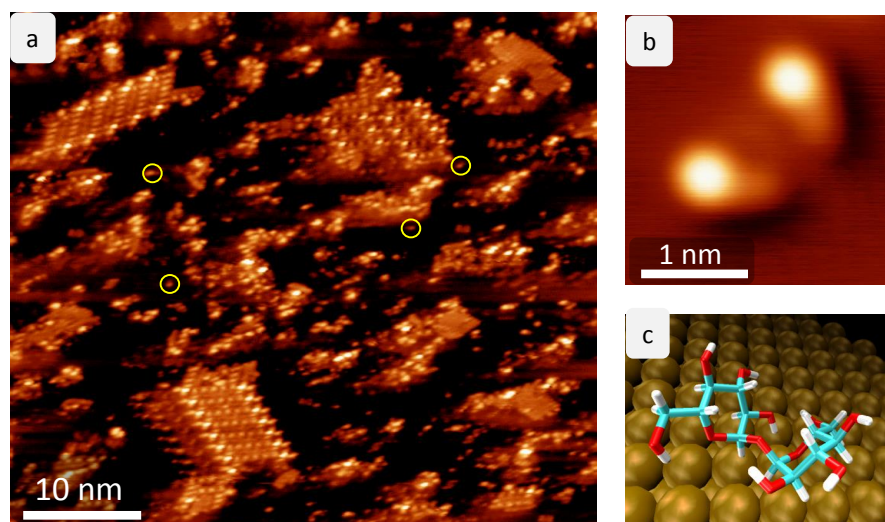


Figure 2.8: STM images of the deposited trehalose molecule on Cu(100) at 40 K. **a:** In addition to the assemblies described above, single molecules are observed and marked with yellow circles. **b:** Magnification of two single molecules with different chirality is presented. **c:** Possible conformation of adsorbed trehalose are found by a global landscape search.

In order to shed light on the different adsorption geometries of disaccharides, a global landscape search is performed² employing a robotics-inspired stochastic method for global optimization and subsequent *ab initio* local optimization of the molecule on the surface.¹⁴⁵ While this is still ongoing work, this approach yields different conformations on the surface, as shown in Fig. 2.8c. Evaluating them by their energy results in a prediction of stable conformations on the surface. Eventually, this will help to understand the adsorption of highly complex molecules on the surface.

2.4 Summary and Outlook

In contrast to peptides or small molecules, only little is known about the adsorption of carbohydrates on surfaces so far, despite being the most abundant class of molecules. In this chapter, we have, for the first time, imaged disaccharides, namely trehalose and sucrose, with submolecular resolution by STM on metal surfaces and explored their adsorption and self-assembly behavior.

The molecules are deposited as deprotonated, negatively charged species because this mode gives the highest intensity in the electrospray and, moreover, the deprotonation site is very specific, yielding only one ion.¹³⁴ Despite their similarity in size and the

²by collaborators JC Schön and J Cortez

location of hydroxyl groups in the periphery of the molecule, the 2D self-assembly is different for trehalose and sucrose. On Cu(100), sucrose forms nearly exclusively porous networks at low temperatures. In these assemblies, the single monosaccharide building block is resolved within the disaccharide. In contrast, three different motifs and their combination are observed for trehalose on Cu(100) as well as unstructured agglomerations. Moreover, single molecules are found manifold on the surface suggesting that the interaction of trehalose is stronger with the surface than the interaction of sucrose.

Employing MD models, the structure of disaccharides on the surface has been investigated, demonstrating that the interaction by hydrogen bonds is primarily responsible for forming the nanostructures. Due to the different monosaccharide building blocks (fructose instead of glucose), the adsorption geometry differs which influences the structure formation and most likely also the spatial resolution that can be achieved by STM for the molecule. This might be due to different interactions with the surface which also influence the hybridization of the electronic structure with the surface.

In order to evaluate the adsorption of disaccharides, a global landscape search is being performed, allowing to study the adsorption on an *ab initio* level.¹⁴⁵ Thus, providing submolecular resolution data will eventually help to evaluate or develop new functions and tools to describe the adsorption of complex molecules on a metal surface.

Moreover, the characterization of single oligosaccharides is a very interesting, yet challenging project to pursue in the near future. In this thesis, I have demonstrated that we can achieve resolution to observe the single monosaccharide building blocks within the molecule. While this depends most likely on the adsorption and interaction with the surface (see sucrose vs. trehalose), it should in principle be possible to resolve the building blocks in oligosaccharides. As a consequence, it is then possible to identify branching positions in longer sequences. This would advance and facilitate the structural analysis of unknown saccharides.

First results are shown in Fig. 2.9. As a first step, we deposited a linear (1,6)-mannose pentamer (structure, see Fig. 2.9a). The *m/z* selected deposition on Cu(100) and subsequent scanning at 5 K yields small agglomerations on the surface, as shown in Fig. 2.9b. Most of these agglomerations are unstructured, but some of them have a very distinct appearance with a straight line and a triangular pattern above it, as marked with blue circles. So far, no structure clarification is performed. However, these first experiments look very promising as small, round features are observed (see Fig. 2.9c)

that fit to the expected size of the monosaccharide building block. In Fig. 2.9c, one of these distinct agglomeration is shown exhibiting two triangular structures. As the pentamer most likely folds up, it could be that one triangular structure corresponds to one folded molecule, as depicted in the overlay of the molecule.

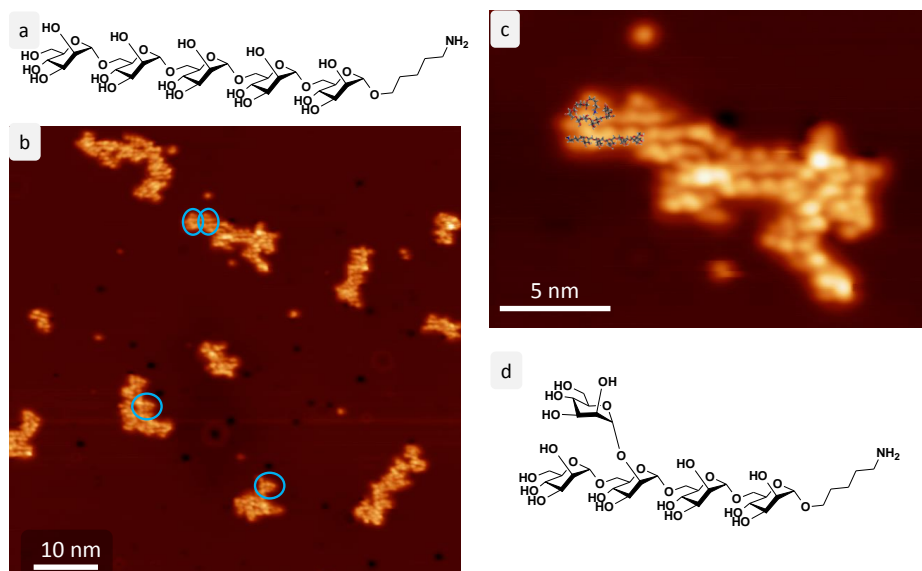


Figure 2.9: First data for mannose pentamer on Cu(100). **a:** Chemical structure of linear (1,6)-mannose pentamer. The alkyl chain with amino functionalization is residual to the synthesis route. **b:** First STM measurements on Cu(100) at 5 K showing agglomerations as well as repeating structures (blue circles) with submolecular resolution. **c:** Proposed model of the pentamer adsorbing on the surface. **d:** Structure of a branched isomer of the mannose pentamer.

In order to facilitate the structural investigation of oligosaccharides, we are trying to decrease the mobility on the surface to yield single molecules. While different surfaces, such as Cu(110) and Au(111), were not successful, we will test out different anchoring groups at the alkyl tail of the oligomer, such as carboxylate, thiol or thiophene. These groups might react with the surface and thus immobilizing the molecule on the landing position. Moreover, with a cold transfer system, cold depositions are possible which also reduce the mobility and might help to obtain isolated species on the surface. Once the linear oligosaccharide can be reliably imaged and characterized, we can compare it to the structures of a branched isomers, such as the one shown in Fig. 2.9d, expecting to see differences due to the branching.

3 Peptides as sequence-controlled building blocks for self-assembly*

3.1 Peptide assembly and nanostructures

In billions of years, nature has evolved a powerful way to build highly functional machines and nanostructures by the self-assembly and folding of peptides and proteins. Peptides and proteins are composed of a linear sequence of amino acids that programs the shape and function of the molecule by non-covalent interactions, such as dispersive interactions, hydrogen bonding and ionic interactions.

In recent years, the engineering of synthetic peptide nanostructures has been extensively investigated. In contrast to DNA, which is used as the prototype for artificially engineered biological nanostructures,^{17,18} peptides are chemical versatile and often exhibit catalytic activity. Moreover, they are flexible and possess a large conformational adaptivity while their synthesis is straightforward due to their sequential structure. This makes peptides attractive as building blocks, but the complex interaction schemes impede a rational engineering of peptide nanostructures.²⁷ Only in 2013, it was demonstrated that distinct 3D cage structures can be engineered by one single protein strand, similar to the vast range of examples shown for DNA origami in the last decade.¹⁴⁷

The engineering of nanostructures by multiple, small oligopeptides has already proven to yield materials with outstanding properties and applications, such as scaffolds for 3D cell culture, coatings, energy conversion and storage or biosensors.^{11,25,148} Several approaches - mostly in liquids - have been applied to yield different structures. As there are extensive reviews about self-assembled peptide nanostructures available,^{12,28,149,150} only a short overview is given here in order to introduce different

*Parts of this chapter are based on a publication¹⁴⁶

motifs for artificial peptide nanostructures.

The first formation of peptide nanotubes was demonstrated by Ghadiri et al. in 1993.¹⁵¹ They employed cyclic peptides with an even number of alternating L and D amino acids. Upon change in the pH, the formation of tubular structures is triggered. Due to the alternating L and D conformation, the residues are exclusively on the outside of the nanostructures. The sub-units can stack in an anti-parallel fashion, giving rise to backbone - backbone hydrogen bonding to produce a contiguous β -sheet structure.

Most other approaches are based on the interaction of the different residues in addition to the β -sheet formation of the backbone. Ionic, self-complementary peptides can be employed for the formation of nanofibers further assembling into macroscopic membranes.^{152,153} In general, this class of peptides exhibits an alternative sequence of polar and nonpolar residues (mostly alanine). Due to the stable β -strand that they form, the residues segregate in a nonpolar hydrophobic side and a hydrophilic side with the polar residues. The polar residues are again distinguished by their ionic character, either positive or negative, and are grouped in ionic areas so that they form complementary patterns (e.g. for EAK16, it is ++ – ++ –). Thus, perpendicular to the β -sheet formation, several β -sheets can now assemble by hydrophobic or ionic interaction of the residues to form extended fibrils and membranes.

Another approach is based on amphiphilic, surfactant-like peptides yielding tubular nanostructures. The amphiphilic character is achieved by a hydrophilic head with one or two charged amino acids and four or more consecutive non-polar amino acids to form the hydrophobic tail. Due to the enormous chemical space, the individual sequence of this class of molecules can vary. Examples for this kind of molecules are the peptide V_6D (V₆VVVD), consisting of 6 non-polar valine residues and a negatively charged aspartic acid residue in addition to the negatively charged C-terminal. Similar to lipids, they can form a bilayer and thus self-assemble into nano-structures.^{154,155}

So far, the formation of nanostructures by oligopeptides is mostly based on general interaction schemes (polar, nonpolar, charged, amphiphilic) that are defined by segments rather than on the specific interaction scheme of single residues. The influence of specific amino acids and their position is tested by a theoretical study that has mapped the chemical space of all tripeptides and could therefore established general rules for gelation of tripeptides.²⁷

Moreover, hydrophobic dipeptides have also been investigated for the self-assembly into nanostructures.¹⁵⁶ The most prominent example is the dipeptide diphenylalanine which readily forms nanotubes in solution¹⁵⁷ as well as by vapor deposition¹⁵⁸. The

formation is due to the $\pi - \pi$ stacking of the aromatic residues which is also thought to be responsible for the formation of amyloid fibrils being connected to the Alzheimer's disease.¹⁵⁹ These nanotubes exhibit outstanding rigidity (20 GPa Young's modulus) which is due to a zipper-like motif of the phenyl groups in the $\pi - \pi$ stacking¹⁶⁰.

So far, all these approaches (except the vapor deposition of Phe-Phe) were made in liquid environment. In this environment, a characterization of the formed structures is mostly performed by spectroscopy such as NMR, IR/Raman and UV/VIS. The peptide nanostructures can also be dried and then characterized in solid state such as by X-ray diffraction, solid state NMR, electron microscopy or AFM. These methods are suited for structural investigations, yet they average over a large ensemble and thus they provide only limited insight into individual molecules and heterogeneities.

Recently, STM was employed to investigate peptide nanostructures on the single molecule level. For very small peptides, such as the dipeptide diphenylalanine, adaptive binding and chiral recognition could be demonstrated to work on the surface.¹⁶¹ The differentiation of different residues by combining STM and molecular modeling has also been demonstrated.¹⁶² Investigating conformational heterogeneities, a flexibility in the conformation could be shown that might be relevant for cooperative fibril-like growth in amyloid fibrils.¹⁶³

While STM proves to be a very powerful tool to investigate peptide assemblies on the molecular and submolecular level, the accessibility of peptides, especially in the polypeptide regime, is very limited in the controlled environment of UHV. However, with the ES-IBD set-up,¹⁶⁴ this limitation can be overcome expanding the available peptide sequences up to protein complexes.^{70,90}

The capacity of STM/ESIBD to investigate proteins has been demonstrated presenting structural details of individual, large proteins of cytochrome C and BSA resolved by STM with sub-nanometer resolution. It could be shown that 2D refolding occurs on Au(111) for unfolded proteins due to the allowed diffusion on the surface.⁷¹ In contrast to the spontaneous refolding, the gas phase conformation of unfolded proteins cytochrome C steers the conformation when soft-landed on Cu(100). However, it could also been shown that there is no specific interaction due the different residues. The polypeptide rather behaves like an unspecific polymer and can be described by the worm-like chain model.⁷⁰

The influence of the chemical moieties can be enhanced by deposition of shorter sequences. Depositing bradykinin, a 9 amino acid peptide, we could show that folding is possible on the surface and can be steered by the diffusion regime on the sub-

strate.¹⁶⁵ This enables to observe a distribution of different conformations of single molecules when the diffusion is limited on Cu(110). Allowing the peptide to diffuse on the surface, the formation of stable dimers in only one specific conformation is found, hinting to an induced fit while folding on the surface.

Thus, for oligopeptides, the interplay between the chemical moiety as well as the interaction with the substrate influence the assembly. In this sequence range, we can investigate binding motifs that are based on the interplay of different residues and rationalize the influence of the sequence on the structure. To this end, we investigate oligopeptides by STM and elucidate the atomic structure by MD modeling for structure clarification.

3.2 Steric aspect to stabilize peptide tecton

As demonstrated by the example of bradykinin and cytochrome C, biomolecules try to form compact assemblies if allowed by diffusion. Due to the low dielectric parameter of the environment, the biomolecules behave like in a non-polar medium and tend to self-passivate. For flexible molecules and no diffusion limitation, self-passivation is possible, meaning that the biopolymer can fold and thus bury all polar groups in the inside. Another possibility to passivate on the surface is dimer formation. This happens, if diffusion is possible but self-passivation is limited. By choosing an amino acid sequence with bulky residues, one can suppress folding and consequently self-passivation, thereby enhancing intermolecular interactions.

Thus, to explore rational design principles for extended peptide nanostructures on surfaces, Angiotensin-I (AT-I, DRVQIHPFHL, $M = 1296$ g/mol) serves as starting point because its sequence is dominated by sterically demanding groups. Moreover, the sequence offers a variety of polar residues that promote non-covalent interactions, such as (Asp1, Arg2, Tyr4, His6,9) as well as non-polar residues for passivation (Val3, Ile5, Phe8, Leu10) (see Fig. 3.1a).

The mass spectrum of At-I is shown in Fig. 3.1d. The molecule is ionized by ESI from a water-ethanol 1:1 solution with approximately 0.1% formic acid at a final concentration of 10^{-4} M. To ensure high quality samples, one species is mass selected before deposition. As the 3+ charge state is most abundant, we selected this molecular ion. For deposition, the ions were decelerated to 10 eV to ensure soft landing. In this energy range, fragmentation of the molecular ions upon deposition can be neglected. During deposition, the sample was held at room temperature and the ion current (50-80 pA) was measured on the sample to estimate the coverage (60 pAh).

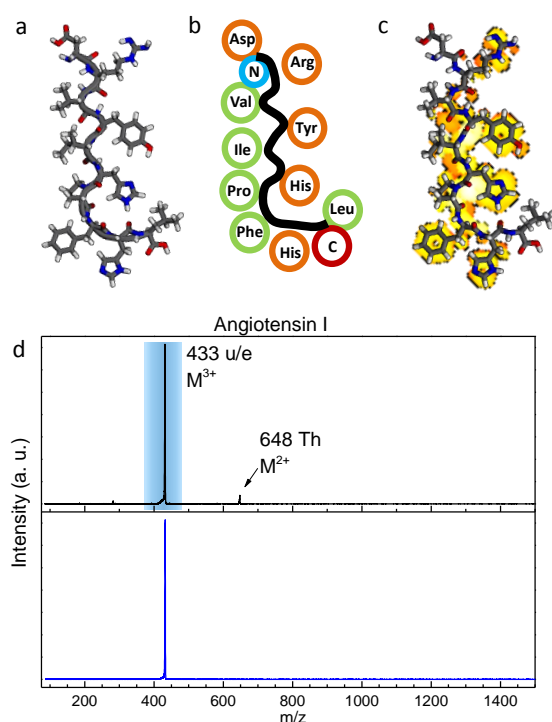


Figure 3.1: Angiotensin-I.

a: L-shaped MD optimized structure on the surface. **b:** Cartoon of the residue's distribution showing a segregation into a mostly polar (orange) and a nonpolar (green) side. **c:** Corresponding simulated STM image of the MD optimized structure. **d:** TOF mass spectrum of the full molecular beam and after mass selection (bottom).

Atomistic MD simulations of the landing process¹ of At-I on the Au(111) surface were performed for the protonated peptides being transported to the surface under the influence of an external electric field yielding a quasi-2D conformation.⁷⁰ For all MD simulations, the GROMACS software¹³⁷ was employed using the OPLS-AA force field¹⁶⁶. To take the interaction between the gold atoms and the peptide into account, the force field was extended by the GoIP force field.¹⁶⁷ The simulation yields a quasi-2D conformation in which the peptide backbone exhibits an L-shape (Fig. 3.1a, b). For clarification of the assembly motifs, this quasi-2D conformation was neutralized and employed as starting conformation. Similar to the configuration in a β -strand, the chemical moieties of the amino acids are located alternately on both sides of the backbone resulting in a near-complete segregation of polar (orange) and nonpolar (green) moieties, as illustrated in Fig. 3.1c. The residues are spread out on the substrate, thus the peptide lies flat to increase the area of interaction with the substrate.

After deposition on Au(111) and cooling to 40 K, we observe small island of two different types of assemblies by STM (see Fig. 3.2). Structure A is built by a linear chain,

¹performed by L. Harnau

while structure B is a compact assembly with rows separated by areas of different contrast. Moreover, single rotors can be observed on the elbow sites which are attributed to individual peptides that are still mobile to rotate but pinned at the elbow sites.

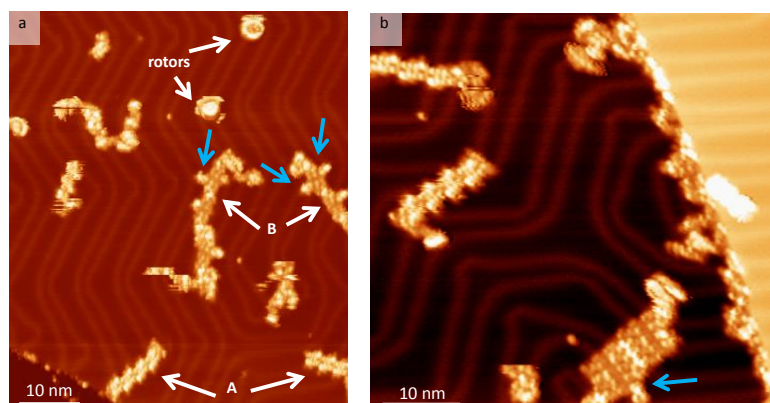


Figure 3.2: STM overview images of AT-I on Au(111). **a:** Assignment of the competing assembly motifs A and B. Rotors at the elbow site are also observable. **b:** Structure A and B on Au(111). In structure B, the different contrast between the rows is visible. The blue arrows point to adsorbed molecules on structure B.

In structure A, the repeating segment within the chain is 2.3 nm wide and 2.7 nm long and characterized by a double row of 4 features with 2 features between the rows. The observed C_2 symmetry of the segment suggests a dimer structure. Only by defect structures, highlighted in Fig. 3.3a, the assignment within the dimer is possible. Here, half of the dimer is observed which can be assigned to a single peptide. Thus, the peptides are aligned anti-parallel along their long axis in the dimer assembly, as indicated by the white arrows.

However, a precise assignment of the residues' position is not possible by STM alone. The information deduced by STM is used to validate the MD model shown in Fig. 3.3c. Furthermore, simulated STM images of the final MD structures without the substrate are calculated by DFT¹⁶⁸ to compare and confirm the MD optimized structures with the experimental STM data.

The simulated STM image reproduces the key features of the experimental data very well as shown in Fig. 3.3e: The rim of the assembly is pronounced due to the peptide backbone, whereas the inter-dimer region shows low intensity because of the low electron density of the C-terminal and the residues of Asp1, Arg2 and Leu10 located there. Moreover, in the inside of the dimer, the simulated STM image also shows pronounced features due to the aromatic moieties of Tyr4 and His6.

3.2. Steric aspect to stabilize peptide tecton

The atomistic MD model of structure A reveals that the nonpolar residues (Val3, Ile5, Pro7 and Phe8) lie on the outside of the assembly (see sketch Fig. 3.3c). Due to the L-shape of the backbone, the polar moieties of the peptides within the dimer cannot come into close proximity. Thus, the polar groups in the inside of the assembly (Tyr4 and His6) do not form hydrogen bonds but stabilize the dimer by weaker polar interactions. Hydrogen bonds are only formed between the C terminal and the Arg2 moiety. This formation also interacts with the His9 residue of the neighboring dimer, indicated by the yellow triangle in 3.3c. In addition, the His9 residues of two adjacent dimers connect diagonally, shown by the blue rectangle. This motif is encapsulated by the phenyl rings of Phe8, which passivates the polar groups from the outside. In combination, the self-passivation on the long side and the binding motif at the short ends lead to the formation of the staggered dimer chain.

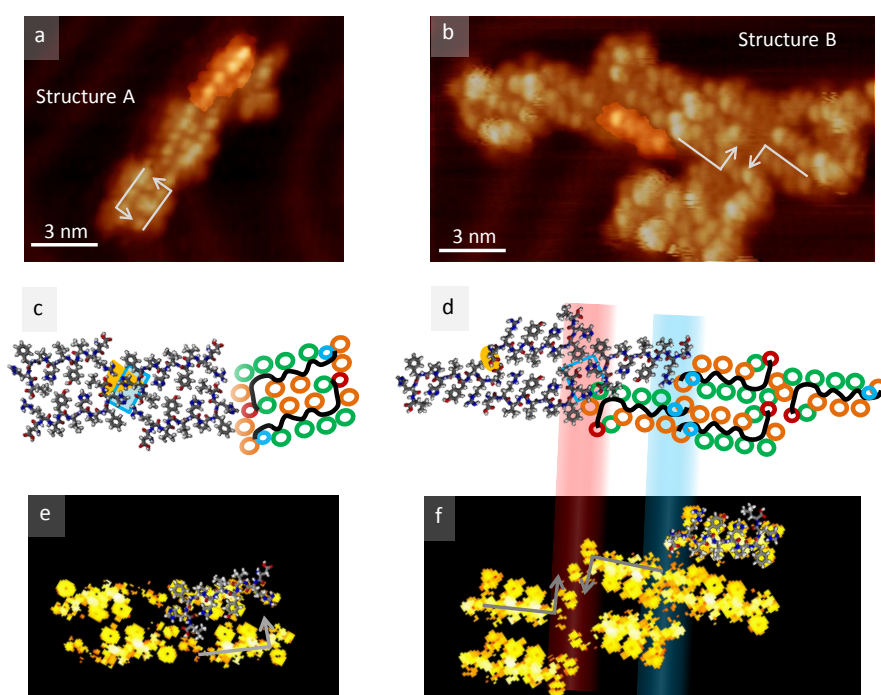


Figure 3.3: Detailed view of AT-I self-assemblies on Au(111). **a:** STM image of structure A. The upper dimer exhibits a defect and the single molecule is highlighted. **b:** STM image of structure B with the single molecule being highlighted. White arrows mark the orientation of the molecules. **c:** MD optimized atomistic model of structure A with important binding motifs (yellow and blue). **d:** MD optimized atomistic model of structure B with important binding motifs (yellow and blue). A cartoon on the right side sketches the motif. **e:** Simulated STM image of structure A. **f:** Simulated STM image of structure B. Red and blue background highlights the C- and N-terminal interfaces for the MD model and the simulated STM image.

In the more compact structure B, the assembly of the peptides is completely different, while the characteristic shape of the molecule can still be found at the edges of the assembly, as highlighted in Fig. 3.3b. From the symmetry in the assembly, we can deduce that the molecules stack in a parallel manner along the long side of the molecule resulting in rows of 3 nm width. The contrast between the rows alternates from very bright to low due to a rotation of adjacent rows by 180°. This means that the interaction at the interface of the rows is alternately via the C-terminal or the N-terminal.

These features are reproduced well in the MD optimized and stable structure shown in Fig. 3.3d. The molecules still have the same backbone geometry as for structure A while they are assembled in parallel stacks. Due to the rotation of adjacent rows, different interaction motifs are formed. At the N-terminal, two adjacent Asp1 residues form hydrogen bonds across the row structure (yellow oval in Fig. 3.3d), while on the other side, the C-terminal binds to the His9 residue of adjacent peptides (blue rectangle). The molecules within the stack can barely interact by strong hydrogen bonding as the polar side of one molecule faces the side with non-polar residues of the next molecule. Thus, the assembly is mostly stabilized by hydrogen bonding across the rows.

However, the interfaces between the rows cannot be assigned unequivocally by MD alone as this does not contain information about the intensity in the STM image, especially at the interfaces between the rows. An unambiguous assignment is given by the simulated STM image shown in Fig. 3.3f in which the alternating contrast at the C- and N-terminal interfaces is clearly reproduced. The interface at the C-terminal shows a wide region of low intensity. Nonetheless, the zipper-like motif within the interface is reproduced by the aromatic rings of His9 residue which carries some electron density. At the N-terminal interface, a row of high contrast is observed in the experimental and also in the simulated STM image due to the electron density at the N-terminal itself and the Asp1 residue.

Interestingly, while structure A seems to be completely passivated on the outside, attached molecules can be found for structure B. They are exclusively attached at the low contrast interface, indicated in the overview images in Fig. 3.2 by the blue arrows. With the introduced model we can deduce that this adsorption site is always at the C-terminal interface. Furthermore, for some assemblies they are observed on both sides, but tilted towards opposite stacks each. This observation fits well with the model as adjacent stacks have their polar, and thus reactive, side of the molecule on different sites.

For both structures, A and B, the assembly tries to passivate due to the non-polar side chains of the peptide which are segregated on one side of the backbone. Moreover, we observe that the conformations of the single molecule appear very similar in shape for both assemblies. The change in contrast might be due to slight variations in the conformation of single residues. This is confirmed by our simulation where the secondary structure of the peptide is the same and does not change with the type of assembly. However, due to the steric hindrance caused by the L-shaped backbone, the peptides can mostly interact via their C- and N- terminals, but only little interaction is possible between the side chains. This yields to the two competing motifs described in this section.

3.3 Sequence-controlled manipulation

As seen in earlier studies, the shape as well as the behavior of the peptide is controlled by the sequence of the amino acids. The manipulation of this sequence is easily accessible thanks to the sequential composition and synthesis of peptides.²² Based on the atomistic understanding of the nanostructures of AT-I on Au(111), we can now rationalize about how to manipulate the sequence to stabilize the assembly. While the segregation of polar and non-polar residues to the alternating order is definitely helpful for the stabilization of nanostructures on surfaces, the steric hindrance due to the bend in the backbone is less desirable. Therefore, keeping the sequence, but shorten it by 2 amino acids to avoid the bend in the backbone, results in a linear building block as pictured in Fig. 3.4a. The resulting sequence (DRVQIHPE, $M = 1046$ g/mol) is angiotensin II (AT-II), the biological successor of AT-I for the angiotensin-resin-circle and the biological more active hormone for blood pressure control.

With a linear backbone configuration, the residues of two peptides can come into closer proximity, resulting in a stabilization of the dimer motif. This is supported by MD simulations of the new sequence as shown in Fig. 3.4b. Due to the removal of the steric hindrance, the Arg2 group can now swing in, so that the residues Arg2, Tyr4 and His6 are aligned in a straight line and stabilize the dimer from the inside by 3 hydrogen bonds. The different residues have varying distances from the peptide backbone, so that Arg2 is the furthest away and His6 closest to the backbone. In the anti-parallel alignment of a dimer, this results in two peptides being wedged together in the way that the Arg2 is interacting with His6 of the other peptide and vice versa. While these optimized interactions stabilize the dimer at the inside, it is still passivated at the outside by the non-polar residues Val3, Ile5 and Pro7. Only the C-terminal and Asp1

are also located at the outside giving rise to possible points for interaction between dimers.

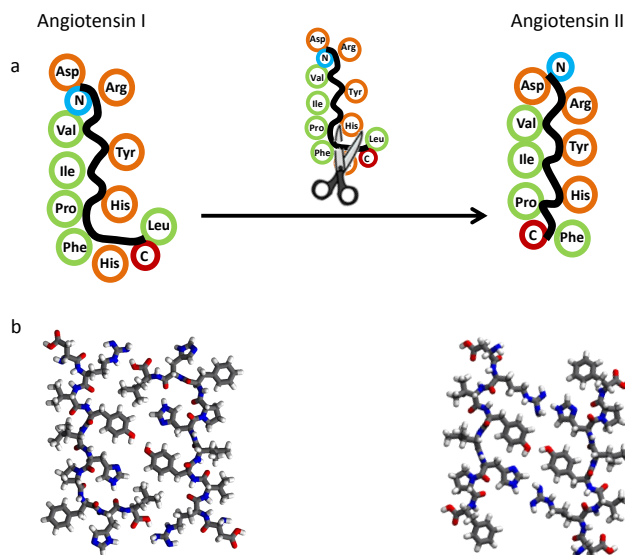


Figure 3.4: Manipulation of the sequence. **a:** Schematic of the removal of His9 and Leu10 at the C-terminal resulting in a linear backbone in AT-II on the surface. **b:** MD-optimized structures of the dimers on the surface.

3.4 Characterization of the honeycomb network

After deposition of the m/z selected At-II on Au(111) and cool-down to 40 K, we find porous honeycomb networks as shown in Fig. 3.5. It is apparent that the controlled manipulation of the building block's sequence changes the assembly behavior dramatically from small patches to long-ranged ordered porous assemblies. The overview image (Fig. 3.5a) shows that the islands grow on the terraces rather than from the step edges. On nearly every terrace, a honeycomb network can be found. Thanks to the high mobility at room temperature, the molecules can come together and crystallize in this regular honeycomb network upon decreasing the temperature. During cool down (approx. at 70 K), single molecules are first attached to the elbow sites of the Au(111) herringbone reconstruction as can be observed in Fig. 3.5b. At the elbow sites, the molecules are still mobile and are thus observed as rotors. These rotors might be the starting seed for the self-assembled network.

3.4. Characterization of the honeycomb network

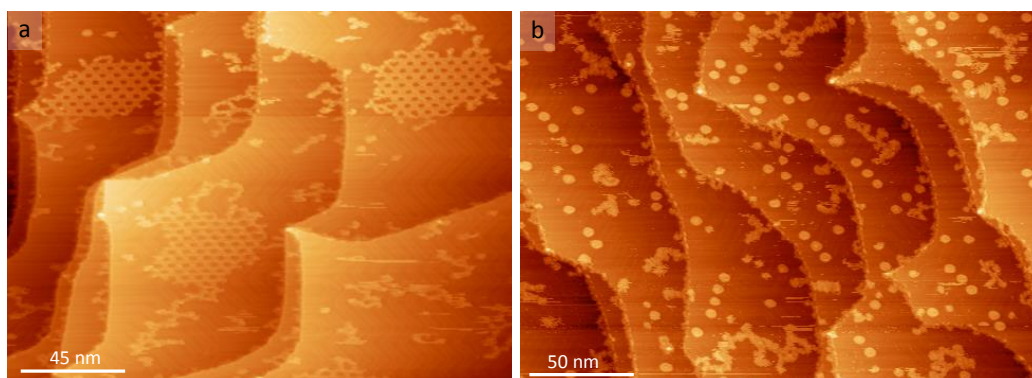


Figure 3.5: STM overview images of AT-II on Au(111). **a:** Honeycomb networks of AT-II form on nearly every terrace after cool-down to 40K. **b:** STM image during cool-down (approx. 70K). Single molecules are attached at the elbow sites of Au(111) and rotate.

Fig. 3.6a shows a single island of the honeycomb network. While there are defects and structural variations at the rim of the assembly, the assembly is completely ordered in the inside. The unit cell of this assembly accounts for $a = b = 5.5 \text{ nm}$ at an angle of 120° . The large hexagonal pores have an area of 4.5 nm^2 . The honeycomb assembly is chiral which is established in the anticlockwise tilt of the pores by 6° with respect to the networks lattice's principle direction, as indicated by the two hexagons in Fig. 3.6a. At the rim of the assembly, the double-walled nature of the network can be seen. However, it is very difficult to distinguish the assembly pattern due to the interwoven nature of the peptide assembly. Nonetheless, defects in the highly ordered assembly reveal the alignment of the assembled molecule as shown in Fig. 3.6b. Here, two different kinds of 0D defects can be observed. The white rectangle marks a defect, in which the dimer assembly is visible due to the enlarged distance between the two molecules of the dimer. By the shape and the distribution of intensity, this clearly shows that the dimer is made by an anti-parallel alignment of two molecules. Three dimers come together forming a vertex. The vertex arrangement of the dimers can be observed at the 1D defects marked with white circles. A tilt of one of the three dimers forming the vertex is responsible for this rotational defect and makes the complex interwoven pattern visible. These observations, including the formation of dimers as well as the dimensions, morphology and symmetry of the honeycomb network, serve as constraints for the structural clarification by MD and validates the model.

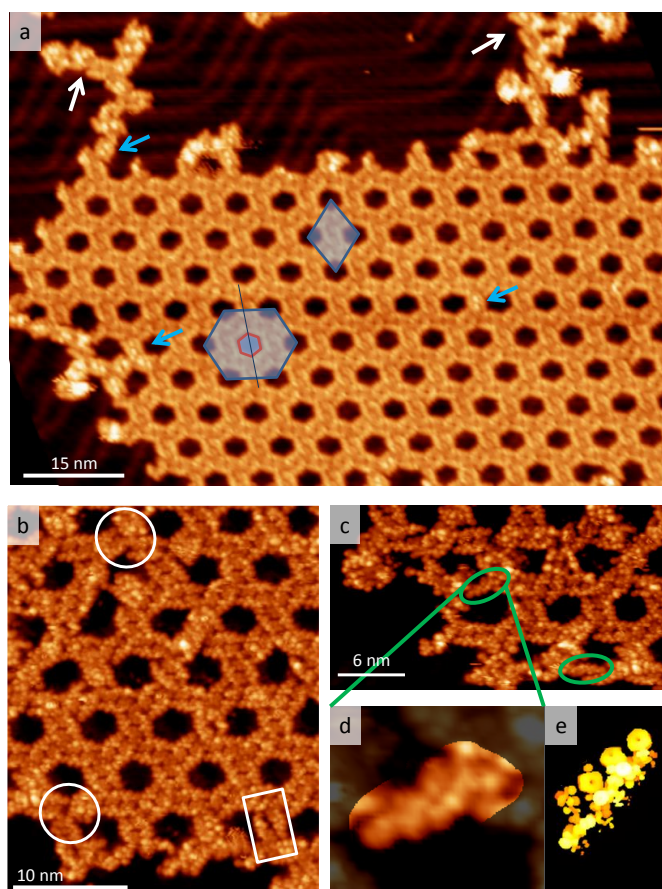


Figure 3.6: Unraveling the honeycomb network of AT-II on Au(111). **a:** STM image of the honeycomb network with the unit cell (blue diamond) and the highlighted chirality (blue and red hexagon). Conformational defects are marked by a blue arrow while linear assemblies of staggered lines are marked with a white arrow. **b:** Defect structures within the network to decipher the pattern of the assembly. **c:** STM images of single molecules at defects in the assembly. **d:** Magnification of a single peptide molecules in high resolution. **e:** Simulated STM image of a single AT-II molecule that reproduces the key features well.

Furthermore, even single molecules can be observed at defect sites within the network, as shown in Fig. 3.6c,d. With a high resolution (Fig. 3.6d), nearly every polar group as well as the backbone produces a feature in the STM. This agrees well with the simulated STM image in Fig 3.6e, where the backbone gives the highest signal due to the electron rich peptide bond and its height. The N-terminal (left) appears thinner as the Asp1 and Arg2 do not carry much electron density. In contrast, the residues close to the C-terminal are intense due to the electron rich groups His6 and Phe8.

3.5 Model of the honeycomb network

In the MD optimized model, the network is built up from the dimers shown in Fig. 3.4. While the dimer is passivated on the outside by non-polar groups, the short ends still have polar groups, such as Asp1 as well as the C- and N-terminal at the periphery. This enables the interaction with other dimers via polar interactions and hydrogen bonds between the C- and N-terminal of adjacent peptides at the vertices and results in the long ranged assembly scheme. Moreover, the three phenyl groups of Phe8

come together in the vertex. Although they are not in close proximity, they seem to have a strong stabilizing effect which can be seen in some defect structures. The interaction between the phenyl groups bears some resemblance to the zipper-like motif of the diphenylalanine nanonstructures. As highlighted in the cartoon in Fig. 3.7, this assembly yields a structure which is close to completely passivated at the pores and also exhibits the chirality by a tilt at the vertex.

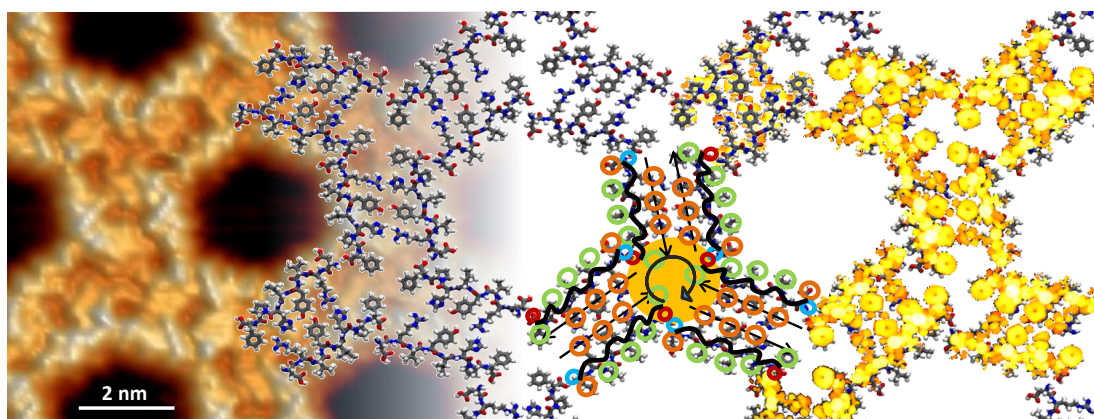


Figure 3.7: High-resolution STM image (1.3V, 35.5 pA; left) overlapped with the MD atomistic model (middle) and the calculated STM image (right).

Staggered linear arrangements of dimers are occasionally observed at the edges of the island, as indicated in Fig. 3.6a with a white arrow in the upper right corner. From the model, a polar - non-polar - polar - binding motif is formed at the ends of the dimer, involving the C-terminal (polar), the Phe8 residue (non-polar) and the Asp1' residue (polar). This motif promotes the staggered linear arrangement, which offer a binding site to add a third dimer and form the vertex. The Asp1 residue still remains present at the rim of the pore, however, the angled arrangement of the two adjacent dimers sterically hinders the approach of another peptide or dimer. Moreover, the discussed influence of the three phenyl groups in a zipper-like motif enhances the stability of the threefold assembly.

While the threefold symmetry of the vertex might be supported by the substrate symmetry, it also lowers the free energy of the system as Coulomb and van-der-Waals interactions have more impact due to the packing into hexamers. Thus, the vertex structure is more favorable for instance over a linear chain arrangement with the same amount of possible hydrogen bonds (see defects).

In addition to the agreement of the symmetry and dimensions with the experimental data, the simulated STM image of this structures reproduces the key features very

well: The high intensity ridge at the edge to the pores in the recorded STM image corresponds to the peptide backbone, while barely any signal is given by the non-polar groups at the periphery. Moreover, the distribution of intensity within the dimer is also in good agreement as the Tyr4 and His6 exhibit the highest signal of all residues in the simulated STM image. Most convincingly, the three phenyl groups in the vertex are rather isolated in the model but exhibit electron density and thus give rise to a signal in the simulated STM image. This fingerprint of the vertex is also observed in the measured STM image, where three distinct protrusions can be observed next to three features with very low intensity.

Taking into account all agreement of the model with the experimental data - symmetry and dimensions as well as features in the STM image - the introduced model is very reliable in the limit of the STM resolution. Moreover, while this model is established for the regular honeycomb network of AT-II, it is also able to account for all types of defects that we observe experimentally, as discussed in the next section.

3.6 Defects

Remarkably, most peptides in the assembly exhibit the same overall adsorption conformation within the network despite being flexible and multi-functional. However, there are some defects occurring that range from different conformations to different assembly motifs.

Conformational changes can be observed, e.g. by small variation of the contrast within the dimer motif as seen in Fig 3.6a (marked with blue arrows). Here, the defective dimer still exhibits the contrast at the rim to the pore, however the inner structure of the dimer is dominated by two bright spots. This demonstrates that the conformation of the peptide backbone is still comparable to the one introduced in the MD model, while the conformation of the residues Arg2, Tyr4 and His6 have changed for both peptides in the dimer. This change might be attributed to a rotation of the His6 ring or a misalignment of the amino acid pairing within the dimer. Nonetheless, as this defect is observed deep in the network as well as at the rim, this change in conformation does not seem to influence the overall assembly demonstrating its robustness.

Fig. 3.8a shows a linear grain boundary as a 1D defect between two domains that are rotated against each other by approximately 1° . This yields an increase in the width of the grain boundary by 1 \AA per hexagonal pore (from left to right). In order to account for the variation of the grain boundary's width, the dimers within the boundary adjust to the new situation. This can be seen in several changes: First, the single peptide

within the dimer is clearly distinguishable as they are partially pulled apart. Thus, one can clearly see that the peptide on the right side of the dimer is interacting with the upper domain and the left sided peptide with the lower domain, respectively. This agrees very well with the introduced model as the peptide on the right side is interconnected in the upper vertex by the phenyl residue (Phe8) and the C-terminal. Second, a variation in the rotation at the vertex motif and shearing within the dimer motif is observed to compensate for the compression/extension within the boundary, respectively. On the left side in the boundary, the dimer is rotated anti-clockwise using the vertex as a hinge to rotate and thus adjusting to the shorter distance between the two domains. On the right side of the boundary, the dimer is in its original rotation, but the two peptides are sheared against each other to make up for the longer distance to bridge the two domains. A combination of both adjustments is found in between. Last but not least, the contrast within the dimer changes from left to right indicating a conformational change of the residues. Taking into account the shearing and rotation of the dimer, a conformational change seems inevitable as the interaction pattern changes.

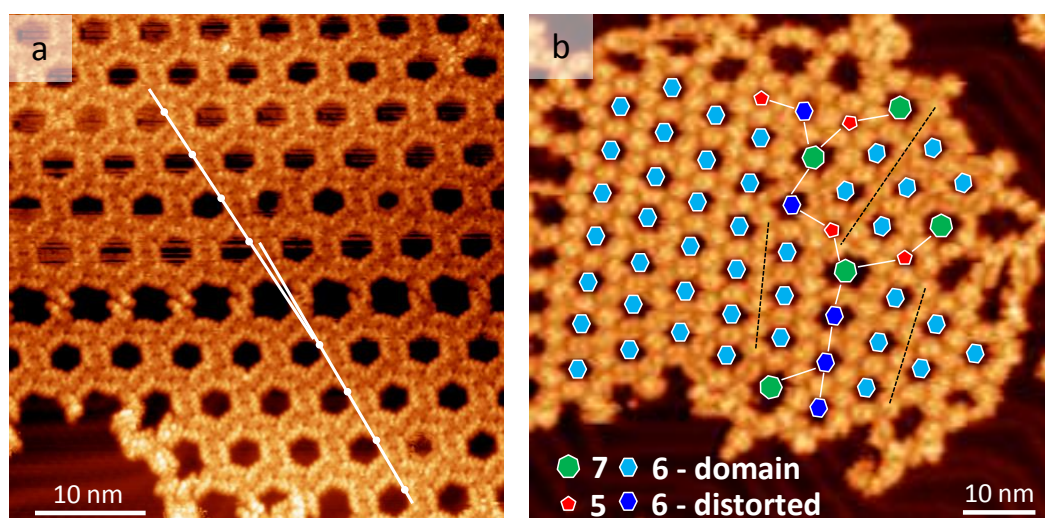


Figure 3.8: Defects in the honeycomb network. **a:** STM image (1.3V, 35.5 pA) of a linear grain boundary running horizontally in the central part of the image. The white lines emphasize the tilt of $\sim 1^\circ$ between the domains. **b:** STM image (2.4 V, 38.7 pA) of grain boundaries with large rotational offset ($\sim 16-25^\circ$). The resulting seven-membered (green) and five-membered (red) rings reduce the stress in the grain boundary.

In case of a more pronounced rotation between the domains, a new way of fault compensation takes place, as shown in Fig. 3.8b for a rotation of $16-25^\circ$ between the

domains. Here, the stress is reduced by forming a grain boundary of alternating seven (green hexagon)- and five (red hexagon)-membered pores. This scheme is fundamental and often found in systems of similar geometry, independent of the length scale, e.g., in graphene¹⁶⁹ or in curved colloidal 2D crystals¹⁷⁰. Unfortunately, the resolution in this measurement is not high enough to explore the possible conformational changes.

As seen for these defects, the peptide assembly is very adaptive and can adjust to different level of stress and mismatch. Interestingly, the peptide backbone seems to be rather undisturbed by the adjustments while the residues can change the inter-peptide binding motif and are thus the key for the adaptivity and fault-tolerance of the peptide nanostructure.

3.7 The role of the substrate

The rotational orientation of the AT-II domains appears to be unrestricted by the underlying Au(111) as many different orientations of the AT-II domains are observed that are not in agreement with a specific symmetry. This experimental finding suggests that the intermolecular peptide-peptide interaction dominates the substrate-peptide interaction. Thus, the Au(111) substrate does not impose a specific substrate-superstructure relation but rather provides 2D confinement. However, both interactions, the substrate-molecule as well as the intermolecular interaction, seem to be rather weak as the formation of an assembly happens only at decreased temperatures. For a computational estimate of the substrate-binding strength, most work in literature has been performed on individual amino acids.^{171–173} However, in this case of an oligopeptide, the peptide is rather rigid on the surface due to the restriction in 2D, the bulky residues and the strong peptide bonds. Thus, the binding of the peptide must be considered as a compound effect as the different affinities of the individual amino acids level out.

Interestingly, other STM studies of oligopeptides found - experimentally as well as computationally - adsorption of peptides on surfaces rotated by 90 °, meaning that the residues point either to or away from the surface rather than adsorbing flat on the surface.^{162,163,174} For most of the work, the difference might lie in the use of different computational approaches and in the sample preparation, as this was mostly done by drop cast on HOPG. In this way, the surface-peptide interaction is changed. Additionally, due to the drop cast, the coverage is high and the substrate-peptide

interface is not controlled. Moreover, the peptides that are mostly studied are related to amyloid fibril formation and thus are known to assemble in β -sheets. Nonetheless, this discrepancy demonstrates how little is understood for the adsorption of peptides on surfaces. Thus, high resolution imaging and other experimental approaches as well as the development of theoretical methods are definitely necessary.

Assembly of AT-II on Ag(111)

An experimental approach to investigate the influence of the substrate is to change the substrate. We have also investigated the assembly of AT-II on Ag(111) and Cu(111). Both of these surfaces do not reconstruct and are more reactive than Au(111) meaning that the interaction between the substrate and the peptide is enhanced. Additionally, on Ag(111) as well as on Cu(111), surface adatoms exist.¹⁷⁵ While they can be neglected for Ag(111) due to their low reactivity, they play a major role for adsorbed molecules on Cu(111).¹⁴¹ Thus, a detailed discussion about the assembly of AT-II on Cu(111) will be given in the next chapter presenting only small dimer and trimer nanostructures on Cu(111) probably due to the metal atom incorporation.

On Ag(111), the STM reveals a different behavior, as shown in Fig. 3.9 for the same sample preparation (clean metal surface, mass selected deposition of AT-II). Instead of the long-range ordered honeycomb network on Au(111), we find many small agglomerations as well as extended, anisotropic assemblies into strands. These assemblies are compact and have a periodic pattern that is seen in a regular wave form at the long edges. While the growth of the assembly is prominently in the direction of the strand, several strands can merge to form a wider strand as shown for a single strand of the width of ~3 nm (marked with A) to a double strand (marked with B) with a width of ~5.5 nm into a triple strand, marked with C and width of ~8.5 nm.

Most interestingly, the self-assembly does also not follow the surface symmetry of the substrate. This can be seen best in the lower left corner of Fig. 3.9. Here, the assembled strands are highly curved. Unfortunately, the low resolution of the images do not allow for a sophisticated analysis of the assembly. Nonetheless, this experiment shows impressively the influence that the substrate can exert.

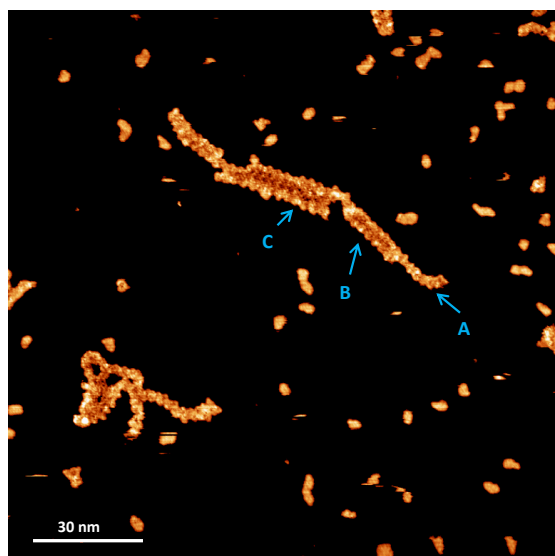


Figure 3.9: STM image of self-assembly of AT-II on Ag(111). In addition to small agglomerations, extended assemblies into strands can be observed. Several strands can come together from a single strand (marked with A) to a double strand (B) and a triple strand (C). The assemblies do not follow the surface symmetry.

3.8 Conclusion and Outlook

The sequence of a peptide programs its self-assembly and hence the expression of specific functionalities through non-covalent interactions, such as ionic interaction, hydrogen bonding, dispersive interaction and hydrophobic/hydrophilic segregation. While the formation of peptide nanostructures has already been explored extensively in the liquid phase, this chapter has provided insight into self-assembly of peptide nanostructures on surfaces in UHV, demonstrating that similar principles are at work on the surface in 2D.

In this chapter, the sequence controlled fabrication of molecular nanostructures has been presented employing peptides as bio-organic building blocks for 2D self-assembly. To this end, the natural peptides AT-I and AT-II were deposited by ES-IBD in UHV and subsequently characterized by STM. We have been able to obtain detailed models that allow for the identification of basic design principles and binding motifs by atomistic MD simulations being constraint by the submolecular STM images and validated by a comparison of the DFT simulated STM image to the experimental data. Based on this understanding, we have demonstrated the controlled change from competing motifs forming compact or linear assemblies (AT-I) to one dominant assembly scheme of long-range ordered, chiral honeycomb networks (AT-II) as a result of removal of steric hindrance by sequence modification. Furthermore, we have learned that using sterically demanding amino acids avoids backfolding and self-passivation of the single peptides on the surfaces. This results in a rigid building block of pre-

dictable behavior in which the residues are able to slightly change their conformation to achieve a high fault tolerance. In addition, a balanced distribution of polar and non-polar moieties is essential to provide specific binding sites, as well as passivation of the assembly, respectively, which is crucial for the long-range order assembly of oligopeptides on surfaces.

Interestingly, several approaches of the formation of peptide nanostructures in liquids are combined in one molecule for our design of the self-assembled long range ordered network: by the alternating sequence of polar and non-polar amino acids, a non-polar (hydrophobic) site and a polar site of the β -strand is formed, similar to the approach for the self-complementary peptides¹⁵². If one substitutes the charges in the self-complementary peptides by hydrogen bond acceptor and hydrogen bond donor, then the Arg2 (acceptor) and His6 (donor) are congruent with this approach. In addition, while there is no π -stacking, the zipper-like motif of the phenyl groups in the vertex resemble the stabilizing interaction of the diphenylalanine nanotube structure.¹⁶⁰

The detailed experimental mapping of peptides and peptide assemblies on surfaces can serve as a model system to validate the emerging theoretical approaches in this field.^{145,176} For modeling of large systems over long time periods, MD is so far the optimal tool. It applies force fields, which are very advanced for aqueous environments, but still lack the precision for metal-peptide interfaces. The quality of new force fields can be validated with this experimental approach. Eventually, a predictive modeling will be possible that enables the rational design of functional peptide nanostructures. Moreover, this new approach to 2D self-assembly on surfaces grants insight at the atomic level that will enable the use of oligo- and polypeptides as large, multi-functional bio-organic building blocks, and opens a new route towards rationally designed, bio-inspired surfaces. In order to further explore the possibilities, the honeycomb network of AT-II is suited perfectly as starting point. In this way, the proposed design principles can be further tested and improved, e.g. by slight changes in the sequence such as substitution of amino acids. Regarding the huge chemical space of the amino acids, a wide range of manipulations is possible. Manipulations in the sequence could be, for example:

- Substituting the polar amino acid Arg2 by the non-polar amino acid Leu which possesses roughly the same steric demand.
- Alterating the sequence by permuting the Arg2 and the His6 residues to investigate if the position is important.

- Capping the C-terminal (N-terminal) to explore the influence of the C and N-terminal interaction. (Is the dimer still formed, but no arrangement into the honeycomb network?)

Another way to explore the binding motifs and investigate the behavior of peptides on surfaces is to make stronger changes in the sequence under investigation. So far, the network only forms at decreased temperatures. However, for most applications, a self-assembly that is stable around room temperature is more suitable. By stronger interactions between the peptides, it might be possible to shift the crystallization temperature towards higher temperatures. Keeping the same binding motifs, this might be possible by doubling the sequence. In this way, there are not only three residues stabilizing the dimer, but six, as shown in Fig. 3.10. Assuming that the peptide backbone remains linear on the surface, this would form the same motif with increased pore size and higher stability at elevated temperatures. A higher stability could then also enable to functionalize the pore by functionalizing one of the non-polar amino acid pointing in the pore.

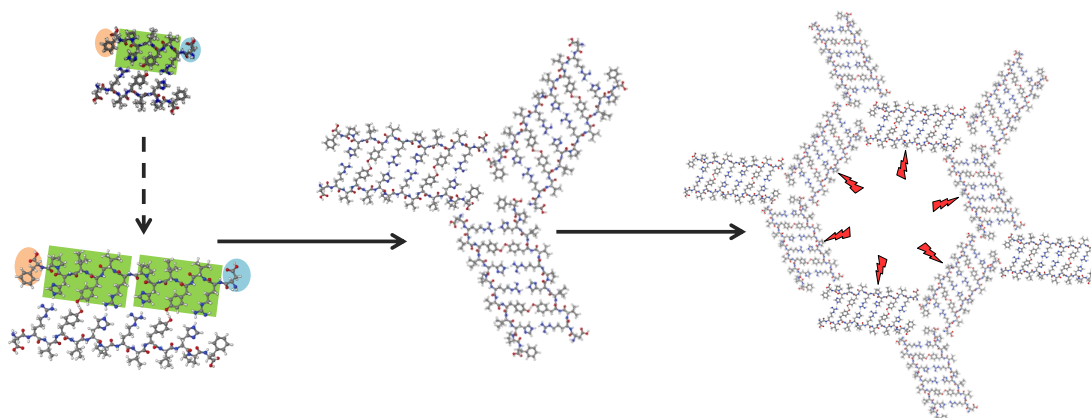


Figure 3.10: Atomistic model of a possible sequence that contains twice the dimer stabilizing motif of AT-II (green). Possible sites for functionalization of the pore are marked with a red flash.

4 Peptides: Metal-coordination to steer the conformation on the surface

Nearly half of all proteins are associated with metal coordination¹⁷⁷ and more than 30% of all proteins exploit metals to perform their specific function.¹⁷⁸ The metal ion can either be implemented in a prosthetic group, e.g. the heme where the metal is contained in a porphyrin, or directly employed as co-factor, e.g. in the zink finger. Owing to the rich variability of metals and proteins, protein-metal binding sites are highly variable and specialized to catalyze different biological processes, such as photosynthesis, signal transducing, molecular oxygen reduction or nitrogen fixation, and many more.¹⁷⁹ For this, the interplay between protein and metal is very important as it determines the electronic structure and redox behavior of the active site, as seen for chlorophyll or cytochrome C. Moreover, metal centers are also employed to steer the conformation¹⁸⁰ and change folding patterns, e.g. in the zink finger motif.¹⁸¹

While protein coordination to metal centers is essential, metal ions are also attributed to poisoning and diseases, such as Wilson's disease or Alzheimer's disease.¹⁸²⁻¹⁸⁴ Metal binding to a protein is suspected to enhance the aggregation and fibrillation. In Alzheimer's or Parkinson's disease, the formation and deposition of amyloid β -protein is strongly influence by metal ions, which is proposed to stabilize nonfibrillar pathogenic aggregate forms.

In STM studies, the coordination of small organic molecules to metal centers on the surface has been investigated extensively in the last decades.^{54,185} A huge variety of different metal organic coordination networks (MOCN) has been fabricated by self-assembly showing the tunability and versatility of this approach. In addition to the structural control of these self-assembled networks, the electronic structure of the metal center can also be addressed, resulting in, e.g. catalytically active nanostructures^{186,187} and controlled charge or spin states of the metal center¹⁸⁸. Employing

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

peptides instead of small rigid organic molecules could enlarge the diversity of the accessible structures and properties. For 3D metal-organic frameworks (MOF) which are built by the same principles as MOCNs, peptides have already been employed as organic building blocks, yielding new structures with interesting properties.¹⁸⁹ Moreover, the synthesis for small organic molecules is very specific, meaning that a new synthesis route has to be established in order to realize small changes in the building block for a MOCN or MOF, such as a different functional group or a different position of the functional group. In contrast to that, the sequence of peptides is linear and its synthesis follows a general scheme. In this way small changes (e.g substituting an amino acid) can be realized by only substituting a reactant in the whole sequential synthesis. This facilitates the accessibility of new building blocks while it also enables fine-tuning of the electronic structure.

In this chapter, I will explore the structural influences that metal coordination can have on peptide self-assembly on the surface. To this end, the self-assembled honeycomb network of AT-II, as introduced in Chapter 2, will serve as reference structure. The AT-II metal coordination complexes are formed in solution, transferred into the gas phase and subsequently deposited by ES-IBD. The self-assembled structures of coordination complexes of different metal ions are compared with the pure AT-II structure on the surface and thus the influence of the metal center can be explored.

4.1 The angiotensin-II-metal complex in the gas phase

The metal coordination to AT-II has been investigated extensively in the gas phase and in solution. Early on, it was found that alkali and earth alkali metal ions such as Na, Li, Ca and Mg increase the biological activity of AT-II in the angiotensin-resin-cycle.¹⁹⁰ In contrast, transition metal ions, such as Co, Cu and Ni, do not change or even decrease the activity. This behavior, being specific to the metal ions, suggests a conformational change of the peptide due to the position of the metal coordination.^{190,191}

The structural investigation of the coordination complex in the gas phase is mainly performed by tandem mass spectrometry. From the fragmentation pattern, the vicinity of the metal center can be inferred. However, between different experiments and groups, there is some discrepancy. Based on the proposed S-shape for the AT-II,¹⁹² DFT and tandem mass spectrometry suggested that Cu and Ni coordinate at the same site, while Mg and Ca coordinated to a different binding pocket.^{193,194} Both transition metals are claimed to coordinate with the oxygen atoms of the first four peptide bonds

4.1. The angiotensin-II-metal complex in the gas phase

(Asp1-Arg2-Val3-Tyr4) in addition to the nitrogen of N-terminal. In contrast, Mg was proposed to bind to the oxygen atoms at the C-terminal and of the last four peptide bonds (Ile5-His6-Pro7-Phe8).

Although earth alkali metal ions are known for their oxophilic behavior, transition metals normally coordinated strongly with nitrogen, such as the N-terminal or electron donating residues, such as the imidazole of His6. This behavior is also described in another tandem mass spectrometry study, in which different coordination sites for Cu and Ni are observed.⁶⁵ From the fragmentation pattern, it was concluded that Ni anchors and strongly binds to the imidazole-N of His6, while it simultaneously coordinates with the adjacent oxygen atoms of the peptide bond (Ile5-His6-Pro7). For Cu, the His6 residue is also an observed binding site, however, the carboxylate group of the C-terminal is preferred.¹⁹⁵

These converse findings, that also differ from proposed coordination structure in solution,^{196–198} demonstrate the manifold possibilities of the coordination environment for a metal center.

4.1.1 Ion beams of AT-II-metal complexes

Before depositing the AT-II metal complexes by ES-IBD, a routine for a stable and intense ion beam has to be established. Although ESI-MS has been demonstrated to be suitable for transferring AT-II metal complexes into the gas phase, we need to adjust the conditions to our machine and ensure an intense beam. Fig. 4.1 depicts the mass spectra of all AT-II metal ion beams before mass selection that we have established on our machine, namely the coordination complexes with Cu, Ni, Mg and Ca.

The coordination complexes are formed by addition of the respective metal salt to a solution of AT-II. As the coordination strength of the metal differs, as well as the sites to which the metal coordinates, the condition under which the complexes form vary drastically. This is represented in the optimized condition that was adjusted for each ion beam individually, as shown in Fig. 4.1 on the right side. While the Cu and Ca coordination complexes are formed with the AT-II solution in neutral pH, the Ni and Mg coordination complexes were only observable after raising the pH by addition of ammonia. For all complexes, we ensured to form only the 2+ coordination complex, which is attributed to the neutral peptide and the metal in oxidation state +II.

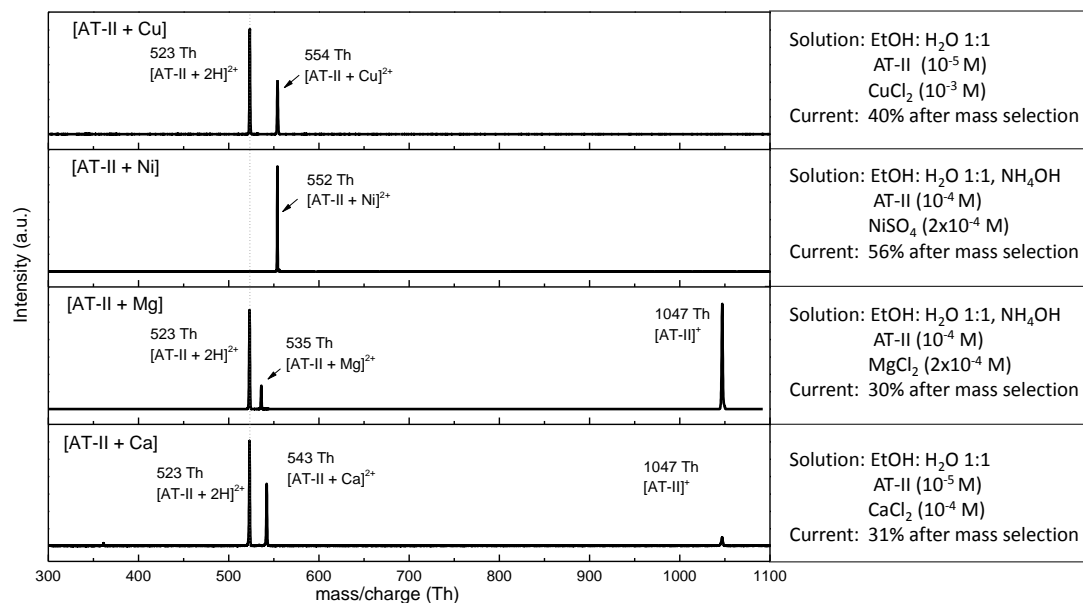


Figure 4.1: Mass spectra of metal peptide coordination complexes before mass selection. On the right side, the optimized conditions as well as the yield after mass selection is presented.

Under optimized conditions, the beams still contain the pure AT-II as main component of the ion beam, except for the Ni coordination complex which is amazingly pure and intense without optimization. All metal coordination complexes are sufficiently abundant, so that m/z selection is possible yielding an ion beam that is intense enough for deposition in UHV. However, due to the small difference in mass between the pure AT-II and the Mg or Ca coordinated metal complexes, respectively, the loss due to mass selection is larger. Thus, the $[AT-II + Cu/Ni]^{2+}$ coordination complexes are easier to operate and exhibit higher ion currents. Therefore, they were employed for deposition and the formed nanostructures were imaged by STM and are discussed in this chapter.

4.2 Peptide-metal coordination on the surface

This section explores nanostructures that are formed by deposition of $[AT-II + Cu/Ni]^{2+}$ coordination complexes on the surface. In order to see the influence of the metal coordination center, I will compare the observed structure with the nanostructures of the pure AT-II on different surfaces (Au(111), Ag(111), Cu(111)). On Au(111), I present the effect of different metal centers, as they prefer different coordination environments. On Cu(111), I discuss the formation of $[AT-II + Cu]$ coordination

complexes by the incorporation of surface adatoms. In order to ensure the intactness of the metal coordination complex, the species are soft-landed on the surface with 10 eV landing energy.

4.2.1 1D nanostructures by coordination to Ni and Cu on Au(111)

[AT-II + Cu] coordination complex

An STM overview image after m/z-selected deposition of the $[\text{AT-II} + \text{Cu}]^{2+}$ on Au(111) and subsequent cooling to 40 K is shown in Fig. 4.2. It is apparent that the deposition of the Cu complex yields a completely different assembly than the pure peptide AT-II (see chapter 2). In contrast to the 2D honeycomb network of the pure AT-II, the overview image shows short (max. 25 nm) 1D assemblies into zig-zag lines. Most of these assemblies follow the herring bone reconstruction within the hcp domain or are pinned to the elbow sites. In some cases, branching of the lines can be observed, as highlighted in Fig. 4.2b by a white circle. These branching points resemble the vertex structure of the AT-II honeycomb network (Fig. 3.7).

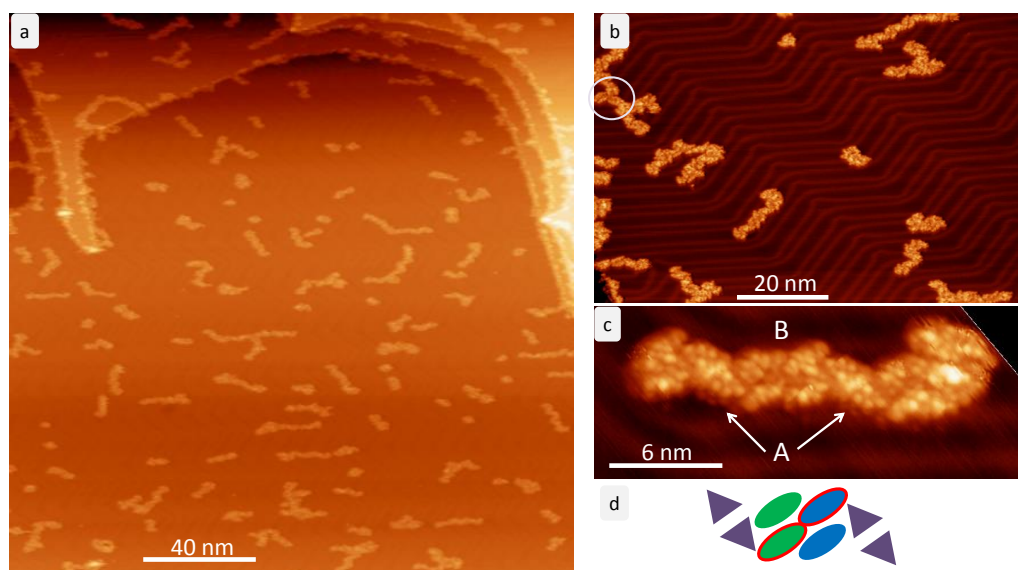


Figure 4.2: STM image of [AT-II + Cu] coordination complex on Au(111). **a:** Overview image showing short 1D lines and some alignment with the reconstruction. **b:** Branching in the lines is observed as highlighted with the white circle. **c:** The lines consist of two motifs labeled A and B. **d:** Simplified cartoon of the motif A and B together.

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

The reoccurring structure is shown in a magnification in Fig. 4.2c and depicted in a cartoon in Fig. 4.2d. The most striking motif of this structure is labeled with **A** and sketched by two violet triangles. Due to its size and the C_2 rotational symmetry, it is assigned to a dimer structure. It consists of a diagonal feature in the middle, and four smaller features on each site, resembling the footprint of a paw. The dimer motif is not completely passivated as it never terminates the assembled lines, but always has another agglomeration or structural motif next to it.

In most cases, motif **B**, a staggered chain of two dimers (sketched in green and blue ovals in Fig. 4.2d, respectively) which resemble the dimers of the pure AT-II structure, connects two of the motifs **A**. For this, the dimers are sheared and can thus partially envelop the non-passivated site of motif **A**. This also accounts for the zig-zag structure due to the angle at which both motifs come together. Due to the staggering of the dimers in motif **B**, two of the peptides can connect diagonally (red outline of the ovals), while the other two molecules in this motif are squeezed.

After annealing to 360 K and subsequent cool-down, the self-assembled structures have changed slightly. As seen in the overview image (Fig. 4.3), the branching has mostly disappeared and the zig-zag lines exclusively align with the hcp domain of the herringbone reconstruction. In addition to the zig-zag lines, straight lines of dimers are formed that exhibit two bright features where the dimers meet.

Within the zig-zag lines, the structure has slightly changed, as can be seen in Fig. 4.3b and the respective cartoons in Fig. 4.3b. While motif **A** remained the same, the conformation within the two staggered dimers of motif **B** changed resulting in two bright features where the two dimers meet. Here, the assembly has changed in the ways that the outstretched molecules that were connected before (red outline in the sketch) moved out and the squeezed molecules changed the conformation yielding the bright lobes (black dots in the cartoon) and coming closer in contact. Due to this conformational change, the pawn print motif of motif **A** can now also be observed in the staggered dimers, as marked by blue dots in Fig. 4.3b. Interestingly, some single molecule in motif **B** can be found that resemble the uncoordinated At-II molecule in the network (highlighted by a light white circle). By this resemblance, an orientation of the molecule and the identification of the C- and N-terminal can be proposed.

Moreover, dimers of the squeezed peptides are found to assemble into linear lines with the same bright features. As seen in the ending points of these lines, the bright features only occur where dimers meet, suggesting that the conformational change is due to interaction with other molecules.

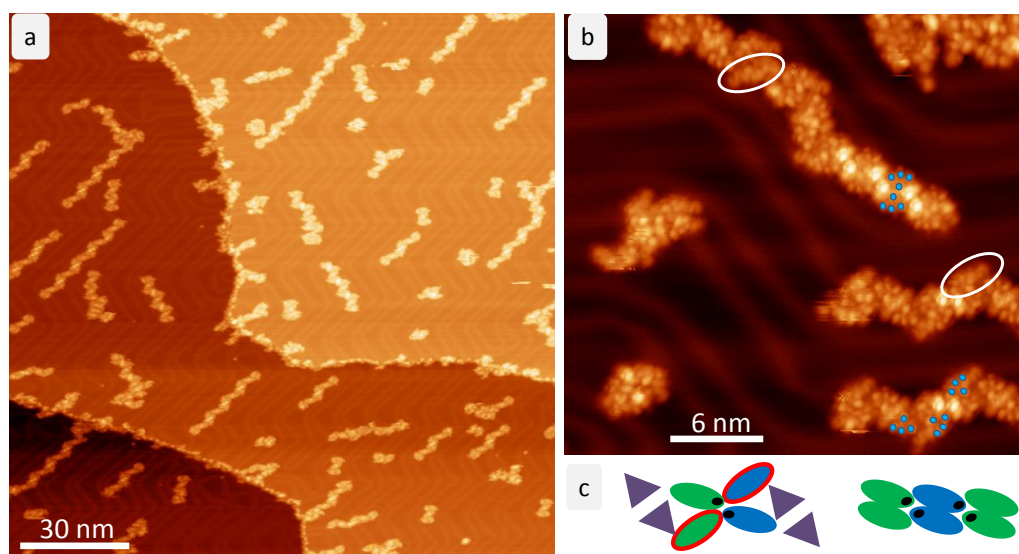


Figure 4.3: STM image of [AT-II + Cu] coordination complex on Au(111) after annealing to 360 K. **a:** Overview image shows the increased alignment of the zig-zag lines with the herringbone reconstruction. **b:** Detailed image of the motifs. Motif A remained the same, while motif B changed, showing now the same paw-print pattern (blue dots) and bright features. Some molecules that resemble the pure AT-II are found (light green circle). Additionally, a new structure is observed, formed by dimers with bright features at the side. **c:** Simplified cartoon of the A+B motif and the new straight dimer line.

Model of the [AT-II + Cu] coordination complex on Au(111)

For a model, the described experimental observations, such as symmetry, defects and change after annealing have to be taken into account as well as the potential coordination sites for the metal center. With all this, a model is developed based on the MD optimized structure of AT-II on Au(111) (see chapter 2). To this end, the structure is manually adjusted for metal coordination and subsequently relaxed by the 3D geometry optimization using the FFU force field implemented in Avogadro¹⁹⁹ to avoid strained structures and chemical or physical inconsistencies. Although the molecule is very complex, this approach is sufficient to rationalize the binding motifs of the structures. A further MD optimization including the substrate interaction will change the conformation of the peptide slightly, however, the binding motifs will most likely remain.

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

Compared to the pure AT-II on Au(111), the single molecule (white circle in Fig. 4.3b) exhibits a higher curvature and has changed mostly on one side: The "thin" end (compare to Fig. 3.6d) at the N-terminal with the Asp1, Arg2 and Val3 is very similar, while the C-terminal side is wider for the [AT-II + Cu] coordination complex. This suggests that the Cu coordination takes place close to the C-terminal, which is also in agreement with one gas phase structure.^{65,195} Including a Cu coordination to the C-terminal and the backbone oxygens of Pro7, His6, Ile5, yet keeping the conformation of the N-terminal site the same than for the pure AT-II, the model has been developed which is pictured in Fig. 4.4a. Due to the coordination, the peptide is not completely extended, and the segregation of polar (orange) and non-polar (green) groups is now incomplete as pictured in the cartoon Fig. 4.4b. With this model, all motifs, which have been observed by STM, can be explained as presented in the following.

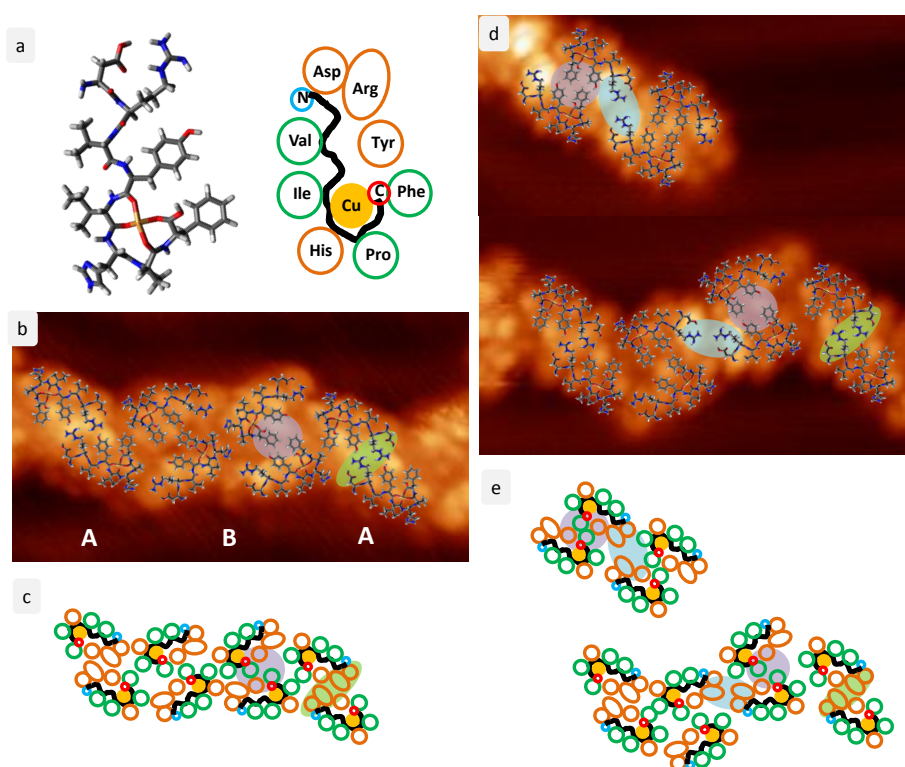


Figure 4.4: Model of [AT-II + Cu] coordination complex on Au(111). **a:** Atomistic model and cartoon showing the Cu center in yellow, polar groups in orange and non-polar groups in green. **b:** Overlay of the proposed model with the STM image before annealing, rationalizing motif A and B and highlighting important motifs. **c:** Cartoon of the atomic model before annealing. **d:** Overlay of the atomic model with a STM image after annealing, rationalizing the change in motif B and highlighting important binding motifs. **e:** Cartoon of the atomic model for the structures after annealing.

4.2. Peptide-metal coordination on the surface

An overlay of this model structure on the STM images (see Fig. 4.4b and d) demonstrates that the model has reasonable interactions sites and is able to rationalize all structures observed with the same model conformation. In motif **A**, two peptides come together in the way that an Arg2-Tyr4' interaction between the peptides is established, and vice versa. The paw-like features are formed by the backbone and the His6, Ile5 and Val3, as well as the C-terminal and the Phe8 residue. The bright diagonal line corresponds to the two Arg2 residues which are almost parallel (highlighted green in Fig. 4.4b,c). In this dimer formation, the non-polar residues Ile5, Val3 form the outside while motif **B** interacts with the His6 and the Asp1 residue, thus, establishing the angle for the zig-zag formation.

In motif **B**, the two dimers can interact with each other via hydrogen bonding between Arg2 and His6', a motif that is already known from the honeycomb network of AT-II, however in a slightly other configuration. Within the dimer, the interaction is mostly vdW interaction between the aromatic residues Phe8 and Tyr4 (highlighted violet), rationalizing the facile shearing of the dimers that can be observed comparing before and after annealing.

After annealing, this motif changes, resulting in a structure with two bright features in the middle, while simultaneously the formation of dimers with two bright features at the sides is observed. In the model, the bright features are attributed to an interaction between two Arg2 residues (highlighted blue in Fig. 4.4d,e). As seen for the compact structure of AT-I (chapter 2 Fig. 3.3b,d,f) where the Arg2 residue contributes to the high intensity interface, Arg-Arg interactions often result in bright features in STM. For this interaction to happen, the diagonally connecting sites between two dimers need to change. This might be initiated by the annealing yielding a further deprotonation (e.g. of the Asp1 residue) and thus a change in interactions strength for some motifs. The same motif is found for the newly establish dimers, where the dimer interactions is by aromatic residues in the inside (highlighted violet) and hydrogen bonding between the dimers via the Arg2-Arg2' interactions that is responsible for the bright features. This motif also explains why no bright features are observed at the end of the dimer line: As no second Arg2' is available at the end, there is no structural motif that gives rise to the bright feature.

[AT-II + Ni] coordination complex

Similar to the [AT-II + Cu] coordination complex, the [AT-II + Ni] coordination complex yields 1D assemblies on Au(111) at 5 K, as shown in Fig. 4.5. The STM overview image

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

in Fig. 4.5a demonstrates that the assemblies remain 1D, even at high coverage and after annealing. As the structures are anchored at the elbow sites and exhibit a variety of different growth directions, linear as well as curved assemblies are observed. Even after annealing, the alignment with the surface does not increase.

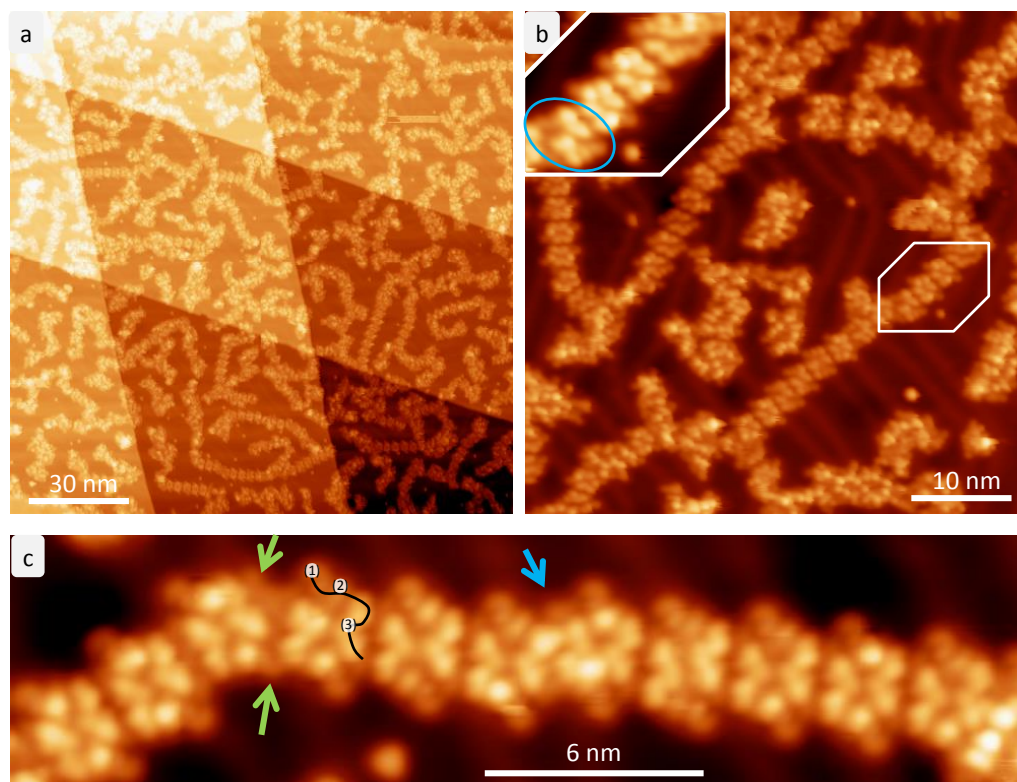


Figure 4.5: STM images of [AT-II + Ni] coordination complex on Au(111). **a:** Overview image showing the long lines can be bent. **b:** Lines of dimers. The inset shows a shearing of one dimer. **c:** Magnification of a linear assembly. Green and blue arrows mark the defect structures due to the bent in the line.

The inner structure of the 1D assemblies (see Fig 4.5b and c) is periodic and nearly rectangular building blocks (2 nm width, 3.9 nm length from tip to tip) can be observed. These building blocks exhibit a C_2 rotational axis suggesting that they are formed by dimers. In general, these dimers are aligned in a parallel manner with their long edge interacting with each other. However, in the bent lines, the dimers are misaligned and rotated against each other. This misalignment results in wider gaps between the dimers (green arrows), as well as defect structures merging two dimers (blue arrow) demonstrating that the interaction within the dimer is stronger than between the dimers. From the binding motif of bradykinin¹⁶⁵ or the pure AT-II¹⁴⁶ (chapter 3),

we know that hydrogen bonding and the interaction of polar residues takes mostly place within the dimer, while the dimer is passivated against further interaction to the outside. As this is suggested to be a general scheme that is also observed for other amphiphilic species such as SDS,¹⁰⁵ it should also apply here and could explain why the misplacement happens preferably between the dimer structures and not within a dimer.

The substructure of the dimer is governed by a diagonal of low intensity going from the top left corner of the dimer to the bottom right, dividing the dimer into two triangular structures. In defect structures (see blue circled structure in left side of inset in Fig. 4.5b) the triangular building blocks are sheared against this line. This suggests that there is no covalent bond across this line, thus separating the single molecules within the dimer. The single molecule has therefore a roughly triangular shape in which a bent line of different features can be seen (line with numbers in Fig. 4.5c). At the very start of this line, a low intensity protrusion (1) is observed that goes further out of the rectangular dimer. In comparison with other peptides on surfaces, such as bradykinin or the pure AT-II, this protrusion is characteristic for the flat lying phenyl residue of Phe8, suggesting that this feature is located in proximity to the C-terminal. Attached to this protrusion is feature (2). This feature is bright in most cases, but can also possess the lowest intensity in other dimers, hinting to a feasible change at this site. The next features (3) always has high intensity, but cannot be further distinguished and thus resemble a V-shape until the end of the single molecule.

Model of the [AT-II + Ni] coordination complex on Au(111)

As there is barely any resemblance in the STM images between the Ni coordinated peptide and the pure AT-II, the model has to be explored based on gas phase structure from literature and subsequently validated by the STM images. A coordination to Ni by the His6 residue and the backbone oxygens of Asp1, Arg2 and Val3 yields a triangular building block. At the periphery, all non-polar (Val3, Ile5, Pro7) groups are passivating the building block on the outside. Most polar groups are in the inside, such as the His6 that interacts further with the Asp1 residue and the C-terminal. This leaves one side that is suitable for interactions, presenting the Arg2, residue as well as the joint interaction site of the Asp1 and C-terminal. Moreover, the Tyr4 is opposite to this reactive side and also suitable for hydrogen bonding interaction.

Overlaying the model with the STM image, as presented in Fig. 4.6b, it is apparent that the overall shape of the model agrees with the experimental data, but some dimen-

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

sions diverge, such as the width of the dimer or the distances between the molecules within the dimer. This divergence is a consequence of the manual adjustments and the 3D geometry optimization. Taking into account the flattening effect of the surface, conformational changes will certainly happen that change the dimensions, such as the adjustment of the backbone or tilting and flatten of the His6 residue. One can assume that this will stretch the molecule along the Arg2-Phe8 directions resulting in a better agreement with the dimensions of the STM image.

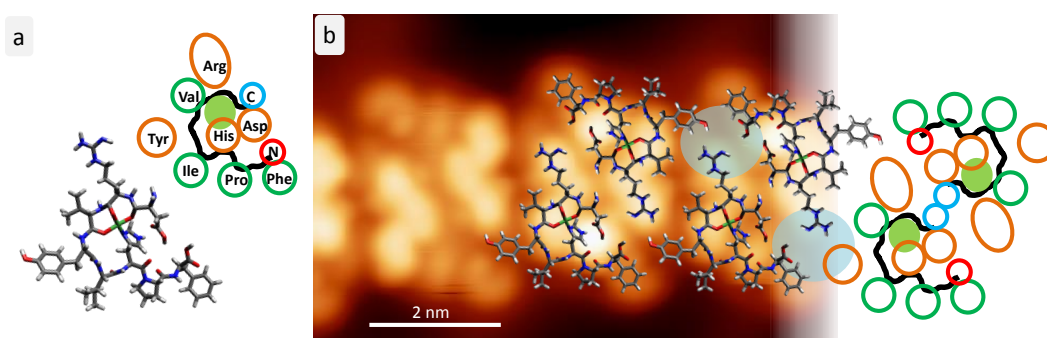


Figure 4.6: Model of [AT-II + Ni] coordination complex on Au(111). **a:** Atomistic model showing the Ni center in green. **b:** Detailed STM image of dimers. On the right an overlay of the model with the STM image demonstrates the fit of the model.

Bringing two of these triangular structure together, a rectangular dimer can be formed that is interacting with the polar side. The Arg2 of one peptide can interact via hydrogen bonding with the C-terminal and the Asp1' of the other peptide (see blue highlighting circle). The straight polar side going diagonal through the dimer also enables shearing of the dimers along this line as observed in some defect structures. Moreover the low intensity protrusion described in the STM image fits in position very well to the Phe8 residue.

The formed dimer is nearly completely passivated on the outside (see sketch), except for one part of the Arg2 residue and the Tyr4 residue. Lining the dimers up next to each other, these two residues can interact and therefore stabilize the inter-dimer interaction. However, this interaction is probably not so strong, so that the dimers can be bent apart or squeezed together to release strain, as seen in the bent lines in Fig. 4.5.

4.2.2 Comparison of structural motifs between the pure peptide and the metal coordination complexes on Au(111)

In order to evaluate the influence of the metal center, a direct comparison between the nanostructures is drawn. Fig. 4.7 presents the cartoon of the monomers for the pure AT-II, the [AT-II + Cu] and the [AT-II + Ni] coordination complex and the atomistic dimer models that were described in the previous sections.

The influence is already seen for the conformation of the single peptide: In contrast to the pure peptide exhibiting an almost linear backbone, the backbone of the metal coordination complexes is curved. For [AT-II + Cu], the curvature only occurs at the C-terminal side, where it coordinates to the Cu, and thus yields still an overall rectangular shape. The strongest curvature is observed for [AT-II + Ni], for which the backbone describes a half circle resulting in a triangular shaped structure on the surface. This strong curvature is enforced by the coordination of the Ni to the backbone at N-terminal side and the His6 residue in the middle of the sequence.

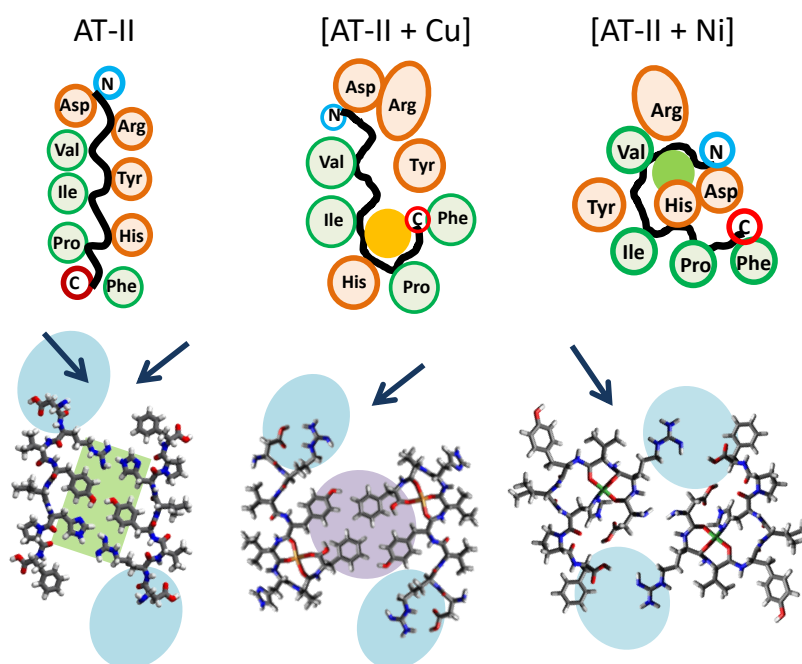


Figure 4.7: Comparison of the different structural motifs on Au(111).

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

The conformation of the backbone influences the conformation of the residues and thus enables different dimer motifs. For the pure AT-II, an almost complete segregation in polar and non-polar residues is achieved. A resulting dimer is stabilized by the hydrogen bonds and interactions of polar groups in the inside (highlighted green). Additionally the interaction between dimers is enabled at the short side of the dimer at the Asp1 residue (highlighted blue) and the C-terminal. Due to the interaction with two dimers on each side (blue arrows), a 2D assembly is established (see chapter 2). The dimer of the [AT-II + Cu] complex looks very similar, but has some distinct differences. Due to the bend in the backbone, the segregation between polar and non-polar groups is incomplete and the dimer is not stabilized by hydrogen bonding or polar interaction, but by the interaction between the Tyr and Phe residues (highlighted in violet). The burying of the C-terminal in the inside of the peptide structure decreases the number of strong interaction sites between dimers. Thus, a dimer only interacts with one other dimer on each site (blue arrow) via the polar residues of Asp1 and Arg2 (highlighted blue), therefore a 1D assembly is established.

In the triangular structure of the [AT-II + Cu] complex, one side is passivated by the residues of Ile5, Pro7 and Phe8. Another side has mostly polar groups, such as the C- and N- terminal, the Asp1 and the Arg2 residue. The dimer formation takes place via this side (highlighted in blue), which leaves only the polar Tyr4 and some of the Arg2 as interaction sites for further dimer-dimer interaction and thus, also yields 1D assemblies.

As seen in this comparison, the conformation of the backbone is pivotal as it determines also the conformation of the residues and thus the possible interaction sites. The general scheme, that the assembly is driven by passivation to the outside, also holds true for all the observed nanostructures of the metal-peptide coordination complexes. While the sequence exhibits bulky residues and is thus rather rigid forming a linear building block for self-assembly, it can be changed drastically by the selective implementation of metal coordination centers.

4.2.3 Close packed nanostructures on Ag(111)

As shown in Fig. 4.8a, the deposition of the AT-II Cu coordination complex on Ag(111) also yields highly anisotropic strand-like assemblies, very similar to the one obtained on Ag(111) without metal coordination. These strands possess bright dots on both sides and do not follow any surface symmetry. Interestingly, in contrast to the pure AT-II on Ag(111), barely any small features are observed that were related to single

peptides.

After annealing to 365 K, the assembly has changed forming large 2D islands of merged long strands (see Fig. 4.8b). While the edges of this assembly are straight and sharp, the assembly is not in registry with the surface, as different angles are observed. The magnification of the dense island shows a regular structure within the strands (see Fig 4.8c and d). Building blocks of a length of 3.5 nm and a width of 1.2 nm are stacked parallel, tilted by 145° with respect to the strand direction. This gives an impression of the strands being braided. The building block exhibits a C_2 symmetry, suggesting that it is actually formed out of two single peptides (highlighted with blue and red dots in Fig. 4.8c) interacting at the ends rather than the side. Depending on the functionalization of the tip, the contrast changes, giving rise to bright features within the strands (Fig. 4.8d). This might hint to the position of the metal atom. Unfortunately, no comparison can be made to the uncoordinated AT-II structures on Ag(111), as no annealing was performed on these samples.

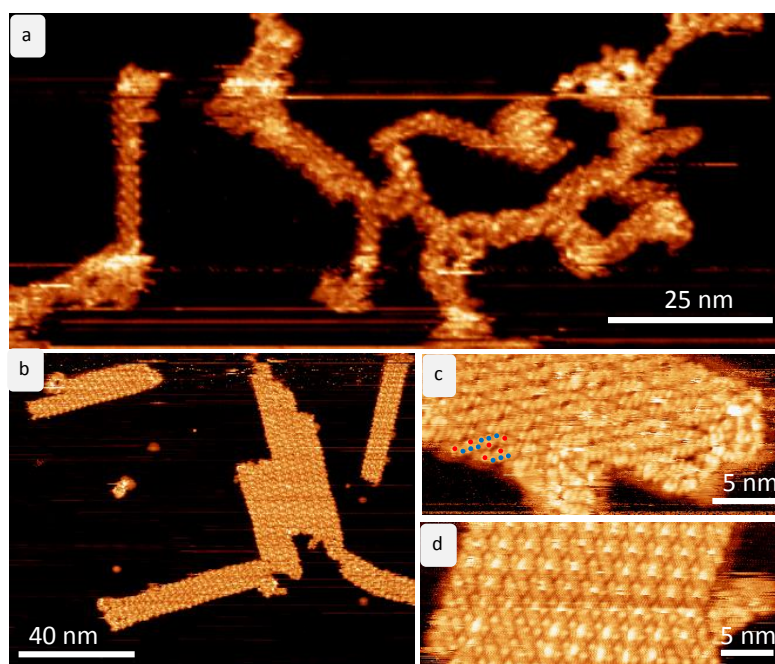


Figure 4.8: STM image of [AT-II + Cu] coordination complex on Ag(111). **a:** Non-regular strands are formed before annealing. **b:** After annealing to 360 K, 2D islands are formed by the strands attaching together. **c:** Detailed STM image of 2D island. Blue and red dots highlight the building block and the single peptide within. **d:** Detailed STM image of 2D island with a different contrast.

4.2.4 Isolated nanostructures of [AT-II + Cu] on Cu(111)

On Cu(111), rectangular dimer and triangular trimer structures (labeled T) are observed on the surface after deposition of the pure peptide (see Fig. 4.9a), aligning with the threefold symmetry of the surface. These structures are stable at room temperature, however, to get high resolution STM images, the samples are cooled down to 40 K. In the dimer, the corners of the rectangle are alternating bright and dim with several features within. Only by defect structures (highlighted with a white circle in Fig. 4.9a) is it possible to deduce the single molecule in the dimer revealing a chiral elongated building block that has two high intensity features at the bright rectangle and an area of less intensity on the other. The part that is outside when forming the dimer exhibits more intensity than the inner part, suggesting that this might be due to the backbone and the non-polar residues. At high coverages, 1D lines are formed that have the same width than the dimers and seemed to be formed by merging the dimer. The triangular structure T has round edges and exhibits a cavity at the middle while the three corners have the highest intensity. The substructure within this assembly suggests that it is formed by the same building blocks as the dimer. Interestingly, the orientation of the monomer with respect to the surface changes from the assembly in the dimer to the assembly in the trimer. Very rarely, other types of trimer structure can be found on the surface (marked with blue/violet circles). The transformation from the originally described trimer T to the other triangular and thermodynamically more favored structures can be enhanced by annealing, demonstrating that the described trimer is only a local minimum and kinetically trapped.

After annealing to 360 K and subsequent He cooling to 40 K, the dimer structure is completely vanished, as shown in Fig. 4.9b. Instead three different triangular assemblies have been formed that have been rarely observed before annealing and differ from the originally observed structure T: a regular triangular shaped (A), a truncated triangular shaped (B), and a wind mill type (C). All of these trimer structures possess some common features such as a similar area and the C_3 symmetry axis which is identically aligned for all three structures. Additionally, a cavity in the very center can be observed for the truncated structure B which is slightly visible in the other two structures, too. Structure A and C also agree in three small, distinctly bright features decorating the cavity in the middle. This feature is point-like for structure A, but elongated for structure C.

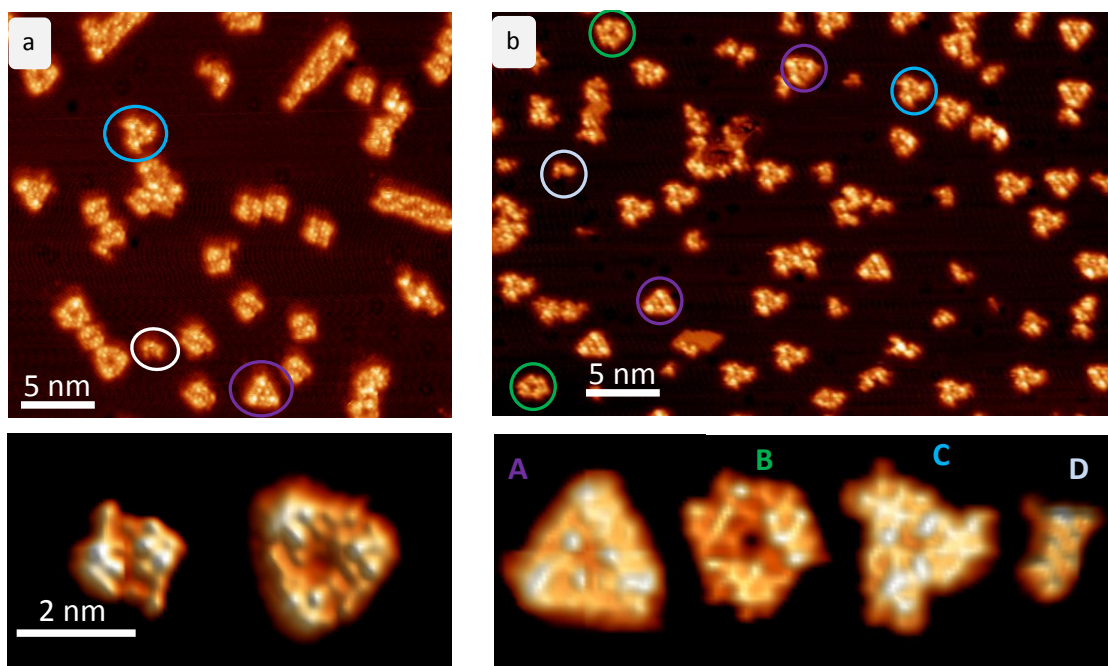


Figure 4.9: STM images of AT-II on Cu(111). **a:** STM image before annealing, depicting the formation of dimers and trimer structures. A defect showing a single molecule is circled in white. The more stable triangular structures are marked with blue and violet circles. At the bottom, high resolution images of the dimer and the trimer before annealing are displayed. **b:** STM image after annealing to 360 K demonstrating the formation of new trimer structures and a monomer. The different nanostructures are also marked with colored circles. At the bottom, high resolution images of these structures (A-D) are shown.

Interestingly, while structure A and structure B appear in 2 orientations (distinguished by a 180° rotation of the C_3 axis and highlighted in Fig. 4.9b) on the surface, structure C is only found in one orientations (but with a C_3 axis). This finding is surprising because the top layer of the Cu(111) surface exhibits a sixfold (D_6) symmetry, suggesting that structure C is influenced by the second surface layer. For chiral molecules as well as single atoms, such a symmetry breaking has been reported and explained by the surface local capacity to screen the dipole moment of a molecule.^{200,201}

In addition to the three trimer structures A-C, an asymmetric structure D is observed. This is identified as a monomer due to its dimensions and shape. It has a curved trapezoidal shape consisting of 2 lines with 4 and 3 distinct features. One of the 3 features in line exhibits the highest intensity. Next to this feature, a small protrusion is found that is very low in intensity, but characteristic for this structure. This single molecule aligns in all 6 symmetry directions of the substrate. Moreover, it can be recognized as smallest building block of the different triangular structures A-C, as shown in Fig. 4.10. Due to an intermediate structure that is probably formed at a

defect site, the head-to-tail assembly forming the triangular structures is observed.

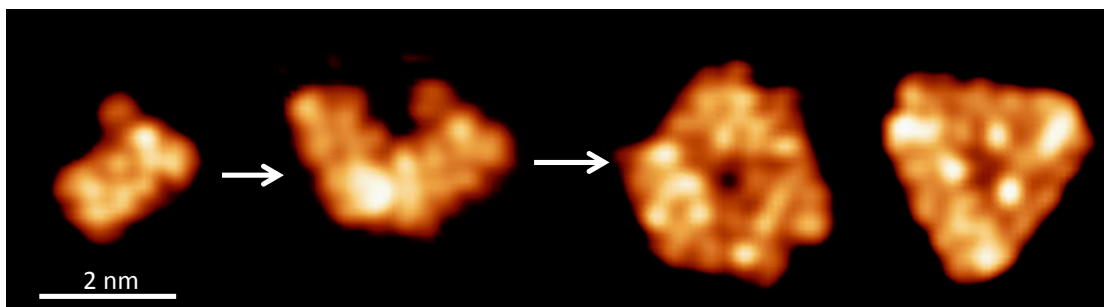


Figure 4.10: STM images showing the assembly of trimer structures of AT-II on Cu(111) from the monomer (left) to a defect structure to triangular structure A or B.

It is very interesting to note that the monomer is only observed after annealing. If the single molecule was a stable configuration on the surface, one would expect to observe it immediately after deposition because the peptide is deposited as single molecule. However, only after annealing, this species is found on the surface. This suggests either a drastic change in conformation due to the elevated temperature, or a deprotonation and subsequent strong binding to the Cu surface or a Cu adatom stabilizing the single monomer configuration. The deprotonation and subsequent coordination is well investigated for small organic molecules in combination with Cu atoms.^{140,141,185} It might also happen for peptides as some residues such as the His, Asp and the C-terminal are acidic and provide good binding sites for metal coordination.¹⁹⁷

4.2. Peptide-metal coordination on the surface

However, depositing the AT-II Cu coordination complex on Cu(111), the same behavior can be observed, as shown in Fig. 4.11: First the formation of dimers and round-edge trimers, and after annealing the formation of three different types of trimers A-C.

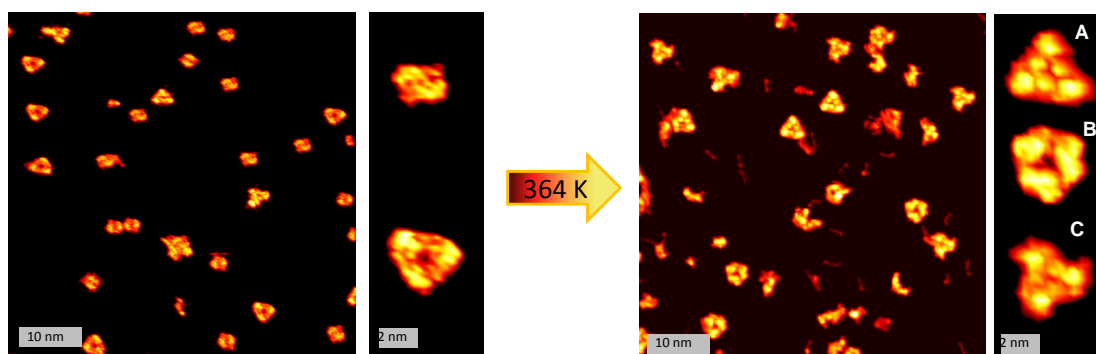


Figure 4.11: STM images of [AT-II+Cu] coordination complex on Cu(111). **a:** STM image before annealing, depicting the formation of dimers and trimer structures as for the non-coordinated peptide. **b:** STM image after annealing to 360 K demonstrating the formation of new trimer structures and a monomer.

The finding that the deposited [AT-II + Cu] complex is reproducing the same self-assembly than the non-coordinated is very surprising and stands in strong contrast to other experiments, such as on Au(111). From these observations on Au(111), it can be excluded that the coordination center is expelled upon deposition. This suggests that the dimer and round triangular structure before annealing already contain Cu. An uptake of adatoms as a metal coordination centers is very likely on Cu(111), but not on Au(111) or Ag(111) because Cu(111) has the most accessible and reactive adatoms of the three surfaces under investigation.^{141,175} As the adatoms are available already at room temperature, at which the deposition takes place, the neutralized peptide is immediately coordinating to a metal center on the surface.

Adatoms are like a 2D gas, so the stoichiometry of the peptide-metal complex is not known. The deposition of the Cu-coordinated peptide yields the same structure as it consists of a neutral peptide and the Cu ion that could relax to the same charge state by charge-transfer with the surface. While the stoichiometry of the Cu-coordinated peptide is 1:1 (peptide:metal) in the gas phase, this can change on the surface by further uptake of more Cu adatoms, if favorable.

From the observations described, another scenario can be hypothesized during the annealing: The formation of the different triangular trimers after annealing can be caused by a conformational change or/via further deprotonation of the peptide. This

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

has to happen for both, the pure AT-II deposition as well as the Cu coordination complex deposition as for both experiments the peptide is positive or neutral in its initial state, respectively. This hypothesis needs to be further tested by, e.g. deposition of the deprotonated AT-II or by calculating the reaction energy of the deprotonation vs. the uptake of metal adatoms.

Evidence for the deprotonation hypothesis is also given by earlier observations on Bradykinin, a nonapeptide:²⁰² Here, it was found that only dimers are formed on Cu(111) and Cu(100) when depositing the protonated peptide. After annealing to 393 K, structures attributed to single peptide were found on Cu(111). Only if deposited in negative mode as [M-H], single molecules could be found on the Cu(100) surface without annealing that resembled the monomer structure on Cu(111) after annealing. This suggests that a deprotonation of the peptide and possible uptake of metal from the surface is happening upon annealing.

For the annealed structures, the combined information about the formation process of the trimer structures (Fig. 4.10) and the model of the [AT-II+Cu] coordination complex on Au(111) yield a model for the structures on Cu(111), as shown in Fig. 4.12. Owing to simplicity, the hypothesized deprotonation of the peptide is not represented in the model and only one metal center is assumed. Most likely, it would enhance the polar interaction. As can be seen, the model fits reasonable well with the monomer. Here, the slight bent in the STM image of structure D is mirrored in the model. The low intensity protrusion corresponds to the phenyl residue of Phe8. The highest intensity is observed close to the coordination center at the Pro7 position, which might also be related to strain in the system.

Employing this building block, the trimer structures can be rationalized. For trimer A, the coordination center is positioned at the corners, where the high intensity is again related to the Pro7 position. Furthermore, the bright features in the center correspond to the Arg2, coming together at the center. At the middle of this structure, the environment of all Arg2 residues coming together could also form a potential coordination site for an additional Cu center from the surface.

For the truncated trimer B, the assembly is not easily elucidated as no distinct feature is observed. In this model, the Arg2 residues interact with peptide backbone, pushing the Phe8 to the inside so that all aromatic residues are decorating the cavity forming a complete circle of aromatic residues. The outside of the structure is still mostly passivated, the His6 at the corner gives a low intensity feature typical for this residue. The windmill trimer C shows again bright features in the center, which correspond to the Arg2 residues in the model and help to adjust the building blocks. In contrast

to structure A, the N-terminal of the peptide is rotated to the inside, so that the Asp1 residue points towards the center. Moreover, the peptides are rotated against each other, giving rise to the windmill shape. Therefore, the Arg2 residues can interact with the residues of Tyr4' and Asp1' of the adjacent peptide. Again, the environment in the center could be a potential coordination site for an additional Cu atom from the center. The structure is passivated on the outside, with His6 corresponding to the small low intensity protrusion at the corner of the windmill's blades.

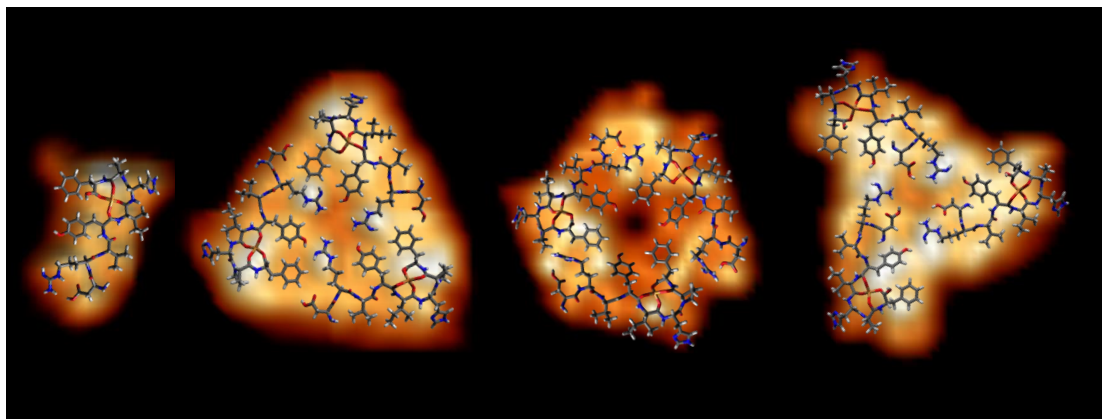


Figure 4.12: Atomistic models of structures formed by [AT-II + Cu] coordination complex on Cu(111). From left to right, it is the monomer D, the triangular trimer A, the truncated trimer B and the windmill structure C.

4.3 Conclusion and Outlook

In nature, the incorporation and formation of peptide metal coordination complexes can enhance the enzymatic activity or lead to toxic structures, as suggested for amyloid fibril formation. This wide range of response demonstrates how powerful and versatile the peptide coordination to metal centers is. In this chapter, the formation and influence of peptide-metal coordination complexes has been explored in an exemplary manner for AT-II coordination complexes with Cu and Ni on Au(111), Ag(111) and Cu(111).

The self-assembled structures of the [AT-II + Cu] coordination complex have been compared with the self-assemblies of the pure peptide. No difference between the coordinated and uncoordinated AT-II assembly could be found on Cu(111). The strongest change was observed on Au(111), where the self-assembly changed from 2D long ranged ordered honeycomb networks to 1D chains with two different motif. This trend follows the interaction strength (reactivity) of the metal surface (Au being more

Chapter 4. Peptides: Metal-coordination to steer the conformation on the surface

inert than Cu). Included in the interaction strength is the concentration and reactivity of the 2D adatom gas on the surface. Therefore, an spontaneous uptake of adatoms to form a [AT-II + Cu] coordination complex is assumed, yielding the same conformation than the deposited [AT-II + Cu] coordination complex.

In addition to the influence of the substrates, the conformational change on the surface by different metal centers has been demonstrated on Au(111) in this chapter. While both, the [AT-II + Cu] and the [AT-II + Ni] coordination complex, show 1D lines on Au(111), the assembly motif is clearly different. As discussed for the models, this is due to different coordination sites. While Cu coordinates with the C-terminal and the oxygens of the backbone (Phe8-Tyr4), Ni preferentially coordinates with the imidazol-N of the His6 and the adjacent backbone oxygens. However, it is challenging to elucidate the position of the metal experimentally in order to verify the preliminary models introduced in this chapter.

X-ray photoelectron spectroscopy (XPS) could help to unravel the charge state of the metal center and might shed some light on the coordination environment. As high resolution is needed, this would need to take place at a synchrotron. Moreover, at a synchrotron, the electronic and magnetic properties of the metal centers could be investigated by X-ray magnetic circular dichroism (XMCD). The electronic structure of the metal center in a peptide complex is very interesting as it determines properties, such as catalytic activity or light absorption.

Moreover, different metal centers could be investigated by further experiments. While there are 83 different metals known, only 13 are employed in natural metal peptide complexes. Each of it has a distinct electronic structure and thus a different influence on the structure and properties of the complex.

In the long run, tuning the structural and electronic properties of metal peptide coordination complexes on the surface could lead to a new class of catalysts. Using nature's tool box and adjusting it to the need of heterogeneous catalysis by materials engineering is a promising path. So far, some MOCN could be demonstrated to work as nature inspired catalyst for the oxygen reduction reaction, employing small and inflexible organic molecules.¹⁸⁶ Being able to fine tune the organic environment of a metal center by exchanging just one amino acid in the sequence might be the key to optimize such properties.

5 Protein imaging by low energy electron holography*

In their native state, proteins are highly functional molecules that take part in nearly all vital processes in the living organism, such as enzymatic reactions or selective transportation (oxygen, or through pores), as well as serving as receptors. In order to understand these processes, the structure of the folded protein has to be solved, including the stoichiometry of protein units in a complex, the connectivity within a complex and to other, small molecules such as drugs or signaling molecules. Moreover, during their function, most of the proteins adapt between different conformations as for example seen in the enzyme ATP Synthase (see Fig. 1) that has three conformation states to promote the synthesis of the biological fuel ATP by binding ADP and phosphate, condensating them to form ATP and finally releasing ATP.⁵ The knowledge of structure and conformational adaption helps to understand molecular processes as well as diseases, and allows to design drugs to specifically target a disfunction.²⁰⁴ For structure determination of proteins and protein complexes, the state-of-the-art methods are X-ray crystallography²⁰⁵ and cryo-electron microscopy (cryo-EM)³⁵. For a long time, X-ray crystallography had a close-to monopoly position in structure determination of biological samples, as can be seen in the contribution to the Protein Data Bank in which about 90 % of the submitted structures this year were provided by this technique (110550 of 123622, update: 18th of October 2016). The technique is based on the diffraction pattern of the incident X-rays caused by the interference with atoms in a crystalline sample. From this diffraction pattern, the 3D structure of the crystal can be solved by computation, yielding a model of the protein crystals with atomic resolution. However, the drawback of this technique is the need for crystalline samples. While

*This chapter is based on a publication²⁰³

Chapter 5. Protein imaging by low energy electron holography

X-ray crystallography is not limited in the size of the molecule, small crystals cause a decrease in inaccuracy as the signal of the diffraction pattern suffers. Finding the right parameters for a protein to crystallize can take several years and might even be impossible for some species, such as membrane proteins which are stabilized by detergent.²⁰⁶

In contrast to X-ray crystallography, EM does not rely on crystalline samples.³⁵ In transmission electron microscopy (TEM), a very thin sample is imaged by high energy electrons (20 keV-400 keV) which are transmitted through the sample. The electron beam is focused by magnetic lenses which also steer the magnification power. By the development of aberration correction of the lenses, the resolution of TEM was enhanced to the low Å regime, thus achieving atomic resolution.^{207,208} TEM achieves the highest resolution for inorganic materials.

For biological samples, the signal is very low, so that averaging is necessary. As the contrast increases with atomic number, staining with heavy metal salts (Mo, U) is routinely used for organic samples. However, enhancing the signal by staining is avoided for structure determination of native proteins because this procedure is known to be able to alter the conformation of the proteins and limit the resolution due to artifacts and the formation of microcrystals.^{209,210} Another possibility is the flash freezing in ethane with liquid nitrogen and subsequent imaging (cryo-EM). In this procedure, biological molecules such as proteins can keep their native conformation while being immobilized in a matrix.³⁴ Moreover, the matrix and the low temperatures protect the species to some extent from radiation damage. Thus, cryo-EM is mostly employed for structure determination of biological molecules.²¹¹ Nonetheless, the large inelastic scattering cross-section of high energy electrons as employed in the newest TEMs inhibits accumulation of sufficient elastic scattering events over time to allow high-resolution reconstruction of just one molecule before it is irreversibly destroyed.³⁴ Therefore, it is still necessary to average over many species.

Several research groups are trying to reduce the radiation damage problem by lowering the electron energy. However, the radiation damage to biomolecules by electrons in the keV range will possibly never allow imaging of truly single proteins at atomic resolution.^{212,213}

The development of direct electron detectors,²¹⁴ that can take images of individual molecules at dozens of frames per second, has advanced the single particle cryo-EM as it allows to reduce the dose of electrons in comparison with a conventional CCD cameras or old-fashioned films. Moreover, sophisticated computer algorithms are employed to categorize species and average over similar (best case identical) specimen

by single particle analysis before reconstruction.^{215,216} This yields 3D structures that match the quality of the crystallographic ones (< 2.9 Å resolution).²¹⁶

However, the averaging process implies that the resolution is highest at the core of the structure where the agreement between different specimen are the highest. At the outside, or at active sites - which are most interesting, yet show the highest variation - the resolution is lower due to inherent variations in the conformation.²¹⁶

In both state-of-the-art methods for structure determination, cryo-EM and X-ray crystallography, the radiation damage is a problem for single molecule studies that has not been solved so far.^{34,52} The development of free electron laser (FEL) as X-ray source with enhanced brightness and short pulses in the fs range has enabled the measurement of small crystal samples in the order of nanometers.^{217,218} Nonetheless, it has also become apparent that averaging over a large number of molecules will not be avoidable with this set-up. Further developments of XFELs with order of magnitude in enhanced brightness and reduced pulse duration might achieve the goal of single molecule imaging in the future.²¹⁹

In contrast, electrons with a lower kinetic energy in the range of 50-250 eV, have been shown to induce no radiation damage in biological samples, such as DNA. Even after long time exposure that yields in summary a dose 5 orders of magnitude higher than the dose used for X-ray crystallography or high energy TEM, no radiation damage could be observed.^{220,221} This observation, combined with the de Broglie wavelength of 0.7-1.7 Å related to the low kinetic energy electrons, makes low energy electron microscopy techniques very suitable for soft matter investigations at atomic scale, thus, even down to the single molecule experiment. Low energy electron holography (LEEH), inspired by Gabor's original idea of holography,²²² is a new approach that has been pioneered by Fink & Longchamp for single molecules.^{221,223}

In this chapter, I present holography measurements of folded proteins performed on single molecules.* While the damage-free radiation of coherent low-energy electrons and the conceptual simplicity of the experimental scheme for holography are appealing, this tool for single protein imaging critically relies on the sample preparation method in UHV. To this end, ES-IBD has been employed to deposit a beam of native proteins on freestanding graphene. Only by the combination of single molecule holography and ESI is it possible to reveal structural details of single native proteins.

*Measurements in collaboration with JN Longchamp and HW Fink

5.1 Low energy electron holography

In 1948, Gabor proposed a new "electron interference microscope", in which the interference between the scattered and unscattered electron wavefronts are employed to record a 3D representation of the investigated object.^{222,224} From the interference pattern, the so-called hologram, the structure of the investigated object can be reconstructed by back propagation. Gabor demonstrated the working principle with a monochromatic light source, however, electron holography employing low energy electrons instead of light, could not be established due to several reasons. First, an appropriate source for monochromatic electrons is necessary, which at that time, did not exist. Moreover, in order to avoid artifacts, the substrate has to be transparent for electrons, while at the same time, it needs to be robust enough to endure the deposition process. This can be glass or anything transparent for visible light, but barely any medium is known being suitable for electron holography.

Nowadays, this microscopy technique, which was inspired by Gabor, is called low-energy electron holography²²⁵ (LEEH). The controlled fabrication of coherent point sources for electron beams²²⁶ enabled the employment of low energy electrons. Additionally, it has been shown recently that ultraclean freestanding graphene fulfills all requirements of an appropriate substrate and has thus opened the way for artifact free measurements of single molecules by LEEH.^{227,228} The set-up is conceptually very simple, as shown in Fig. 5.1a. It consists of a point source for coherent low energy electrons that interfere with the sample. The interference pattern is imaged at a phosphorous screen behind the sample. From the hologram itself, as shown in Fig. 5.1b, the structure and shape of the object cannot be identified directly. Every object on the graphene results in a main feature in the hologram and the spreading interference fringes, similar to the rings on a water surface when a droplet hits the surface. These spreading interference fringes can also overlap. A CCD camera digitalized the hologram (see Fig. 5.1b) and enables thus the computational reconstruction of the imaged object.^{229,230}

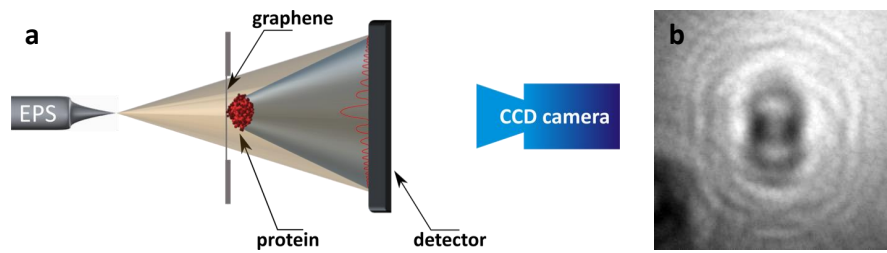


Figure 5.1: Schematic set-up of the LEEH microscope. **a:** A beam of coherent low energy electrons illuminates the complete sample, the interference pattern of the reference and object wave are collected at an phosphorous detector and recorded by a CCD camera. **b:** Exemplary hologram showing the interference fringes that are necessary for reconstruction.

For measurements, the samples are presented to a divergent beam of highly coherent low energy electrons which are generated by an atomically sharp W-tip oriented in (111) direction. This tip performs as electron point source (EPS)^{226,231} emitting single electrons. The ratio between the detector-source distance and the distance between the sample and the source controls the magnification which can be as high as 5 orders in magnitude. While the distance between the sample and the detector is kept constant, the position of the tip varies between 2 mm and 100 nm in front of the sample and is controlled by a 3-axis piezo positioning system. The electron energy for a given electron emission current is given by the sample-source distance. The acquisition time to yield a sufficient signal-to-noise ratio depends on the thickness of the sample because the scattered and unscattered electrons are contributing to the image formation. It can be between hundred microseconds for thin samples such as graphene²³² and hundred milliseconds for thicker samples such as 3D structures of proteins making holography a suitable tool for looking into structural dynamics. All holograms presented here were recorded with electron energies between 60 eV and 140 eV and an emission current of typically 50 nA. A typical acquisition time for a hologram amounts to 100–200 ms.

In contrast to a diffraction pattern, the taken hologram also contains the phase relation of the incident (reference wave) and the scattered wave (object wave). Thus, an unambiguous reconstruction of the investigated object is possible. The transformation from the hologram to the object is performed by numerical reconstruction, yielding the corresponding micrographs as shown in Fig. 5.2. To this end, the back-propagation of to the objective plane is achieved by evaluating the Fresnel-Kirchoff integral transformation.^{229,230} For instance, in Fig. 5.2d, an individual protein is displayed after hologram reconstruction. The diffuse rings around the reconstructed

object are due to the presence of the out-of-focus twin image inherent to in-line holography.²²⁹ Nonetheless, as apparent from the high contrast images 5.2d,e,f, also details of proteins in different orientations can be imaged, even while forming agglomerates of two or three proteins.

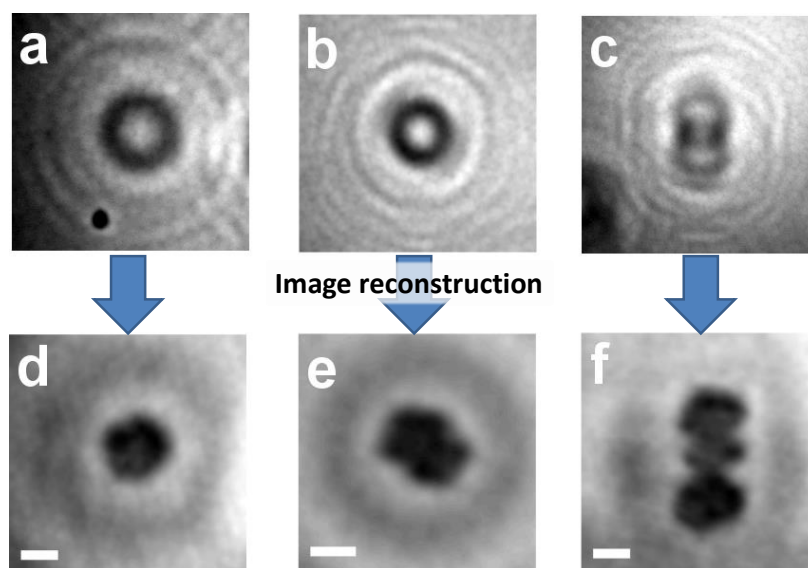


Figure 5.2: Low energy electron holograms of CytC and their reconstructed images. **a-c:** Holograms of Cyt C showing the typical interference fringes. **d-f:** Reconstructed images of a single molecule (d), a dimer aggregation (e) and a trimmer aggregation (f). The scale bars correspond to 2 nm.

The spatial resolution attained in the hologram is estimated by the Abbe criterion^{233,234} and by measuring the largest angle under which interference fringes are observable.^{230,235} In Fig. 5.2, a resolution of 7-8 Å is calculated and agrees with the value taken from the reconstructed images by measuring the edge response over the protein structure.²³⁶ A similar resolution is estimated for all other micrographs presented in this chapter.

5.2 Sample preparation and transfer

While the damage-free radiation of coherent low-energy electrons and the conceptual simplicity of holography are appealing, this technique for single protein imaging critically relies on the sample preparation method. The sample substrate for this is ultraclean, free-standing graphene.²³⁷ In addition, also deposition of the analyte species has to be performed under well-controlled conditions and is thus done by ES-IBD. It is important that the samples remain in UHV for the complete time of the experiment to avoid any contamination of the sample which could be mistaken for the analyte species. A scheme of the experimental workflow is shown in Fig. 5.3, starting with the substrate characterization by holography followed by the m/z selected deposition and eventually the holography measurements of the deposited proteins.

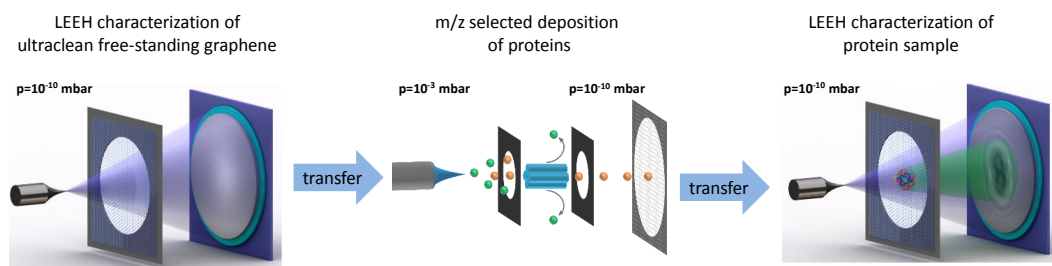


Figure 5.3: Schematic work-flow of the experiment. From left to right: Characterization of an ultraclean graphene sample by LEEH. Deposition of m/z filtered beam of proteins onto freestanding graphene by ES-IBD. Imaging of the proteins within the previously characterized region by LEEH.

For the substrate preparation, graphene is transferred onto a 100 nm thick SiN membrane containing $500 \times 500 \text{ nm}^2$ apertures and mounted in a standard TEM grid/ sample holder. The freestanding graphene is cleaned by the Pt-metal catalysis method, which employs deposited Pt to catalytically remove the carbon residues when the sample is heated to 350°C in ambient conditions.²³⁷ Subsequently, while still being hot, the sample is transferred into the LEEH microscope at 10^{-10} mbar and subsequently characterized by holography to make sure that no contamination is on the substrate. A UHV suitcase (Ferrovac GmbH, Zurich, see also chapter 1), shown in Fig. 5.4a, is used to maintain the vacuum conditions during the sample transfer between the holography set-up and the ES-IBD instrument, which both work in UHV (10^{-10} mbar). For the transfer, a load lock is installed on each instrument where the suitcase can be attached. In addition to a turbo pump, a home-made cryo pump (charcoal trap at

LN₂ temperature) is installed in each load lock to speed up the pumping time to reach 10⁻⁹ mbar in the load lock.

In order to ensure that the sample is not contaminated during these transfer processes, the procedure was tested once with a clean substrate without an actual deposition taking place, but characterizing the substrate before and after the complete transfer cycle by LEEH. As can be seen in Fig. 5.4, there is no relevant change related to contamination visible between before (Fig. 5.4b) and after the transfer (Fig. 5.4c).

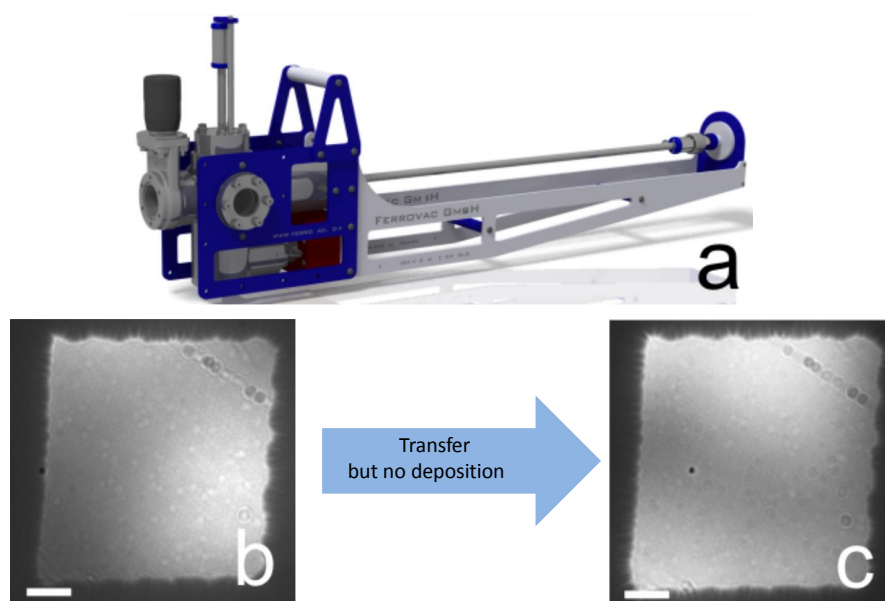


Figure 5.4: Transfer standard tested once for a sample without deposition. Compared to before the deposition (b), no additional features are found on the ultraclean free-standing graphen after the transfer to the ES-IBD without deposition (c). The scale bars correspond to 100 nm.

5.3 Native mass spectrometry with the ES-IBD set-up

Non-covalent interactions, such as electrostatic, hydrophobic, or van-der-Waals interaction, play a vital role in biology and the correct folding and function of proteins. Individually, these interactions are weaker than covalent bonds, yet in large numbers they stabilize the specific 3D conformation, called native conformation, of a protein in solution. However, the conformation of proteins and protein complexes is susceptible to temperature, pH of the solution as well as the salt concentration in the solution. Therefore, neither strongly acidic conditions nor organic solvents can be used for native ESI-MS. In order to balance the preservation of the specific non-covalent in-

5.3. Native mass spectrometry with the ES-IBD set-up

interactions with the preparation of a suitable solution for ionization by electrospray, volatile buffers, such as ammonium acetate, are employed. It has been demonstrated, that ESI is a suitable technique to transfer proteins and protein complexes into the gas phase while preserving their non-covalent interactions.⁴⁷

The preservation in the gas phase of the conformational state that was reached by solvation²³⁸ is only possible because this state is very compact and some forces, such as polar interactions and hydrogen bonds are even stronger without solvation. Moreover, there are no driving forces for unfolding in vacuum.²³⁹ This means that folded structures, protein complexes and even membrane proteins²⁴⁰ can be ionized while maintaining their conformation. The wide applicability of native MS to samples of different mass, symmetry or polydispersity and its capability to investigate different oligomeric states of the same protein without the need to chemically label the species makes it a powerful tool for biocomplexes.

Soft-landing of native proteins such as cytochrome C, lysozyme and trypsin has demonstrated that they mostly retain their biological activity after the complete procedure and subsequent dissolving in aqueous media.^{61,66} Even complete viruses were shown to remain intact upon soft-landing in high vacuum.⁶³ However, re-dissolving the deposited species allows also for re-folding in solution. Thus, the question has to be raised of what the influence of the surface is.^{200,241} While it is suggested that proteins keep their tertiary and quaternary conformation upon deposition,⁶¹ another study demonstrated the enhancement of only one peptide conformation on the surface of SAMs²⁴¹ demonstrating the severe influence that the substrate can have on the conformation. However, the influence of the surface on protein structures in vacuum has not yet been investigated deeply on an atomic scale. For weakly interacting substrates, such as graphene, we can assume that the folded structures are very similar to their conformation in the gas phase.

Here, we exploit native mass spectrometry in our ES-IBD setup to transfer the proteins cytochrome C (CytC), bovine serum albumin (BSA) and hemoglobin (Hb) in their folded state in the gas and subsequently deposit them on free-standing ultraclean graphene to image them by LEEH. We chose these three proteins as they possess different shapes and sizes: CytC is the smallest of the three proteins and has a globular shape. By deposition on Cu(100) and Au(111) and subsequent imaging by STM, we have already demonstrated that we can deposit this protein in the folded state.⁷¹ BSA has a very distinct shape in its native conformation that is often described as heart-shaped.²⁴² This protein is thus very helpful to explore the resolution as well as

the different orientations of the protein after the soft-landing. Hb is the largest of the three proteins under investigation and exists in the native state as tetramer. It consists of two α and two β subunits which are non-covalently, but specifically bonded. By the high mass of this protein complex, we are exploring the detection limit of our home-build TOF. Moreover, depositing and imaging a protein complex is the first step towards a characterization of connectivity in a protein complex by LEEH.

For CytC (bovine, Fluka 30398, $M = 12384 \text{ g/mol}$), a solution of 0.15 mg/mL in 50 mM ammonium acetate buffer (final concentration of $1 \cdot 10^{-5} \text{ M}$) was used at a flow rate of $25 \mu\text{L/h}$ yielding a current of 1.1 nA on the TOF aperture before mass selection. As seen in Fig. 5.5a, low charge states from +5 to +7 ($2476.8\text{--}1769.1 \text{ m/z}$) are present in the mass spectrum with high intensity. These charge states correspond to CytC in native state.⁶⁴ In addition, high charge states at the low m/z range are also present as well as some contamination in the higher m/z range. Note that the peaks of the native CytC are strongly distorted as the detector had to be set to high amplification voltages due to the limited dynamic range of the home-built TOF-MS in the ES-IBD instrument.

In order to obtain a beam of native protein, mass selection is performed. To this end, the m/z selective quadrupole is tuned to select a window from 1250 Th to approximately 3500 Th by setting the 0.5 MHz RF-amplitude at 700 V with a differential dc voltage at 5% . As seen in Fig. 5.5b, this results in a spectrum in which the native protein (charge state +5 to +7) is dominating, while unfolded proteins and undefined agglomerations are removed. The ion current measured is decreased by the m/z selection, resulting in 20 pA on the sample. In total, a charge of 20 pAh was deposited, yielding a submonolayer coverage of 3D folded CytC on free-standing graphene.

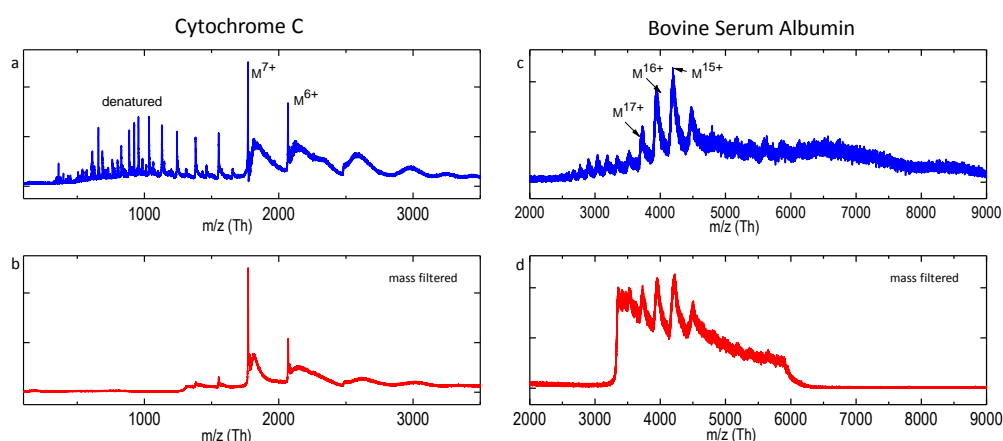


Figure 5.5: Mass spectra of CytC and BSA. The unfiltered RF-only mass spectra of CytC (a) and BSA (c) are presented with the peak assignment of the charge states. In the bottom row, the m/z filtered spectra for CytC (b) and BSA (d) are shown.

5.3. Native mass spectrometry with the ES-IBD set-up

Ion beams of BSA (Sigma Aldrich A4919, $M = 66430$ g/mol) are prepared in the same way as for CytC employing 0.66 mg/mL in 50 mM ammonium acetate for a final concentration of $1 \cdot 10^{-5}$ M. In the unfiltered mass spectrum (Fig. 5.5c), intense peaks around 4000 Th are observed corresponding to charge states $+17$ to $+15$ and are related to the folded protein. In the lower m/z range, higher charge states of (partially) unfolded BSA protein is found, whereas the signal > 6000 Th is attributed to unspecific agglomerations. Thus, the region between 3250 Th and 6000 Th is selected for deposition, as seen in the lower panel of Fig. 5.5d. A total of 20 pAh of the mass selected ion beam is deposited on free-standing graphene samples in UHV.

For hemoglobin (Hb, Sigma H2500, $M = 64500$ g/mol), a solution of 0.3 mg/ml in 50 mM ammonium acetate buffer yield 600 pA on the TOF aperture. The corresponding mass spectrum in Fig. 5.6 is very complex and can be only partially resolved due to the limited performance of the resolution and the dynamic range of the home-built TOF at the ES-IBD.

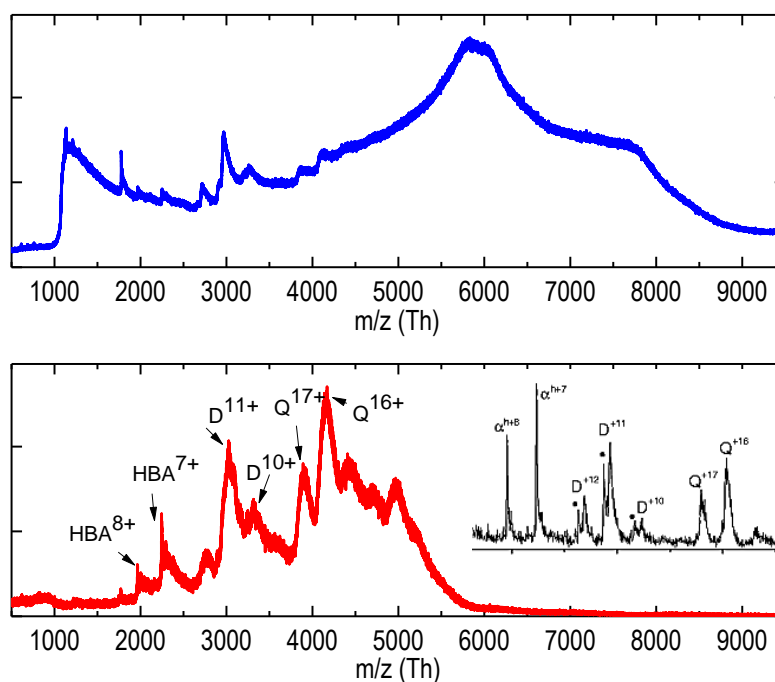


Figure 5.6: Mass spectrum of Hb. top: rf-only mass spectrum showing the full range in which specific peaks are resolved only partially. bottom: mass spectra of the m/z selected beam in which the tetrameric Hb protein complex is found as main species (Q), along with some dimer complexes (D) and monomers of Hb (HBA). The inset shows a reference spectra to assign the peaks correctly.²⁴³

Nonetheless, by comparison with a reference mass spectrum from literature²⁴³ (see inset), characteristic peaks can be identified, such as the intact tetrameric complex with charge state +16 and +17 at 4031 Th and 3794 Th respectively. After removal of the unspecific agglomeration in the high m/z range (>5000 Th), these peaks are more pronounced in the mass spectrum (lower panel Fig. 5.6). The beam transmitted for deposition contains native intact HG complexes (Q^{+17} and Q^{+16}), as well as some dimers (labeled D) and monomers (HBA) that are found in the lower m/z range. In total 70 pAh of this beam were deposited on the substrate.

5.4 Imaging folded protein structures by holography

After deposition, the sample is transferred to the LEEH microscope for imaging. Fig. 5.7 presents the results of the CytC deposition. Before deposition (Fig. 5.7 a), the area of free-standing graphene is clean and shows only few contamination on the homogeneous surface. On the ES-IBD set-up, a sub-monolayer coverage of CytC was deposited with a mass-selected ion beam (Fig. 5.7b). The hologram of the same area of graphene after deposition (Fig. 5.7c) shows that the deposition was successful, as a variety of small globular objects possessing the same dimensions (~ 2 nm) are found on the surface. Larger objects with a linear shape are formed by non-specific agglomeration due to diffusion of the protein on graphene. As discussed before, these features are truly from the deposition because the complete transfer process without deposition did not yield any new features on the graphene.

In the cases when the objects are well-separated, they can be attributed to single CytC proteins. At high magnification (Fig. 5.7 d-j), the shape of the individual protein can be observed. It becomes clear that the shape of the objects differ from each other, while having about the same dimension in size that fits to a single CytC protein. These differences in shape are due to the random adsorption orientation of the protein as a consequence of the deposition process. A comparison of the micrographs with the structural model obtained from X-ray crystallography (Protein Data Bank, pbd id: 1HRC) agrees very well for the different orientations (see lower panel in Fig. 5.7).

5.4. Imaging folded protein structures by holography

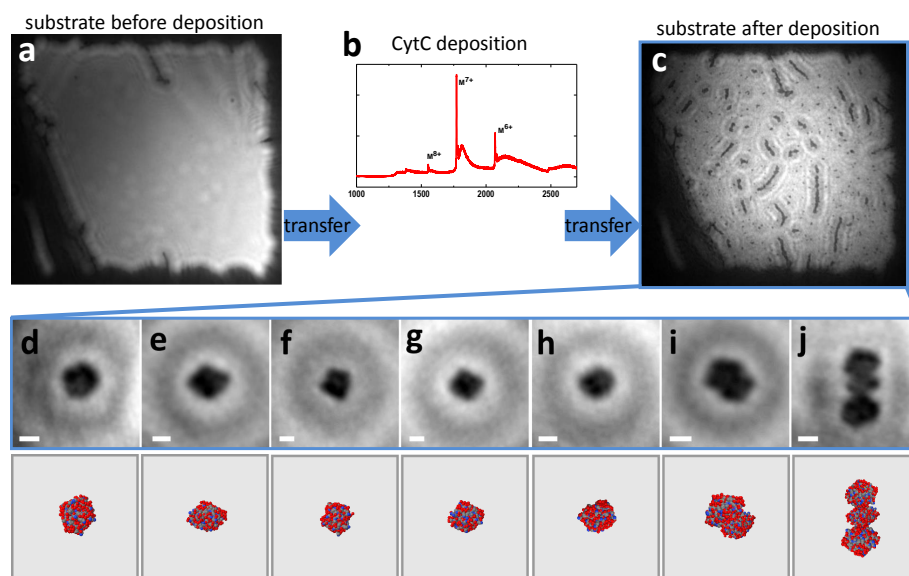


Figure 5.7: Complete data set for the imaging of CytC. **a:** Low-energy electron image of ultraclean graphene covering an $500 \times 500 \text{ nm}^2$ aperture before protein deposition. **b:** Mass-spectrum of the mass-selected CytC beam. **c:** Survey image of the very same free-standing graphene region after deposition of CytC. **d-j:** Low-energy electron micrographs with suggestions for possible protein orientations based on the averaged protein structure derived from X-ray crystallography data and documented in the protein data bank (pdb id: 1HRC). The scale bars correspond to 2 nm.

Measurement over an extended period of time did not reveal any changes that could be attributed to decomposition or beam damage in any kind as can be seen in Fig. 5.8. Here, a time series is presented. While the distinct structure in the middle of the image serves as reference point, the diffusion of the upper right species is clearly visible demonstrating that this technique can also image time resolved processes.

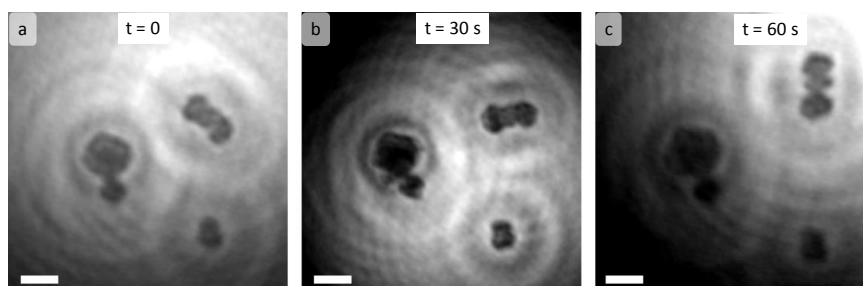


Figure 5.8: Time evolution of the orientation of CytC complexes. The time lapse between subsequent observations amounts to 30sec. From these images it is evident that at least some of the deposited proteins are mobile on freestanding graphene. Scale bars account for 5 nm.

For BSA, the same experimental work-flow was used. In the micrographs of single BSA (see Fig. 5.9 upper panel), the random orientation of molecules on graphene is clearly visible. In contrast to CytC, which is nearly spherical, BSA possesses a distinct 3D shape that is often described as heart-shaped.²⁴² Thus, the micrographs of BSA exhibit very distinct features. The agreement between the micrographs and the structural model (pdb id: 3V03) with the according orientation demonstrates the credibility of the imaging method. Moreover, this finding also shows that the conformation of deposited folded protein on a surface in UHV is very closely related to the native structure as observed in the protein crystal.

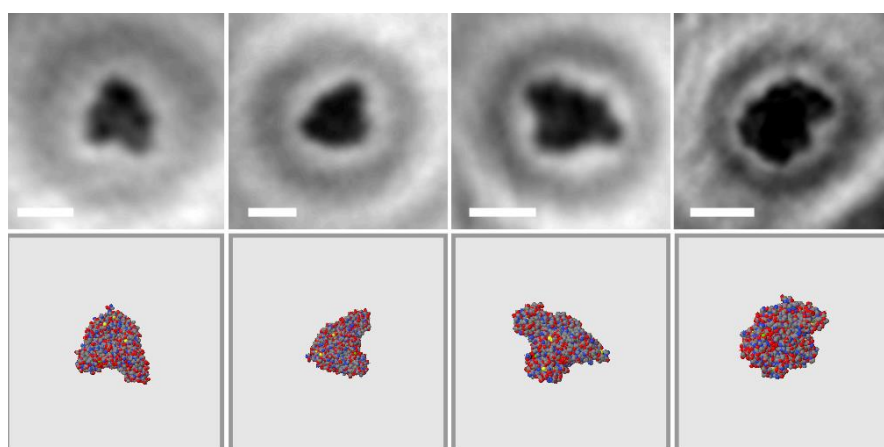


Figure 5.9: Low-energy electron micrographs of BSA in different orientations on graphene. Top, low-energy electron micrographs of BSA. Bottom, the atomic model of BSA (pdb id: 3V03) in the corresponding orientations. The scale bars correspond to 5 nm.

Another way to qualitatively estimate the agreement/quality of this approach is the comparison of the micrographs with simulated electron density maps at the corresponding resolution. To this end, the electron density of BSA was simulated and represented at an isosurface of constant electron density at a resolution of 8 Å using Chimera,²⁴⁴ a software originally developed for the analysis of cryo-EM images. In Fig. 5.10, micrographs of two differently oriented proteins are compared to a isosurface of constant electron density from top (middle panel) and from the side view (bottom panel). The shape of the micrographs fit very well with the top view of the isosurface, yielding the same information as Fig. 5.9. By enhancing the contrast of the micrographs (compared to Fig. 5.9), the change in contrast within the molecule gets more pronounced. This contrast is due to different absorption of the low-energy electrons. For dark regions, the sample is absorbing more electrons. In order to

5.4. Imaging folded protein structures by holography

associate the difference in absorbency and associate it with structural features, the information from the side view of the electron density isosurface is very helpful. The side view in the lower panel is created looking at the molecule from the direction marked with the red arrow in the top view. For clarity, the free-standing graphene is indicated by a blue line. Of specific interest is the side view of the protein presented in the right column of Fig.5.10 because it shows that in this specific orientation, BSA is of strongly varying thickness with a maximum towards the center. The darker region observed in the micrograph evidently corresponds to the higher absorption due to the increased protein thickness in this area. This observation and analysis demonstrates that already at this resolution, information about the 3D structure of the protein can be gained. It also illustrates the future ability to gain a complete 3D structure from a single low-energy electron hologram of proteins at least as large as 66 kDa.

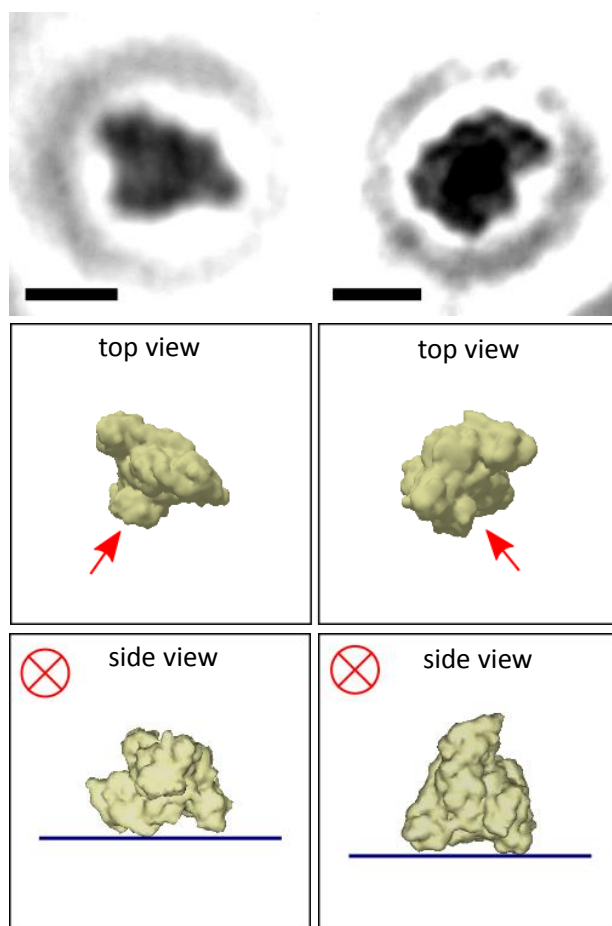


Figure 5.10: Comparison of low-energy electron micrographs of BSA with simulated electron density isosurface. Top: low-energy electron micrographs of BSA. The scale bars correspond to 5 nm. Middle: Electron density isosurface simulated at a resolution of 8 Å and rotated to match the orientation of the proteins presented in the top row. Bottom: Side view of the electron isosurface structure along the directions of observation indicated by the arrows shown in the middle row. The graphene is sketched as blue line.

In order to take it one step further, we also deposited and imaged the non-covalent protein complex Hb. It has been extensively shown that ESI is capable of ionizing and transferring protein complexes into the gas phase while preserving their stoichiometry and native conformation.⁴⁷ The structural characterization of non-covalent protein complexes is of uttermost importance as they reveal position for interactions with small molecules, which is important for drug design,²⁴⁵ or to investigate different states of a functional protein complex, such as ATPase⁹. The data of Fig. 5.7i, j demonstrates that protein agglomerations formed by two or three CytC can be resolved.

As a protein complex that is suitable for our ES-IBD set-up, we investigated Hb. Fig. 5.11 depicts micrographs of two different Hb tetramers on free-standing graphene with the corresponding atomic model (lower panel, pdb id: 2QSS). While the dimensions agree with the expected value, the found structures deviate from the model and do not match as well as for the single proteins CytC and BSA. This might have several reasons. For one, the structure of the protein complex might deviate from that of the native structure due to the influence of surface. However, as we have not experienced this with CytC or BSA, we exclude strong deviations.

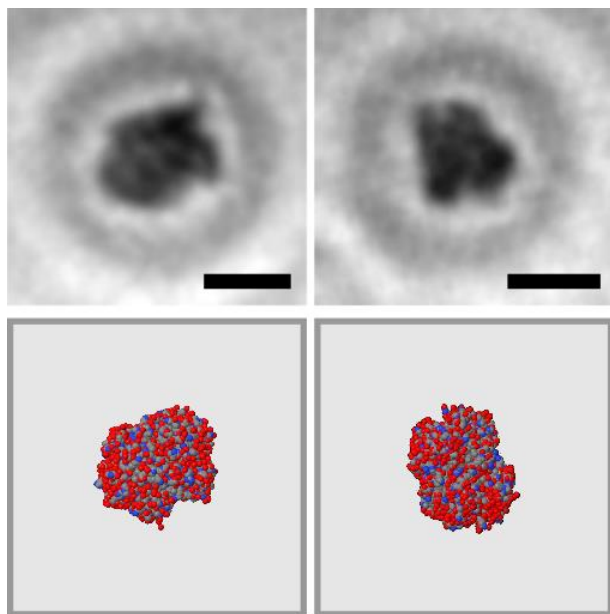


Figure 5.11: Low-energy electron micrographs of two individual Hb and the atomic model in the corresponding orientations. Top: Two micrographs of Hb soft-landed onto freestanding graphene. Bottom: Suggestions for possible orientations based on the averaged protein structure derived from X-ray crystallography data and documented in the protein data bank (pdb id: 2QSS). The scale bars correspond to 5 nm.

Moreover, a range of different species is deposited. Referring to the mass spectrum of the deposition (Fig. 5.6), it is apparent that in addition to the tetramer complex, the beam also contains single and dimer species. Additionally, the tail at the high m/z range is most likely due to some tetramer species with buffer adducts from the solvent. While we can exclude by the size of the imaged feature that it is a dimer or monomer species, an adduct attachment cannot completely be ruled out. This demonstrates how important a well-controlled sample preparation is.

Nonetheless, one has to keep in mind that the measured data of one single protein, which possesses a flexible conformation is compared to the averaged structure over millions of molecules. When an averaging process is involved in the imaging of a highly flexible protein, discrete conformations cannot be distinguished and only an average structure evolves. However, with a technology capable of imaging individual proteins, like LEEH, the entire conformational landscape is revealed. It is therefore not surprising that structural differences between the low-energy electron micrographs and the atomic model are apparent. This finding emphasizes the great strength of LEEH by which we can image single species with $\sim 8 \text{ \AA}$ resolution and present their variation.

Furthermore, a much larger set of images will be needed in the future to address the full conformational landscape of flexible proteins as for instance hemoglobin and to be able determine a protein structure and judge the abundance of the variants. Here, we present two images of hemoglobin that could be unambiguously associated to its atomic model in a specific orientation.

5.5 Outlook and future plans

Imaging a single protein is possible by combining the potential of ES-IBD for UHV sample preparation of native proteins with the high resolution and non-destructive imaging capabilities of LEEH. By means of three examples, the potential of this combined approach has been demonstrated. Up to now, the comparison of the low energy electron micrographs with atomic models available from the protein data bank, or the corresponding electron density maps, has the character of a control experiment. Nonetheless, this first step already shows the feasibility of this novel methodology. The potential for structure determination is clearly demonstrated in these experiments, but also reveals some targets for future improvement.

The LEEH microscope set-up is convincingly simple. In principle, the imaging resolution is only limited by the wavelength. The achieved resolution of $7\text{-}8 \text{ \AA}$ can thus

be further improved. Right now, the limiting factor is found in mechanical vibrations, as the spacing between consecutive interference fringes gradually decreases towards higher orders which carry high-resolution details. Thus, in order to increase the mechanical stability and approach atomic resolution, improved damping of the set-up is necessary. As 3D information is encoded in a single in-line hologram, improved spatial resolution will already permit to determine the (x,y,z) spatial coordinates of every atom of a protein from this very record. A complementary strategy to reveal the complete 3D structure of a single protein is to add tomographic capability to the experimental set-up by an additional rotation axis of the sample. This experimental approach is also known for TEM set-ups.

Further improvement in resolution might be gained by a low-temperature measurement. By cooling the sample stage with liquid nitrogen or helium, diffusion and vibrations of the proteins can be avoided.

Moreover, fundamental questions on the influence of the substrate have to be addressed in the future concerning the deposition of native proteins and protein complexes. For proteins, graphene in vacuum represents a significantly different environment than the aqueous medium of a cell and might induce some changes in the conformation. Although the forces and interaction strengths change, there is strong evidence that proteins and protein complexes can be soft-landed intact while preserving their tertiary, respectively quaternary, structure. The additional interaction with the surface is another huge influence and strongly depends on the substrate. Preliminary STM measurements of native Hb, as shown in Fig. 5.12 suggest that on Cu(100), the protein unfolds as Cu(100) represents a strongly interacting substrate.

This is in strong contrast to earlier measurements for CytC, but demonstrates that the interaction with the surface has a severe influence. Thus, the choice of graphene is not only necessary for LEEH, but also provide a less interacting surface that does not force the protein to unfold. The low-energy electron micrographs presented here are strong evidence that proteins in a folded state are stable on graphene in UHV.

Another interesting approach for future experiments is the deposition on cold substrates, going along with the measurement at low temperature. While the measurements at low temperatures can increase the resolution, also the deposition is interesting as it freezes out the conformation upon landing. This also opens ways to coevaporate molecules, such as water, in UHV and investigate the structural changes, such as a hydration shell.

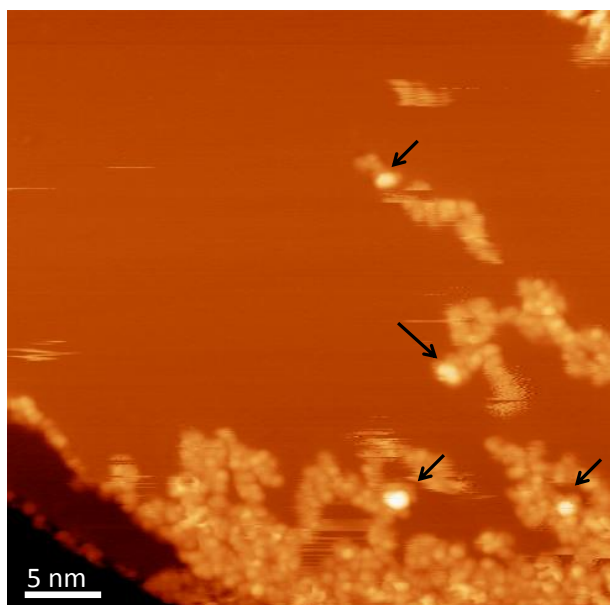


Figure 5.12: STM image of Hb on Cu(100). Arrows mark the features that could correspond to partially folded structures.

The ultimate goal of directly imaging structures of unknown proteins or protein complexes and determining their structure at the atomic levels also requires instrumental efforts at the ES-IBD instrument towards a better chemical and conformational selectivity. Here, the limitations of the ES-IBD set-up are clearly in the mass spectrometer. The home-built TOF mass spectrometer is not designed for extremely large molecules or complexes and thus, the dynamic range and the resolution for high m/z range is limited as seen in this chapter. While an improved mass spectrometer can resolve the chemical constitution, a control of the conformation is only possible by other means, such as an ion-mobility set-up. However, all instrumental changes have to conform with the purpose of deposition, meaning that a high transmission and a continuous beam is preferable.

Additionally, the expertise for native sample handling needs to be expanded for stable beams of native proteins. The unspecific agglomerations that are observed in the mass spectrum accounting for the broadening of the peaks and the intensity at the higher m/z range (Fig. 5.6 and Fig. 5.13a) can be reduced by the right treatment of the solution and the ESI process. In collaboration with Joseph Gault (Robinson Group, University of Oxford), we could establish a stable beam of native BSA, as shown in Fig. 5.13b by purifying the sample solution by buffer exchange and filtering before spraying as well as employing silver-coated silica emitters with no external flow rate for nano-ESI. Although much smaller amounts are used for this technique, the ion current after m/z

selection is comparable to the previous experiments as the beam is much cleaner and the ionization yield higher. However, for native ion beams, the sample preparation is not straight forward and needs to be optimized separately for different species. For specific protein samples, this can even start with the expression of the protein itself. With the recent advances in electrospray ionization and native mass spectrometry of large protein complexes, in particular membrane proteins, the structure of these biologically relevant, but reluctant to crystallize entities may become accessible by combined ES-IBD and LEEH in the future.

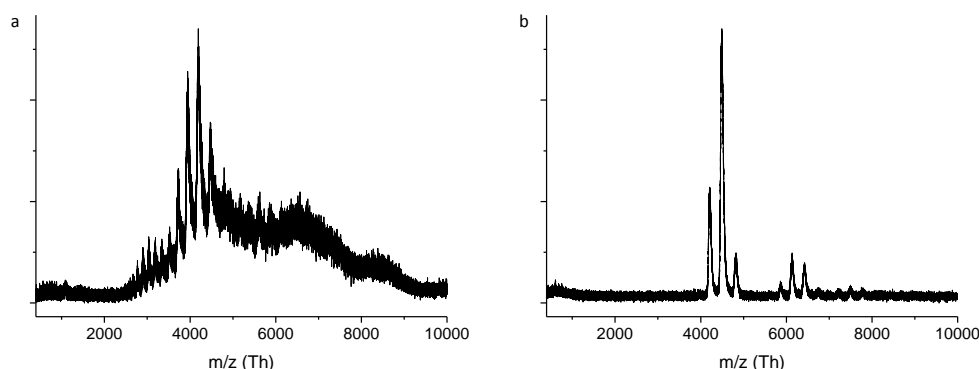


Figure 5.13: Comparison of mass spectra of BSA. **a:** Mass spectrum of BSA, as used for deposition (see also Fig. 5.5). **b:** Mass spectrum of BSA with proper sample treatment for native beams. Sharp peaks are obtained and no unspecific agglomeration at high m/z range is observed.

Conclusion and Outlook

The combination of ES-IBD with UHV imaging techniques is very powerful and marks a new approach that, in the future, might be able to answer pending questions of biology as well as enable to explore new materials. In this thesis, I have presented how the vacuum preparation by ES-IBD can enable UHV tools for the molecular imaging of biomolecular nanostructures.

For isobaric species, where classic mass spectrometry is ambiguous, STM can prove powerful for the structural characterization. As seen for the disaccharides in Chapter 2, STM is able to resolve the single building block units. Consequently, STM imaging might enable to see branching and thus help to identify unknown sequences.

Moreover, STM is suitable to investigate the adsorption and self-assembly of oligopeptides. Exploiting the sequential nature of these biopolymers in Chapter 3, I have demonstrated the sequence controlled self-assembly resulting in a long-ranged ordered honeycomb network. MD and DFT calculations showed that the molecule is adsorbed flat on the surface, so that the residues are stretched out to both sides and are approachable for intermolecular interaction. Employing a sequence with bulky residues avoids self-interaction and folding, and allows for the peptide to be used as building block with a rigid backbone, yet flexible residues to be fault-tolerant.

ES-IBD is powerful as deposition technique because it is very versatile. Thus, in order to manipulate the conformation and bring in more potentially active sites, I have investigated the soft-landing of peptide metal complexes in Chapter 4. By comparison with the pure peptide, we have verified that the peptide-metal coordination complex is actually intact. Due to the preferred coordination environment, various metal centers form different nanostructures on the surface, which can be a tool to control the nanostructures of peptides on the surface.

STM has been employed to investigate relatively small biomolecules in a non-native conformation, however, it is not the appropriate tool to investigate the 3D structure of folded peptides/proteins. In Chapter 5, I have presented the single protein imaging by

Conclusion and Outlook

LEEH. With this emerging imaging technique, it is possible to get images of complex 3D molecules without averaging. Thus, this technique might have the potential to compete with the state-of-the-art structure determination by cryo-EM.

There are two major paths that one can follow for the future, dividing into (i) improving and developing of the instrumentation and (ii) employing the given set-up for relevant questions. As the latter issue has been addressed already at the outlook of every chapter, I would like to focus more on the instrumentation here.

As the sample preparation by ES-IBD is based on mass spectrometry technology, it can also profit from the substantial development in this field. Recently, ion mobility mass spectrometry has been demonstrated for biological samples to give new insights into structure and has consequently gained a lot of attention. In this technique, the selection of species is not only possible by mass-to-charge ratio, but also by the cross section of the species, which relates to the shape.^{246,247} Thus, a conformation selection is possible and enables to draw conclusions about the structure of the analyte. While an implementation for deposition might be challenging because it can reduce the transmission resulting in a loss of ion-beam intensity, it enables the investigation of distinct conformations. A conformation selective deposition allows to address the long-standing question how much, and in which way, the surface influences the conformation. In previous work, it could be demonstrated that the charge state of a protein, which influences the conformation in the gas phase, can also be used to steer the conformation on a surface, if diffusion is restricted.⁷⁰ However, for short sequences, such as the nonapeptide Bradykinin, folding and induced fit could be shown to happen at the surface.¹⁶⁵ This means that the conformation can be changed by the influence of the surface. By selecting different conformations and image them independently on a surface, we can shed light on how the conformation is influenced by the surface.

Investigating, preserving and manipulating the conformation is a current issue in mass spectrometry and the field of structural biology. The huge success of ESI is also due to the fact that this technique is able to do all this: by the many control parameters that are accessible with ESI, the preservation as well as the controlled manipulation of conformation is possible. As shown in the Chapter 5 of this thesis, deposition of folded proteins via native MS opens the path to biological relevant questions by imaging with LEEH. Even if atomic resolution is not yet reached, questions about the connectivity of protein complexes can be answered. This is especially important for drug design

where not only the folded shape of a new drug matters, but also the sites at which it connects to the biological target.^{245,248}

For successful ESI of native proteins, the protein expression and purification is as important as the choice of solution parameters and the control parameters of the spray for gentle ionization and optional activation. The right treatment enables to work on biological samples and tackle structure elucidation that is not possible otherwise, such as integral membrane proteins (IMP). These proteins fulfill important functions in the cell, such as intracellular compartmentalization of organelles or the regulation of cell signaling through surface receptors. However, their structure is very difficult to elucidate as they are stabilized within a lipid membrane and possess an amphiphilic nature and thus are not easily accessible for structure determination, e.g. by crystallization.²⁴⁹ Only recently could it be shown that using detergents to build up a micelle, the IMP can be transferred into the gas phase in the native state.²⁴⁰ An activation is necessary to release the IMP from this micelle environment, either by collision with gas atoms in a collision cell or by increase backing pressure in the initial high-pressure region.²⁵⁰ By the gained access of these protein complexes in the gas phase by ESI, it is worthwhile to investigate their structure after soft-landing.

In cooperation with the Robinson group at the university of Oxford, we started working on integral membrane proteins and could replicate the mass spectrometry experiment for the trimeric ammonium transporter (AmtB)^{250,251} on our machine. Without activation, we obtain a spectra of the non-activated protein complex in the micelle (Fig. 5.14 a), which shows very wide and unspecific features due to the micelles. Upon activation between the ion funnel and the first quadrupole ion guide Q1, we were able to activate and release the protein AmtB of the micelle, obtaining a mass spectra of the trimeric peptide complex with very specific features. This is the first time, that we have employed our home-build time-of-flight mass spectrometer for a high m/z range of more than 6 kTh. As shown in Fig. 5.14b, the protein complex exists mainly in charge state +17. The successful replication of the sophisticated ESI of IMPs demonstrates how important the control of all parameters, such as voltages, energy and gas pressure, along the ion beam's path through the instrument is and can be seen as a first promising step towards the deposition and subsequent imaging of IMPs by LEEH.

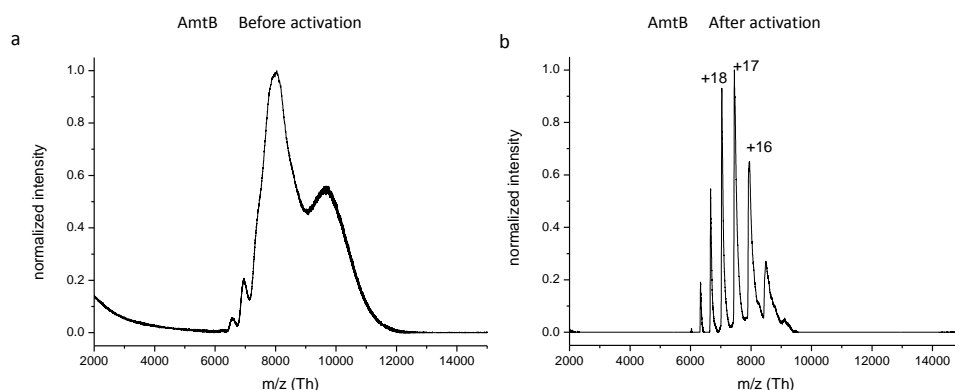


Figure 5.14: Native mass spectrometry of membrane protein AmtB (trimer = 126788 Da) enabled by micelle. **a:** Mass spectra before activation of the micelle shows very broad and unspecific features due to the different sizes and charge states of the micelle. **b:** After activation and release of the trimeric protein complex of AmtB, the mass spec shows very specific peaks corresponding to different charge states of the intact protein complex and reproduces the high-resolution native mass spectra²⁵¹.

The versatility of the ESI in combination with the appropriate imaging instrument gives wide access to methods resolving complex structures at the sub-molecular level. I have presented this approach for biopolymers, but it is, in principle, also applicable to artificial biopolymers and complex nanostructures opening the path for new materials as well as new methods to gain deeper understanding in structure and binding motifs on the atomic scale.

Bibliography

- [1] Nobelprize.org. The 2016 nobel prize in chemistry - press release. Press Release. URL http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2016/press.html.
- [2] Browne, W. R. & Feringa, B. L. Making molecular machines work. *Nat Nano* **1**, 25–35 (2006).
- [3] Eelkema, R. *et al.* Molecular machines: Nanomotor rotates microscale objects. *Nature* **440**, 163–163 (2006).
- [4] Kudernac, T. *et al.* Electrically driven directional motion of a four-wheeled molecule on a metal surface. *Nature* **479**, 208–211 (2011).
- [5] Weber, J. Structural biology: Toward the ATP synthase mechanism. *Nat. Chem. Biol.* **6**, 794–795 (2010).
- [6] Hess, H. & Bachand, G. D. Biomolecular motors. *Mater. Today* **8**, 22–29 (2005).
- [7] Kinbara, K. & Aida, T. Toward intelligent molecular machines: Directed motions of biological and artificial molecules and assemblies. *Chem. Rev.* **105**, 1377–1400 (2005).
- [8] Feynman, R. P. There's plenty of room at the bottom. *Engineering and science* **23**, 22–36 (1960).
- [9] Kühlbrandt, W. & Davies, K. M. Rotary ATPases: A new twist to an ancient machine. *Trends Biochem. Sci.* **41**, 106–116 (2016).
- [10] Petrey, D. *et al.* Template-based prediction of protein function. *Curr. Opin. Struct. Biol.* **32**, 33–38 (2015).
- [11] Berger, O. *et al.* Light-emitting self-assembled peptide nucleic acids exhibit both stacking interactions and Watson-Crick base pairing. *Nat. Nanotechnol.* **10**, 353–360 (2015).
- [12] Adler-Abramovich, L. & Gazit, E. The physical properties of supramolecular peptide assemblies: from building block association to technological applications. *Chem. Soc. Rev.* **43**, 6881–6893 (2014).
- [13] Caruthers, M. H. Deciphering the protein-DNA recognition code. *Acc. Chem. Res.* **13**, 155–160 (1980).

Bibliography

- [14] Zamenhof, S., Brawerman, G. & Chargaff, E. On the desoxypentose nucleic acids from several microorganisms. *Biochim. Biophys. Acta* **9**, 402–405 (1952).
- [15] Watson, J. D., Crick, F. H. *et al.* Molecular structure of nucleic acids. *Nature* **171**, 737–738 (1953).
- [16] Saiki, R. K. *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487 (1988).
- [17] Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99**, 237–247 (1982).
- [18] Seeman, N. C. DNA in a material world. *Nature* **421**, 427–431 (2003).
- [19] Marras, A. E., Zhou, L., Su, H.-J. & Castro, C. E. Programmable motion of DNA origami mechanisms **112**, 713–718 (2015).
- [20] Bath, J. & Turberfield, A. J. DNA nanomachines **2**, 275–284 (2007).
- [21] Nangreave, J., Han, D., Liu, Y. & Yan, H. DNA origami: a history and current perspective. *Curr. Opin. Chem. Biol.* **14**, 608–615 (2010).
- [22] Merrifield, R. B. Solid phase peptide synthesis. I. the synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
- [23] Merrifield, R. B., Stewart, J. M. & Jernberg, N. Instrument for automated synthesis of peptides. *Anal. Chem.* **38**, 1905–1914 (1966).
- [24] Fosgerau, K. & Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discovery Today* **20**, 122–128 (2015).
- [25] Matson, J. B. & Stupp, S. I. Self-assembling peptide scaffolds for regenerative medicine. *Chem. Commun.* **48**, 26–33 (2012).
- [26] Zhao, X. & Zhang, S. Designer self-assembling peptide materials. *Macromol. Biosci.* **7**, 13–22 (2007).
- [27] Frederix Pim W. J., M. *et al.* Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nat. Chem.* **7**, 30–37 (2015).
- [28] Gazit, E. Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev.* **36**, 1263–1269 (2007).
- [29] Ghaemmamghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
- [30] Schmidt, F. R. Recombinant expression systems in the pharmaceutical industry. *Appl. Microbiol. Biotechnol.* **65**, 363–372 (2004).
- [31] Whitford, D. *Proteins: Structure and Function* (Wiley, 2013).
- [32] Dobson, C. M. Protein folding and misfolding. *Nature* **426**, 884–890 (2003).

- [33] Slabinski, L., Jaroszewski, L., Rodrigues, A. P. C., Rychlewski, L., Wilson, I. A., Lesley, S. A. & Godzik, A. The challenge of protein structure determination—lessons from structural genomics. *Protein Sci.* **16**, 2472–2482 (2007).
- [34] Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193 (1995).
- [35] Callaway, E. The revolution will not be crystallized: a new method sweeps through structural biology. *Nature* **525**, 172–174 (2015).
- [36] Klemm, D., Heublein, B., Fink, H.-P. & Bohn, A. Cellulose: Fascinating biopolymer and sustainable raw material. *Angew. Chem. Int. Ed.* **44**, 3358–3393 (2005).
- [37] Parodi, A. J. Protein glucosylation and its role in protein folding. *Annu. Rev. Biochem.* **69**, 69–93 (2000).
- [38] Varki, A. *et al.* (eds.) *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, 2009), 2nd edn.
- [39] Hart, G. W. & Copeland, R. J. Glycomics hits the big time. *Cell* **143**, 672–676 (2010).
- [40] Robyt, J. F. *Essentials of Carbohydrate Chemistry*. Springer Advanced Texts in Chemistry (Springer New York, 2012).
- [41] Seeberger, P. H. The logic of automated glycan assembly. *Acc. Chem. Res.* **48**, 1450–1463 (2015).
- [42] Hofmann, J., Hahm, H. S., Seeberger, P. H. & Pagel, K. Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature* **526**, 241–244 (2015).
- [43] Plante, O. J., Palmacci, E. R. & Seeberger, P. H. Automated solid-phase synthesis of oligosaccharides. *Science* **291**, 1523 (2001).
- [44] Seeberger, P. H. Chemical glycobiology: why now? *Nat. Chem. Biol.* **5**, 368–372 (2009).
- [45] Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
- [46] Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D. & Alice, M. B. Molecular beams of macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968).
- [47] Heck, A. J. R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* **5**, 927–933 (2008).
- [48] Bohrer, B. C., Merenbloom, S. I., Koeniger, S. L., Hilderbrand, A. E. & Clemmer, D. E. Biomolecule analysis by ion mobility spectrometry. *Annual Rev. Anal. Chem.* **1**, 293–327 (2008).
- [49] Mitchell Wells, J. & McLuckey, S. A. Collision-induced dissociation (cid) of peptides and proteins. In *Biological Mass Spectrometry*, vol. 402, 148–185 (Academic Press, 2005).

Bibliography

- [50] Dongre, A. R., Somogyi, A. & Wysocki, V. H. Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *J. Mass Spectrom.* **31**, 339–350 (1996).
- [51] Bai, X.-C., McMullan, G. & Scheres, S. H. How cryo-em is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57 (2015).
- [52] Egerton, R. F., Li, P. & Malac, M. Radiation damage in the TEM and SEM. *Micron* **35**, 399–409 (2004).
- [53] Barth, J. V., Costantini, G. & Kern, K. Engineering atomic and molecular nanostructures at surfaces. *Nature* **437**, 671–679 (2005).
- [54] Barth, J. V. Molecular architectonic on metal surfaces. *Annu. Rev. Phys. Chem.* **58**, 375–407 (2007).
- [55] Schuler, B., Meyer, G., Peña, D., Mullins, O. C. & Gross, L. Unraveling the molecular structures of asphaltenes by atomic force microscopy. *J. Am. Chem. Soc.* **137**, 9870–9876 (2015).
- [56] Franchetti, V., Solka, B., Baitinger, W., Amy, J. & Cooks, R. Soft landing of ions as a means of surface modification. *Int. J. Mass Spectrom. Ion Phys.* **23**, 29 – 35 (1977).
- [57] Johnson, G. E., Hu, Q. & Laskin, J. Soft landing of complex molecules on surfaces. *Annual Rev. Anal. Chem.* **4**, 83–104 (2011).
- [58] Johnson, G. E., Gunaratne, D. & Laskin, J. Soft- and reactive landing of ions onto surfaces: Concepts and applications. *Mass Spectrom. Rev.* **35**, 439–479 (2016).
- [59] Rader, H. J., Rouhanipour, A., Talarico, A. M., Palermo, V., Samori, P. & Mullen, K. Processing of giant graphene molecules by soft-landing mass spectrometry. *Nat. Mater.* **5**, 276–280 (2006).
- [60] Kley, C. S. *et al.* Atomic-scale observation of multiconformational binding and energy level alignment of ruthenium-based photosensitizers on TiO₂ anatase. *Nano Lett.* **14**, 563–569 (2014).
- [61] Ouyang, Z. *et al.* Preparing protein microarrays by soft-landing of mass-selected ions. *Science* **301**, 1351 (2003).
- [62] Feng, B., Wunschel, D. S., Masselon, C. D., Pasa-Tolic, L. & Smith, R. D. Retrieval of DNA using soft-landing after mass analysis by ESI-FTICR for enzymatic manipulation. *J. Am. Chem. Soc.* **121**, 8961–8962 (1999).
- [63] Siuzdak, G., Bothner, B., Yeager, M., Brugidou, C., Fauquet, C. M., Hoey, K. & Change, C.-M. Mass spectrometry and viral analysis. *Chemistry & Biology* **3**, 45–48 (1996).
- [64] Clemmer, D. E., Hudgins, R. R. & Jarrold, M. F. Naked protein conformations: Cytochrome c in the gas phase. *J. Am. Chem. Soc.* **117**, 10141–10142 (1995).

- [65] Hu, P. & Loo, J. A. Gas-phase coordination properties of Zn^{2+} , Cu^{2+} , Ni^{2+} , and Co^{2+} with histidine-containing peptides. *J. Am. Chem. Soc.* **117**, 11314–11319 (1995).
- [66] Volný, M., Elam, W. T., Branca, A., Ratner, B. D. & Tureček, F. Preparative soft and reactive landing of multiply charged protein ions on a plasma-treated metal surface. *Anal. Chem.* **77**, 4890–4896 (2005).
- [67] Laskin, J., Wang, P., Hadjar, O., Futrell, J. H., Alvarez, J. & Cooks, R. G. Charge retention by peptide ions soft-landed onto self-assembled monolayer surfaces. *Int. J. Mass Spectrom.* **265**, 237–243 (2007).
- [68] Hu, Q., Wang, P. & Laskin, J. Effect of the surface on the secondary structure of soft landed peptide ions. *Phys. Chem. Chem. Phys.* **12**, 12802–12810 (2010).
- [69] Mikhailov, V. A., Mize, T. H., Benesch, J. L. P. & Robinson, C. V. Mass-selective soft-landing of protein assemblies with controlled landing energies. *Anal. Chem.* **86**, 8321–8328 (2014).
- [70] Rinke, G., Rauschenbach, S., Harnau, L., Albarghash, A., Pauly, M. & Kern, K. Active conformation control of unfolded proteins by hyperthermal collision with a metal surface. *Nano Lett.* **14**, 5609–5615 (2014).
- [71] Deng, Z., Thontasen, N., Malinowski, N., Rinke, G., Harnau, L., Rauschenbach, S. & Kern, K. A close look at proteins: Submolecular resolution of two- and three-dimensionally folded cytochrome c at surfaces. *Nano Lett.* **12**, 2452–2458 (2012).
- [72] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization—principles and practice. *Mass Spectrom. Rev.* **9**, 37–70 (1990).
- [73] Fenn, J. B., Mann, M., Chin-kai, M., Shek-fu, W. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64 (1989).
- [74] Fenn, J. B. Electrospray wings for molecular elephants (nobel lecture). *Angew. Chem. Int. Ed.* **42**, 3871–3894 (2003).
- [75] Wilm, M. & Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **68**, 1–8 (1996).
- [76] Wilm, M. S. & Mann, M. Electrospray and Taylor-cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Processes* **136**, 167–180 (1994).
- [77] Zeleny, J. The electrical discharge from liquid points, and a hydrostatic method of measuring the electric intensity at their surfaces. *Phys. Rev.* **3**, 69–91 (1914).
- [78] Taylor, G. Disintegration of water drops in an electric field. In *Proc. R. Soc. London A*, vol. 280, 383–397 (The Royal Society, 1964).
- [79] Fernández de la Mora, J. The fluid dynamics of Taylor cones. *Annu. Rev. Fluid Mech.* **39**, 217–243 (2006).

Bibliography

- [80] Cole, R. B. *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation, and Applications* (Wiley, 1997).
- [81] Nemes, P., Marginean, I. & Vertes, A. Spraying mode effect on droplet formation and ion chemistry in electrosprays. *Anal. Chem.* **79**, 3105–3116 (2007).
- [82] Gross, J. & Roepstorff, P. *Mass Spectrometry: A Textbook* (Springer Berlin Heidelberg, 2011).
- [83] Rayleigh, L. F. XX. On the equilibrium of liquid conducting masses charged with electricity. *Philosophical Magazine Series 5* **14**, 184–186 (1882).
- [84] Cole, R. B. Some tenets pertaining to electrospray ionization mass spectrometry. *Journals of Mass Spectrometry* **35**, 763–772 (2000).
- [85] Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the mechanism of electrospray ionization. *Anal. Chem.* **85**, 2–9 (2013).
- [86] Kebarle, P. & Verkerk, U. H. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898–917 (2009).
- [87] Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics* **64**, 2287–2294 (1976).
- [88] Ahadi, E. & Konermann, L. Modeling the behavior of coarse-grained polymer chains in charged water droplets: Implications for the mechanism of electrospray ionization. *J. Phys. Chem. B* **116**, 104–112 (2012).
- [89] Rauschenbach, S., Stadler, F. L., Lunedei, E., Malinowski, N., Koltsov, S., Costantini, G. & Kern, K. Electrospray ion beam deposition of clusters and biomolecules. *Small* **2**, 540–547 (2006).
- [90] Rauschenbach, S., Ternes, M., Harnau, L. & Kern, K. Mass spectrometry as a preparative tool for the surface science of large molecules. *Annual Rev. Anal. Chem.* **9**, 473–498 (2016).
- [91] Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.* **49**, 57–61 (1982).
- [92] Binnig, G. & Rohrer, H. Scanning tunneling microscopy—from birth to adolescence (nobel lecture). *Angew. Chem. Int. Ed.* **26**, 606–614 (1987).
- [93] Chen, C. *Introduction to Scanning Tunneling Microscopy*. Oxford Series in Optical and Imaging Sciences (Oxford University Press, 1993).
- [94] Meyer, E., Hug, H. & Bennewitz, R. *Scanning Probe Microscopy: The Lab on a Tip*. Advanced Texts in Physics (Springer Berlin Heidelberg, 2013).
- [95] Giaever, I. Electron tunneling between two superconductors. *Phys. Rev. Lett.* **5**, 464 (1960).

- [96] Bardeen, J. Tunnelling from a many-particle point of view. *Phys. Rev. Lett.* **6**, 57–59 (1961).
- [97] Tersoff, J. & Hamann, D. R. Theory and application for the scanning tunneling microscope. *Phys. Rev. Lett.* **50**, 1998–2001 (1983).
- [98] Binnig, G., Rohrer, H., Gerber, C. & Weibel, E. Tunneling through a controllable vacuum gap. *Appl. Phys. Lett.* **40**, 178–180 (1982).
- [99] Tersoff, J. & Hamann, D. R. Theory of the scanning tunneling microscope. *Phys. Rev. B* **31**, 805–813 (1985).
- [100] Scientaomicron. Fermi SPM. URL <http://www.scientaomicron.com/en/products/fermi-spm/instrument-concept>.
- [101] Miller, P. E. & Denton, M. B. The quadrupole mass filter: Basic operating concepts. *J. Chem. Educ.* **63**, 617 (1986).
- [102] Douglas, D. J. Linear quadrupoles in mass spectrometry. *Mass Spectrom. Rev.* **28**, 937–960 (2009).
- [103] Wiley, W. & McLaren, I. H. Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* **26**, 1150–1157 (1955).
- [104] Dubey, G., Urcuyo, R., Abb, S., Rinke, G., Burghard, M., Rauschenbach, S. & Kern, K. Chemical modification of graphene via hyperthermal molecular reaction. *J. Am. Chem. Soc.* **136**, 13482–13485 (2014).
- [105] Rauschenbach, S. *et al.* Crystalline inverted membranes grown on surfaces by electro-spray ion beam deposition in vacuum. *Adv. Mater.* **24**, 2761–2767 (2012).
- [106] Horcas, I., Fernández, R., Gomez-Rodriguez, J., Colchero, J., Gómez-Herrero, J. & Baro, A. WSXM: a software for scanning probe microscopy and a tool for nanotechnology. *Rev. Sci. Instrum.* **78**, 013705 (2007).
- [107] Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.* **37**, 205–211 (1951).
- [108] Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
- [109] Lovell, S. C. *et al.* Structure validation by $C\alpha$ geometry: ϕ , ψ and $C\beta$ deviation. *Proteins* **50**, 437–450 (2003).
- [110] Baldwin, J. Structure and cooperativity of haemoglobin. *Trends Biochem. Sci.* **5**, 224–228 (1980).
- [111] Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).

Bibliography

- [112] Brown, R. M. Cellulose structure and biosynthesis: What is in store for the 21st century? *J. Polym. Sci. A Polym. Chem.* **42**, 487–495 (2004).
- [113] Sugiyama, J., Vuong, R. & Chanzy, H. Electron diffraction study on the two crystalline phases occurring in native cellulose from an algal cell wall. *Macromolecules* **24**, 4168–4175 (1991).
- [114] Kim, J., Yun, S. & Ounaies, Z. Discovery of cellulose as a smart material. *Macromolecules* **39**, 4202–4206 (2006).
- [115] Dwek, R. A. Glycobiology: Toward understanding the function of sugars. *Chem. Rev.* **96**, 683–720 (1996).
- [116] Slawson, C. & Hart, G. W. O-GlcNAc signalling: implications for cancer cell biology. *Nat. Rev. Cancer* **11**, 678–684 (2011).
- [117] Dall’Olio, F., Malagolini, N., Trinchera, M. & Chiricolo, M. Mechanisms of cancer-associated glycosylation changes. *Front. Biosci.* **17**, 670 (2012).
- [118] Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A. & Dwek, R. A. Glycosylation and the immune system. *Science* **291**, 2370 (2001).
- [119] Nagase, H., Ogawa, N., Endo, T., Shiro, M., Ueda, H. & Sakurai, M. Crystal structure of an anhydrous form of trehalose: Structure of water channels of trehalose polymorphism. *The Journal of Physical Chemistry B* **112**, 9105–9111 (2008).
- [120] Sussich, F., Skopec, C., Brady, J. & Cesàro, A. Reversible dehydration of trehalose and anhydrobiosis: from solution state to an exotic crystal? *Carbohydr. Res.* **334**, 165–176 (2001).
- [121] Brown, G. M., Rohrer, D. C., Berking, B., Beevers, C. A., Gould, R. O. & Simpson, R. The crystal structure of α,α -trehalose dihydrate from three independent X-ray determinations. *Acta Crystallogr Sect B* **28**, 3145–3158 (1972).
- [122] Brown, G. M. & Levy, H. A. Further refinement of the structure of sucrose based on neutron-diffraction data. *Acta Crystallogr Sect B* **29**, 790–797 (1973).
- [123] Jeffrey, G. A. & Takagi, S. Hydrogen-bond structure in carbohydrate crystals. *Acc. Chem. Res.* **11**, 264–270 (1978).
- [124] Duus, J. Ø., Gotfredsen, C. H. & Bock, K. Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem. Rev.* **100**, 4589–4614 (2000).
- [125] Dell, A. & Morris, H. R. Glycoprotein structure determination by mass spectrometry. *Science* **291**, 2351–2356 (2001).
- [126] Daikoku, S., Widmalm, G. & Kanie, O. Analysis of a series of isomeric oligosaccharides by energy-resolved mass spectrometry: a challenge on homobranched trisaccharides. *Rapid Commun. Mass Spectrom.* **23**, 3713–3719 (2009).

- [127] Yamagaki, T. & Sato, A. Isomeric oligosaccharides analyses using negative-ion electrospray ionization ion mobility spectrometry combined with collision-induced dissociation MS/MS. *Anal. Sci.* **25**, 985–988 (2009).
- [128] Baker, A., Helbert, W., Sugiyama, J. & Miles, M. New insight into cellulose structure by atomic force microscopy shows the I α crystal phase at near-atomic resolution. *Biophys. J.* **79**, 1139 – 1145 (2000).
- [129] Fletcher, M. & Floodgate, G. An electron-microscopic demonstration of an acidic polysaccharide involved in the adhesion of a marine bacterium to solid surfaces. *Microbiology* **74**, 325–334 (1973).
- [130] Nicolson, G. L. & Singer, S. Ferritin-conjugated plant agglutinins as specific saccharide stains for electron microscopy: application to saccharides bound to cell membranes. *Proc. Natl. Acad. Sci.* **68**, 942–945 (1971).
- [131] Kunitake, M., Uemura, S., Ito, O., Fujiwara, K., Murata, Y. & Komatsu, K. Structural analysis of C60 trimers by direct observation with scanning tunneling microscopy. *Angew. Chem. Int. Ed.* **41**, 969–972 (2002).
- [132] Alemani, M., Peters, M. V., Hecht, S., Rieder, K.-H., Moresco, F. & Grill, L. Electric field-induced isomerization of azobenzene by stm. *J. Am. Chem. Soc.* **128**, 14446–14447 (2006).
- [133] Schendel, V. *et al.* Remotely controlled isomer selective molecular switching. *Nano Lett.* **16**, 93–97 (2016).
- [134] Garozzo, D., Impallomeni, G., Spina, E., Green, B. N. & Hutton, T. Linkage analysis in disaccharides by electrospray mass spectrometry. *Carbohydr. Res.* **221**, 253 – 257 (1991).
- [135] Takáts, Z., Wiseman, J. M., Gologan, B. & Cooks, R. G. Electrosonic spray ionization. A gentle technique for generating folded proteins and protein complexes in the gas phase and for studying ion-molecule reactions at atmospheric pressure. *Anal. Chem.* **76**, 4050–4058 (2004).
- [136] Zhang, Z. Atomistic processes in the early stages of thin-film growth. *Science* **276**, 377–383 (1997).
- [137] van der Spoel, D., Marklund, E. G., Larsson, D. S. D. & Caleman, C. Proteins, lipids, and water in the gas phase. *Macromol. Biosci.* **11**, 50–59 (2011).
- [138] Oostenbrink, C., Villa, A., Mark, A. E. & Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676 (2004).
- [139] Schlickum, U. *et al.* Chiral Kagomé lattice from simple ditopic molecular bricks. *J. Am. Chem. Soc.* **130**, 11778–11782 (2008).
- [140] Lin, N., Dmitriev, A., Weckesser, J., Barth, J. V. & Kern, K. Real-time single-molecule imaging of the formation and dynamics of coordination compounds. *Angewandte Chemie* **114**, 4973–4977 (2002).

Bibliography

- [141] Lin, N., Payer, D., Dmitriev, A., Strunskus, T., Wöll, C., Barth, J. V. & Kern, K. Two-dimensional adatom gas bestowing dynamic heterogeneity on surfaces. *Angew. Chem. Int. Ed.* **44**, 1488–1491 (2005).
- [142] Fernandez-Torrente, I., Monturet, S., Franke, K. J., Fraxedas, J., Lorente, N. & Pascual, J. I. Long-range repulsive interaction between molecules on a metal surface induced by charge transfer. *Phys. Rev. Lett.* **99**, 176103 (2007).
- [143] Tiwari, R. K., Otálvaro, D. M., Joachim, C. & Saeys, M. Origin of the contrast inversion in the STM image of CO on Cu(1 1 1). *Surf. Sci.* **603**, 3286–3291 (2009).
- [144] Vitali, L., Levita, G., Ohmann, R., Comisso, A., De Vita, A. & Kern, K. Portrait of the potential barrier at metal-organic nanocontacts. *Nat. Mater.* **9**, 320–323 (2010).
- [145] Schön, J. C., Oligschleger, C. & Cortes, J. Prediction and clarification of structures of (bio)molecules on surfaces. *Zeitschrift für Naturforschung B* **71** (2016).
- [146] Abb, S., Harnau, L., Gutzler, R., Rauschenbach, S. & Kern, K. Two-dimensional honeycomb network through sequence-controlled self-assembly of oligopeptides. *Nat. Commun.* **7**, 10335 (2016).
- [147] Gradišar, H. *et al.* Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat. Chem. Biol.* **9**, 362–366 (2013).
- [148] Kim, S., Kim, J. H., Lee, J. S. & Park, C. B. Beta-sheet-forming, self-assembled peptide nanomaterials towards optical, energy, and healthcare applications. *Small* **11**, 3623–3640 (2015).
- [149] Liu, L., Busuttill, K., Zhang, S., Yang, Y., Wang, C., Besenbacher, F. & Dong, M. The role of self-assembling polypeptides in building nanomaterials. *Phys. Chem. Chem. Phys.* **13**, 17435–17444 (2011).
- [150] Zhang, S. Fabrication of novel biomaterials through molecular self-assembly. *Nat Biotech* **21**, 1171–1178 (2003).
- [151] Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324–327 (1993).
- [152] Zhang, S., Holmes, T., Lockshin, C. & Rich, A. Spontaneous assembly of a self-complementary oligopeptide to form a stable macroscopic membrane. *Proc. Natl. Acad. Sci.* **90**, 3334–3338 (1993).
- [153] Altman, M., Lee, P., Rich, A. & Zhang, S. Conformational behavior of ionic self-complementary peptides. *Protein Sci.* **9**, 1095–1105 (2000).
- [154] Hartgerink, J. D., Beniash, E. & Stupp, S. I. Self-assembly and mineralization of peptide-amphiphile nanofibers. *Science* **294**, 1684 (2001).

- [155] Vauthey, S., Santoso, S., Gong, H., Watson, N. & Zhang, S. Molecular self-assembly of surfactant-like peptides to form nanotubes and novovesicles. *Proc. Natl. Acad. Sci.* **99**, 5355–5360 (2002).
- [156] Görbitz, C. H. Nanotube formation by hydrophobic dipeptides. *Chemistry – A European Journal* **7**, 5153–5159 (2001).
- [157] Reches, M. & Gazit, E. Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **300**, 625 (2003).
- [158] Adler-Abramovich, L. *et al.* Self-assembled arrays of peptide nanotubes by vapour deposition. *Nat. Nanotechnol.* **4**, 849–854 (2009).
- [159] Görbitz, C. H. The structure of nanotubes formed by diphenylalanine, the core recognition motif of alzheimer's β -amyloid polypeptide. *Chem. Commun.* 2332–2334 (2006).
- [160] Azuri, I., Adler-Abramovich, L., Gazit, E., Hod, O. & Kronik, L. Why are diphenylalanine-based peptide nanostructures so rigid? Insights from first principles calculations. *J. Am. Chem. Soc.* **136**, 963–969 (2014).
- [161] Lingenfelder, M., Tomba, G., Costantini, G., Colombi Ciacchi, L., De Vita, A. & Kern, K. Tracking the chiral recognition of adsorbed dipeptides at the single-molecule level. *Angew. Chem. Int. Ed.* **46**, 4492–4495 (2007).
- [162] Claridge, S. A., Thomas, J. C., Silverman, M. A., Schwartz, J. J., Yang, Y., Wang, C. & Weiss, P. S. Differentiating amino acid residues and side chain orientations in peptides using scanning tunneling microscopy. *J. Am. Chem. Soc.* **135**, 18528–18535 (2013).
- [163] Kalashnyk, N. *et al.* Scanning tunneling microscopy reveals single-molecule insights into the self-assembly of amyloid fibrils. *ACS Nano* **6**, 6882–6889 (2012).
- [164] Rauschenbach, S. *et al.* Electrospray ion beam deposition: Soft-landing and fragmentation of functional molecules at solid surfaces. *ACS Nano* **3**, 2901–2910 (2009).
- [165] Rauschenbach, S. *et al.* Two-dimensional folding of polypeptides into molecular nanostructures. *ACS Nano* (2017).
- [166] Kaminski, G., Duffy, E. M., Matsui, T. & Jorgensen, W. L. Free energies of hydration and pure liquid properties of hydrocarbons from the OPLS All-Atom model. *J. Phys. Chem.* **98**, 13077–13082 (1994).
- [167] Iori, F., Di Felice, R., Molinari, E. & Corni, S. GoIP: An atomistic force-field to describe the interaction of proteins with Au(111) surfaces in water. *J. Comput. Chem.* **30**, 1465–1476 (2009).
- [168] Neese, F. The ORCA program system. *WIREs Comput Mol Sci* **2**, 73–78 (2012).
- [169] Banhart, F., Kotakoski, J. & Krashenninnikov, A. V. Structural defects in graphene. *ACS Nano* **5**, 26–41 (2011).

Bibliography

- [170] Dinsmore, A. D., Hsu, M. F., Nikolaides, M. G., Marquez, M., Bausch, A. R. & Weitz, D. A. Colloidosomes: Selectively permeable capsules composed of colloidal particles. *Science* **298**, 1006 (2002).
- [171] Feng, J., Pandey, R. B., Berry, R. J., Farmer, B. L., Naik, R. R. & Heinz, H. Adsorption mechanism of single amino acid and surfactant molecules to Au(111) surfaces in aqueous solution: design rules for metal-binding molecules. *Soft Matter* **7**, 2113–2120 (2011).
- [172] Hoeffling, M., Iori, F., Corni, S. & Gottschalk, K.-E. Interaction of amino acids with the Au(111) surface: Adsorption free energies from molecular dynamics simulations. *Langmuir* **26**, 8347–8351 (2010).
- [173] Hoeffling, M., Iori, F., Corni, S. & Gottschalk, K.-E. The conformations of amino acids on a gold(111) surface. *Chem. Phys. Chem.* **11**, 1763–1767 (2010).
- [174] Liu, L., Zhang, L., Mao, X., Niu, L., Yang, Y. & Wang, C. Chaperon-mediated single molecular approach toward modulating a β peptide aggregation. *Nano Lett.* **9**, 4066–4072 (2009).
- [175] Kołaczkiwicz, J. & Bauer, E. Clausius-clapeyron equation analysis of two- dimensional vaporization. *Surf. Sci.* **155**, 700–714 (1985).
- [176] Heinz, H., Lin, T.-J., Kishore Mishra, R. & Emami, F. S. Thermodynamically consistent force fields for the assembly of inorganic, organic, and biological nanostructures: The interface force field. *Langmuir* **29**, 1754–1765 (2013).
- [177] Lippard, S. J. & Berg, J. *Principles of Bioinorganic Chemistry*. Principles of Bioinorganic Chemistry (University Science Books, 1994).
- [178] Gomes, C. & Wittung-Stafshede, P. *Protein Folding and Metal Ions: Mechanisms, Biology and Disease* (CRC Press, 2016).
- [179] Gray, H. B. Biological inorganic chemistry at the beginning of the 21st century. *Proc. Natl. Acad. Sci.* **100**, 3563–3568 (2003).
- [180] Liu, D., Seuthe, A. B., Ehrler, O. T., Zhang, X., Wyttenbach, T., Hsu, J. F. & Bowers, M. T. Oxytocin-receptor binding: Why divalent metals are essential. *J. Am. Chem. Soc.* **127**, 2024–2025 (2005).
- [181] Frankel, A. D., Berg, J. M. & Pabo, C. O. Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc. Natl. Acad. Sci.* **84**, 4841–4845 (1987).
- [182] Migliorini, C., Porciatti, E., Luczkowski, M. & Valensin, D. Structural characterization of Cu^{2+} , Ni^{2+} and Zn^{2+} binding sites of model peptides associated with neurodegenerative diseases. *Coord. Chem. Rev.* **256**, 352–368 (2012).
- [183] Savellieff, M. G., Lee, S., Liu, Y. & Lim, M. H. Untangling amyloid- β , tau, and metals in alzheimer’s disease. *ACS Chem. Biol.* **8**, 856–865 (2013).
- [184] Bush, A. I. The metallobiology of alzheimer’s disease. *Trends Neurosci.* **26**, 207 – 214 (2003).

- [185] Barth, J. V., Weckesser, J., Lin, N., Dmitriev, A. & Kern, K. Supramolecular architectures and nanostructures at metal surfaces. *Appl. Phys. A* **76**, 645–652 (2003).
- [186] Grumelli, D., Wurster, B., Stepanow, S. & Kern, K. Bio-inspired nanocatalysts for the oxygen reduction reaction. *Nat. Commun.* **4**, 2904 (2013).
- [187] Čechal, J. *et al.* CO₂ binding and induced structural collapse of a surface-supported metal-organic network. *J. Phys. Chem. C* **120**, 18622–18630 (2016).
- [188] Skomski, D., Tempas, C. D., Cook, B. J., Polezhaev, A. V., Smith, K. A., Caulton, K. G. & Tait, S. L. Two- and three-electron oxidation of single-site vanadium centers at surfaces by ligand design. *J. Am. Chem. Soc.* **137**, 7898–7902 (2015).
- [189] Manton, A., Massüger, L., Rabu, P., Palivan, C., McCusker, L. B. & Taubert, A. Metal-peptide frameworks (mpfs): “bioinspired” metal organic frameworks. *J. Am. Chem. Soc.* **130**, 2517–2526 (2008).
- [190] Schaechtelin, G., Walter, R., Salomon, H., Jelinek, J., Karen, P. & Cort, J. H. Enhancement of the activity of angiotensin II by certain cations. *Mol. Pharmacol.* **10**, 57 (1974).
- [191] Craig, L. C., Harfenist, E. J. & Paladini, A. C. Dialysis studies. VII. the behavior of angiotensin, oxytocin, vasopressin, and some of their analogs*. *Biochemistry (Mosc.)* **3**, 764–769 (1964).
- [192] Spyroulias, G. A., Nikolakopoulou, P., Tzakos, A., Gerothanassis, I. P., Magafa, V., Manessi-Zoupa, E. & Cordopatis, P. Comparison of the solution structures of angiotensin I & II. *Eur. J. Biochem.* **270**, 2163–2173 (2003).
- [193] Kim, J. Y., Kim, M. J. & Kim, H. T. Determination of a binding site of Cu, Ni, Mg, and Ca metal ions with angiotensin II peptide by electrospray tandem mass spectrometry. *Bull. Korean Chem. Soc.* **31**, 1377–1380 (2010).
- [194] Kim, B.-R. & Kim, H.-T. Dft study of the [Cu-Angiotensin II]₂⁺ complex structure. *Bull. Korean Chem. Soc.* **33**, 2407–2410 (2012).
- [195] Prudent, M. & Girault, H. H. On-line electrogeneration of copper-peptide complexes in microspray mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **19**, 560–568 (2008).
- [196] Reverend, B. D.-L., Liman, F., Livera, C., Pettit, L. D., Pyburn, S. & Kozłowski, H. A potentiometric and spectroscopic study of the interaction of angiotensin II and some of its peptide fragments with copper(II). *J. Chem. Soc., Dalton Trans.* 887–894 (1988).
- [197] Bal, W., Jezowska-Bojczuk, M., Kozłowski, H., Chruscinski, L., Kupryszewski, G. & Witczuk, B. Cu(II) binding by angiotensin II fragments: Asp-Arg-Val-Tyr-Ile-His and Arg-Val-Tyr-Ile-His. competition between amino group and imidazole nitrogens in anchoring of metal ions. *J. Inorg. Biochem.* **57**, 235–247 (1995).
- [198] Pettit, L. D., Pyburn, S., Kozłowski, H., Decock-Le Reverend, B. & Liman, F. Co-ordination of nickel(II) ions by angiotensin II and its peptide fragments. a potentiometric, proton nuclear magnetic resonance and circular dichroism spectroscopic study. *J. Chem. Soc., Dalton Trans.* 1471–1475 (1989).

Bibliography

- [199] 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. *J. Cheminf.* **4**, 1–17 (2012).
- [200] Ohmann, R., Levita, G., Vitali, L., De Vita, A. & Kern, K. Influence of subsurface layers on the adsorption of large organic molecules on close-packed metal surfaces. *ACS Nano* **5**, 1360–1365 (2011).
- [201] Wang, S. C. & Ehrlich, G. Self-adsorption sites on a close-packed surface: Ir on Ir(111). *Phys. Rev. Lett.* **62**, 2297–2300 (1989).
- [202] Rinke, G. *Electrospray Ion Beam Deposition of Complex Non-Volatile Molecules*. Ph.D. thesis, École Polytechnique Fédérale de Lausanne (2013). URL <https://infoscience.epfl.ch/record/188400>.
- [203] Longchamp, J.-N., Rauschenbach, S., Abb, S., Escher, C., Latychevskaia, T., Kern, K. & Fink, H.-W. Imaging proteins at the single-molecule level. *Proceedings of the National Academy of Sciences* (2017).
- [204] Soto, C. Protein misfolding and disease; protein refolding and therapy. *FEBS Lett.* **498**, 204–207 (2001).
- [205] Drenth, J. *Principles of Protein X-Ray Crystallography*. SpringerLink: Springer e-Books (Springer New York, 2007).
- [206] George, A. & Wilson, W. W. Predicting protein crystallization from a dilute solution property. *Acta Crystallographica Section D* **50**, 361–365 (1994).
- [207] Rose, H. Correction of aperture aberrations in magnetic systems with threefold symmetry. *Nucl. Instrum. Methods Phys. Res.* **187**, 187–199 (1981).
- [208] Haider, M., Rose, H., Uhlemann, S., Kabius, B. & Urban, K. Towards 0.1 nm resolution with the first spherically corrected transmission electron microscope. *J. Electron Microsc.* (Tokyo) **47**, 395–405 (1998).
- [209] Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative staining and image classification — powerful tools in modern electron microscopy. *Biological Procedures Online* **6**, 23–34 (2004).
- [210] Cheng, Y., Grigorieff, N., Penczek, P. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438 – 449 (2015).
- [211] van Heel, M. *et al.* Single-particle electron cryo-microscopy: towards atomic resolution. *Q. Rev. Biophys.* **33**, 307–369 (2000).
- [212] Henderson, R. Realizing the potential of electron cryo-microscopy. *Q. Rev. Biophys.* **37**, 3–13 (2004).
- [213] Chamberlain, T. W., Biskupek, J., Skowron, S. T., Bayliss, P. A., Bichoutskaia, E., Kaiser, U. & Khlobystov, A. N. Transmission electron microscopy: Isotope substitution extends the lifetime of organic molecules in transmission electron microscopy (small 5/2015). *Small* **11**, 510–510 (2015).

-
- [214] Milazzo, A.-C. *et al.* Initial evaluation of a direct detection device detector for single particle cryo-electron microscopy. *J. Struct. Biol.* **176**, 404–408 (2011).
- [215] Hohn, M. *et al.* SPARX, a new environment for cryo-EM image processing. *J. Struct. Biol.* **157**, 47–55 (2007).
- [216] Fischer, N., Neumann, P., Konevega, A. L., Bock, L. V., Ficner, R., Rodnina, M. V. & Stark, H. Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520**, 567–570 (2015).
- [217] Cho, A. What shall we do with the X-ray laser? *Science* **330**, 1470–1471 (2010).
- [218] Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77 (2011).
- [219] Miao, J., Chapman, H. N., Kirz, J., Sayre, D. & Hodgson, K. O. Taking X-Ray diffraction to the limit: Macromolecular structures from femtosecond X-Ray pulses and diffraction microscopy of cells with synchrotron radiation. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 157–176 (2004).
- [220] Germann, M., Latychevskaia, T., Escher, C. & Fink, H.-W. Nondestructive imaging of individual biomolecules. *Phys. Rev. Lett.* **104**, 095501 (2010).
- [221] Longchamp, J.-N., Latychevskaia, T., Escher, C. & Fink, H.-W. Non-destructive imaging of an individual protein. *Appl. Phys. Lett.* **101** (2012).
- [222] Gabor, D. *et al.* A new microscopic principle. *Nature* **161**, 777–778 (1948).
- [223] Latychevskaia, T., Longchamp, J.-N., Escher, C. & Fink, H.-W. Holography and coherent diffraction with low-energy electrons: A route towards structural biology at the single molecule level. *Ultramicroscopy* **159**, 395–402 (2015).
- [224] Gabor, D. Microscopy by reconstructed wave-fronts. *Proc. R. Soc. London A* **197**, 454–487 (1949).
- [225] Fink, H.-W., Stocker, W. & Schmid, H. Holography with low-energy electrons. *Phys. Rev. Lett.* **65**, 1204–1206 (1990).
- [226] Fink, H., Stocker, W. & Schmid, H. Coherent point source electron beams. *Journal of Vacuum Science & Technology B* **8**, 1323–1324 (1990).
- [227] Nair, R. R. *et al.* Graphene as a transparent conductive support for studying biological molecules by transmission electron microscopy. *Appl. Phys. Lett.* **97** (2010).
- [228] Longchamp, J.-N., Latychevskaia, T., Escher, C. & Fink, H.-W. Low-energy electron transmission imaging of clusters on free-standing graphene. *Appl. Phys. Lett.* **101** (2012).
- [229] Kreuzer, H. J., Nakamura, K., Wierzbicki, A., Fink, H.-W. & Schmid, H. Theory of the point source electron microscope. *Ultramicroscopy* **45**, 381–403 (1992).

Bibliography

- [230] Latychevskaia, T. & Fink, H.-W. Practical algorithms for simulation and reconstruction of digital in-line holograms. *Appl. Opt.* **54**, 2424–2434 (2015).
- [231] Morin, R. & Fink, H. Highly monochromatic electron point-source beams. *Appl. Phys. Lett.* **65**, 2362–2364 (1994).
- [232] Germann, M., Latychevskaia, T., Escher, C. & Fink, H.-W. Pulsed electron holography. *Appl. Phys. Lett.* **102**, 203115 (2013).
- [233] Abbe, E. VII.–On the estimation of aperture in the microscope. *J. R. Microsc. Soc.* **1**, 388–423 (1881).
- [234] Abbe. The relation of aperture and power in the microscope (continued).*. *J. R. Microsc. Soc.* **2**, 460–473 (1882).
- [235] Mutus, J. Y. *et al.* Nanoscale structuring of tungsten tip yields most coherent electron point-source. *New J. Phys.* **15**, 073038 (2013).
- [236] Smith, S. W. *The Scientist and Engineer's Guide to Digital Signal Processing* (California Technical Pub., 1997).
- [237] Longchamp, J.-N., Escher, C. & Fink, H.-W. Ultraclean freestanding graphene by platinum-metal catalysis. *Journal of Vacuum Science & Technology B* **31** (2013).
- [238] Ruotolo, B. T. & Robinson, C. V. Aspects of native proteins are retained in vacuum. *Curr. Opin. Chem. Biol.* **10**, 402–408 (2006).
- [239] Wolynes, P. G. Biomolecular folding in vacuo!!!(?) **92**, 2426–2427 (1995).
- [240] Barrera, N. P., Di Bartolo, N., Booth, P. J. & Robinson, C. V. Micelles protect membrane complexes from solution to vacuum. *Science* **321**, 243 (2008).
- [241] Wang, P. & Laskin, J. Helical peptide arrays on self-assembled monolayer surfaces through soft and reactive landing of mass-selected ions. *Angew. Chem. Int. Ed.* **47**, 6678–6680 (2008).
- [242] Carter, D. C. & He, X. M. Structure of human serum albumin. *Science* **249**, 302 (1990).
- [243] Kang, Y., Terrier, P. & Douglas, D. J. Mass spectra and ion collision cross sections of hemoglobin. *J. Am. Soc. Mass. Spectrom.* **22**, 290–299 (2011).
- [244] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. Ucsf chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- [245] Babine, R. E. & Bender, S. L. Molecular recognition of protein-ligand complexes: Applications to drug design. *Chem. Rev.* **97**, 1359–1472 (1997).
- [246] Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V. & Ruotolo, B. T. Collision cross sections of proteins and their complexes: A calibration framework and database for gas-phase structural biology. *Anal. Chem.* **82**, 9557–9565 (2010).

- [247] Bleiholder, C., Dupuis, N. F., Wyttenbach, T. & Bowers, M. T. Ion mobility-mass spectrometry reveals a conformational conversion from random assembly to β -sheet in amyloid fibril formation. *Nat. Chem.* **3**, 172–177 (2011).
- [248] Salentin, S., Haupt, V. J., Daminelli, S. & Schroeder, M. Polypharmacology rescored: Protein–ligand interaction profiles for remote binding site similarity assessment. *Prog. Biophys. Mol. Biol.* **116**, 174–186 (2014).
- [249] Caffrey, M. Membrane protein crystallization. *J. Struct. Biol.* **142**, 108–132 (2003).
- [250] Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protocols* **8**, 639–651 (2013).
- [251] Gault, J. *et al.* High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nat. Methods* **13**, 333–336 (2016).

Acknowledgements

If you give a man a fish, he is hungry again in an hour. If you teach him to catch a fish, you do him a good turn. – Anne Isabella Thackeray Ritchie

After the 4 years of my thesis, there are quite some people I would like to thank as they have taught me a lot and supported me through all ups and downs of my PhD.

First, I would like to thank Prof. Klaus Kern for giving me the opportunity to do my doctoral studies in his group at the Max Planck Institute for Solid State Research. I am thankful to have worked on such an enthralling subject which I could present and discuss on several conferences. I really enjoyed the freedom to explore my own ideas and the trust you put in me. I also thank the members of my jury committee, Prof. Michele Ceriotti, Prof. Giovanni Costantini and Prof. Florian Klappenberger for their time.

A special thanks belongs to my daily supervisor, Dr. Stephan Rauschenbach, who build the machine and always knew how to tame the beast. Together with Dr. Gordon Rinke, my predecessor, he introduced me to the machine and showed me all the little tricks. Thank you, Stephan, for all the stories, discussions and advice - scientific as well as personal. You gave me lots of room to explore, but were always there when necessary.

Moreover, I would like to thank the whole ESI team, Dr. Gordon Rinke, Dr. Girjesh Dubey, Elise Duquesne, Ahmed ben Falleh, Dominik Lorocho and Dr. Suman Sen who shared some time at the instrument with me. Especially the time with Elise working on some peptide-metal complexes was very exciting.

Only with the help of our collaborators are we able to unravel the complex behavior of biomolecules on surfaces in detail. Dr. Ludger Harnau was a big help employing MD to understand the peptide assemblies on surfaces as well as the saccharides. Thank you for the many constructive discussions and emails! Dr. Rico Gutzler as well as Prof. Talat Rahman and Dr. Duy Lee have performed the DFT calculation and simulated STM images, so that we are able to compare the developed models with our actual data. Thank you! Another thank you goes to Dr. Juan Cortes and Prof. Christian Schön, who developed new algorithms to explore the adsorption behavior of disaccharides. Dr. Uta Schlickum, Dr. Soon Jung Jung,

Bibliography

Sebastian Koslowski, Tomasz Michnowicz and Christian Dette all kindly introduced me to their low temperature STM instruments and let me measure some of the data presented in this thesis. From Ester Martin and Joseph Gault, I have learned a lot about native MS and the right handling of protein solutions.

Without our mechanical engineers Arthur Küster, Wolfgang Stiepany, Peter Anders and Marko Memmler, things would not have been as smooth as they were, especially during the exchange of the STM and the design of new sample holders. In administration, this part was taken over by our secretary Sabine Birtel who always had things under control (and chocolate). *Thank you.*

I have really enjoyed working in the department thanks to the encouraging environment created by all group members: Dr. Markus Ternes, for discussions about the electronics and the set-up with our STM; the tea-sessions with Sebastian Koslowski and Dr. Thomas White discussing about adsorbed peptides; poetry and rhymes with Dr. Christian Ast; without Matthias Muenks, the ScienceSlam would have not been have as much fun; the regular Sushi evenings were motivation for the whole week and of course the legendary cooking sessions with Diana and Andy!

I would also like to thank Dr. Uta Schlickum, Dr. Christian Schön, Sebastian Koslowski and of course Dr. Stephan Rauschenbach for proof-reading my thesis. Special thanks go to my office mates, Dr. Alexander Hoyer and Alessio Scavuzzi, and our colorful and fun lunch group, including Dr. Christopher Kley, Dr. Christoph Große, Dr. Gordon Rinke, Dr. Gutzler, Dr. Christian Dette, Dr. Claudius Morchutt, Patrick Alexa, Anna Roslawska, Andreas Topp, Diana Hötger, Shai Mangel, and everyone who ever joined our table and was in for discussions.

Furthermore, I would like to express my gratitude to my parents for their faith in me and their support. You have always supported me in all decisions I made. For this, and your never ending love, I owe you all my gratitude. I also want to thank my sister Martina for all her care and support. You were always a role model to me, encouraging me to go abroad and dare to get to know the world.

Last, but not least, I like to thank Hauke. Your support, advice and unwavering love are undeniably the bedrock upon which the past years of our life have been built. Let's see what we can built on this cornerstone in the future!

Curriculum Vitae

Name Sabine Abb
Address Galileistrasse 30
70565 Stuttgart
Germany
Birth 10 October 1986
in Friedrichshafen, Germany

EDUCATION

Feb 13 – present **Doctoral Studies (PhD)**
Ecole Polytechnique Federale de Lausanne, Switzerland
Doctoral School: Materials Science and Engineering
Research assistant in the group of Prof. K. Kern
Max Planck Institute for Solid State Research Stuttgart

Jun 10 – Oct 10 **Exchange studies**
University of Southampton (UK), Group of Prof. A. Kanaras
Growth of Au nanoparticles and their self-assembly

Aug 09 – Jun 10 **Exchange studies**
Indiana University Bloomington (USA), Group of Prof. S.L. Tait
Self-assembly of metal organic networks

Oct 06 – Dec 12 **Diploma Studies in Chemistry**
Eberhard Karls Universität Tübingen, Germany
Diploma thesis: *Controlled growth of organic magnets on a highly ordered insulator surface*
in the group of Prof. Th. Chassé; supervisor: M.B. Casu

1997 – 2006 **Gymnasium (High School)**
Karl-Maybach-Gymnasium, Friedrichshafen

Bibliography

AWARDS & HONORS

- Aug 16* **2nd Price ScienceSlam BaWü-Stiftung**
 "Fluglotse für molekulare Elefanten"
 Wanted, ScienceSlam der Baden-Württemberg Stiftung
- Sept 15* **Student Presentation Award ECOSS 2015**
 "Folding and Self-Assembly of Peptides on surfaces"
 European conference on surface science, Barcelona
- Mar 13* **Student Poster Award Ascona 2013**
 "Understanding the peptide folding on surfaces in UHV
 by electrospray-ion beam deposition"
 Conference on Molecular Nanostructures, Ascona, 2014
- Mar 07 – Mar 12* **Stipend of the sdw**
 Stiftung der deutschen Wirtschaft

CONFERENCES & SEMINARS

- Sept 16* Indiana University Bloomington, seminar
 "Electrospray-ion beam deposition for high-resolution imaging of biomolecules
 by STM" (invited talk)
- Aug 16* ACS fall meeting, Philadelphia, PA
 "Electrospray-ion beam deposition for high-resolution imaging of biomolecules
 by STM" (talk)
- Jul 16* IMPRS, Berlin
 "High-resolution imaging and self-assembly of saccharides at surfaces"
 (poster)
- Jun 16* BaWü Stiftung
 "Fluglotse für molekulare Elefanten" (ScienceSlam)
- Mar 16* DPG spring meeting, Regensburg
 "Disaccharide self-assembly on metal surfaces" (talk)
- Aug 15* ECOSS 2015, Barcelona
 "Folding and Self-Assembly of Peptides on surfaces" (talk)
- Mar 15* DPG spring meeting, Berlin
 "Peptides as bio-organic building blocks for self-assembly on surfaces" (talk)
- Aug 14* ICMS, Geneva
 "Peptide folding and assembly on metal surfaces by soft landing in UHV"
 (poster)
- Mar 13* Conference on Molecular Nanostructures, Ascona
 "Understanding the peptide folding on surfaces in UHV by electrospray-ion
 beam deposition" (poster)

PUBLICATIONS

1. S. Abb, L. Harnau, J. Cortes, J.C. Schön, N.Tarrat, K. Kern, "Characterization and Self-assembly of disaccharides" *in preparation*
2. S. Rauschenbach, G. Rinke, R. Gutzler, S. Abb, A. Albarghash, D. Le, T.S. Rahman, M. Dürr, L. Harnau, K. Kern, "Two-Dimensional Folding of Polypeptides into Molecular Nanostructures at Surfaces" *ACS Nano*, **2017**.
3. J-N. Longchamp, S. Rauschenbach, S. Abb, C. Escher, T. Lатычевскаиа, K. Kern, H.W. Fink, "Imaging proteins at the single molecule level" *Proceedings of the National Academie of Science*, **2017**.
4. S. Abb, L. Harnau, R. Gutzler, S. Rauschenbach, K. Kern, "Two-dimensional honeycomb network through sequence-controlled self-assembly of oligopeptides" *Nature Communications*, **2016**, 7, 10335.
5. G. Dubey, R. Urcuyo, S. Abb, G. Rinke, M. Burghard, S. Rauschenbach, K. Kern, "Chemical Modification of Graphene via Hyperthermal Molecular Reaction" *Journal of the American Chemical Society*, **2014**, 136 (39), 13482.
6. S. Abb, S.-A. Savu, A. Caneschi, T. Chassé, M. B. Casu, "Paramagnetic nitronyl nitroxide radicals on Al₂O₃(11-20) single crystals: Nanoscale assembly, morphology, electronic structure and paramagnetic character towards future applications" *ACS Applied Materials & Interfaces*, **2013**, 5, 13006.
7. S.-A. Savu, S. Abb, S. Schundelmeier, J. D. Saathoff, J.M. Stevenson, C. Tönshoff, H. F. Bettinger, P. Clancy, M. B. Casu, T. Chassé, "Pentacene-based nanorods on Au(111) single crystals: Charge transfer, diffusion and step-edge barriers" *Nano Research*, **2013**, 6, 449.
8. U. Aygöl, H. Hintz, H.-J. Engelhaaf, A. Distler, S. Abb, H. Peisert, T. Chassé "Energy Level Alignment of a P3HT/Fullerene Blend during the initial Steps of Degradation" *Journal of Physical Chemistry C*, **2013**, 117, 4992.
9. S.-A. Savu, M. B. Casu, S. Schundelmeier, S. Abb, C. Tönshoff, H. F. Bettinger, T. Chassé, "Nanoscale assembly, morphology and screening effects in nanorods of newly synthesized substituted pentacenes" *RSC Advances*, **2012**, 2, 5112.
10. K. M. Schulz, S. Abb, R. Fernandes, M. Abb, A. G. Kanaras, O. L. Muskens, "Formation and Plasmonic Response of Self-Assembled Layers of Colloidal Gold Nanorods and Branched Gold Nanoparticles" *Langmuir*, **2012**, 28(24), 8874.
11. D. Skomski, S. Abb, S. L. Tait, "Robust Surface Nano-Architecture by Alkali-Carboxylate Ionic Bonding" *Journal of the American Chemical Society*, **2012**, 134 (34), 14165.