Investigation of Biomolecular Assemblies by Ion Mobility-Mass Spectrometry and Gas-Phase Infrared Spectroscopy

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)
submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Waldemar Hoffmann
from Dzhambul, Kazakhstan

2018

The work reported in this thesis was performed from April 2015 to July 2018, at the Freie Universität Berlin and at the Fritz Haber Institute of the Max Planck Society of Berlin, under the supervision of Prof. Dr. Kevin Pagel.

1st Reviewer:	Prof. Dr. Kevin Pagel Freie Universität Berlin
2nd Reviewer:	Prof. Dr. Christoph A. Schalley Freie Universität Berlin
Date of Defense:	15.10.2018

Danksagung

In den letzten Jahren haben mich eine Reihe von Menschen während meines akademischen Werdegangs begleitet und in unterschiedlichster Weise zum Erfolg dieser Arbeit beigetragen. Zuallererst möchte ich mich bei meinem Erstgutachter und Betreuer Prof. Dr. Kevin Pagel für die sehr frühe Aufnahme in seiner Arbeitsgruppe im Jahr 2012 als auch für die Vergabe verschiedenster und sehr interessanter Projekte bedanken. Ich bin sehr dankbar für die Freiheiten und Ratschläge, die er mir während meiner Forschung gegeben hat, aber auch für die Einblicke in die verschiedensten Aspekte einer akademischen Laufbahn sowie für die Gelegenheit, an zahlreichen Konferenzen im In- als auch Ausland teilzunehmen, und auch für die Möglichkeit, einen längeren Zeitraum meiner Promotion im Ausland forschen zu können.

Weiterhin bedanke ich mich bei Prof. Dr. Christoph Schalley für die stetige Übernahme des Zweitgutachters schon bereits während meiner Bachelorarbeit und letztlich auch meiner Dissertation.

Als nächstes möchte ich dem Fritz-Haber-Institut und allen Mitgliedern der Molekülphysik, insbesondere Prof. Dr. Gerard Meijer und Prof. Dr. Gert von Helden dafür danken, dass ich stets als Gast erwünscht war, an den vielseitigen Geräten arbeiten durfte und den so wunderbaren Schreibtisch behalten konnte. Ich habe sowohl die freundschaftliche und familiäre Atmosphäre als auch den wissenschaftlichen Austausch am Fritz-Haber-Institut geschätzt. In diesem Sinne bedanke ich mich auch bei Sandy Gewinner und Henrik Haak für die immensen Kicker-Trainingsstunden. Dr. Wieland Schöllkopf bin ich vor allem dankbar für die zahlreichen Beamzeiten und Pizzen, die wir zusammen an diesen Abenden verschlungen haben. Ein besonderer Dank geht an alle Mitglieder der Pagel und von Helden Gruppe. Ihr habt die Zeit am FHI unvergesslich gemacht und ich habe mich durch das freundschaftliche Klima und dem Zusammenhalt in der Gruppe immer gut aufgehoben gefühlt; angefangen beim wissenschaftlichen Austausch, über zu den vielen Spieleabenden, den gemeinsamen Unternehmungen, bis hin zu unseren Skikurztrips – bei denen ich allein ziemlich aufgeschmissen wäre (Wink an Johannas Rettung nach Jakob und Leo) – über das Team AG Pagel/von Helden während der zahlreichen Konferenzen. Besonderer Dank gilt Dr. Stephan Warnke, der eine Menge Geduld aufgebracht hat, um mich am *i*Mob und seinen Eigenheiten einzuweisen. Ich möchte mich auch bei meinen „Office Buddies“ Johanna Ho(f)fmann, Christian Manz und Rayoon Chang für die wunderbaren Ablenkungen, die witzigen Momente, die Freundschaft und die vielen E-Mail Erinnerungen bedanken! Auch den anderen

Mitgliedern der Gruppe, insbesondere Leo, Melanie, Jakob, Eike, Alexandra, Daniel, Malerz, Márkó, Maike und Chris bin ich für das Korrekturlesen der Arbeit als auch die tolle Zeit während und neben der Arbeit dankbar. Ich danke auch allen weiteren Personen, die mich während der Zeit begleitet haben und für die ich leider keine Zeit finde sie namentlich zu nennen. Tausend Dank für die Unterstützung! Weiterhin möchte ich auch allen Praktikanten und Bacherlor-/Masterstudenten für den Fleiß und eure Anstrengungen danken!

During my Ph.D. I had the chance to do some research for a few months at one of the nicest locations on earth, namely Santa Barbara. Thus, I would like also to acknowledge the DAAD for funding and Prof. Michael T. Bowers for hosting me during that time. I really appreciated the atmosphere in the lab and I could learn a new way of scientific thinking during that time that I am convinced will strengthen my future career. I thank you Michael T. Bowers for all fruitful discussions, all project ideas and your support. I also thank the whole Bowers lab for welcome me warmly, but in particular I thank Michael, Thomas, Matt, Alex, Veronica and Zach for the introduction into all instruments, projects and all social activities. Moreover, I thank also Nur Harell and Nathalie Koch for spending so much time together at the west coast. I will never forget the time and hope that I can see you soon back in Germany!

Zuletzt möchte ich mich sowohl bei meinen Eltern, meiner Schwester, meiner Tante, dem Rest der Familie als auch meinen Freunden für die Unterstützung in allen Lebenslagen bedanken! Ich bin so froh, dass ihr mir stets mit Rat und Tat zur Seite steht und immer für mich da seid. Ihr habt zu einem großen Teil zum Erfolg der Arbeit beigetragen. Ein besonderer Dank gilt Janine, die mich nun seit fast 6 Jahren meines Lebens begleitet, es mit mir aushält und durch dick und dünn geht!

Kurzzusammenfassung

Proteine gehören zu den wichtigsten Biomolekülen und sind an zahlreichen unerlässlichen Prozessen in lebenden Organismen beteiligt. Ihre räumliche Struktur ist eng mit ihrer Funktion verbunden, weshalb Proteinmissfaltungen nicht nur entscheidende biologische Prozesse behindern, sondern auch menschliche Funktionsstörungen wie zum Beispiel Alzheimer, Parkinson oder Typ-II-Diabetes auslösen können. Die beteiligten Proteine aggregieren und wandeln sich von meist ungeordneten Konformationen zu β-Faltblatt-reichen, unlöslichen Amyloid-Fibrillen um. Neuste Forschungen deuten jedoch an, dass nicht die vollentwickelten Fibrillen, sondern eher kurzlebige, polydisperse und polymorphe Oligomere die toxischen Spezies in den oben genannten Krankheiten darstellen. Demnach ist das Verständnis über Schlüsselfaktoren wie inter- und intramolekularen Wechselwirkungen, die die native Proteinstruktur stabilisieren sowie die Amyloidbildung beeinflussen oder gar verhindern, entscheidend für die zukünftige Wirkstoffentwicklung.

Traditionelle Methoden, die aus der kondensierten Phase bekannt sind, können nur bedingt für die strukturelle Charakterisierung amyloider Oligomere eingesetzt werden, da sie nur gemittelte Informationen eines Ensembles anstatt eines einzelnen Oligomerzustandes liefern. Im letzten Jahrzehnt hat sich Ionenmobilitäts-Massenspektrometrie (IM-MS) als eine bedeutende Alternative für die Untersuchung amyloidbildendender Systeme hervorgetan. Mittels IM-MS können Ionen nicht nur nach ihren Masse-zu-Ladung Verhältnis (m/z), sondern auch nach Unterschieden in Größe, Gestalt und Ladung, getrennt werden. Ferner kann IM-MS mit anderen Techniken wie Infrarot (IR) Spektroskopie in der Gasphase orthogonal gekoppelt werden, um komplementäre und hoch diagnostische Informationen über inter- und intramolekulare Wechselwirkungen von m/z- und größenselektierten Ionen zu erhalten. Diese Kombination ermöglicht es daher, wichtige Einblicke in Sekundär- als auch Tertiärstrukturübergänge für individuelle Oligomerzustände zu erhalten.

In dieser Arbeit wurden Biomolekülaggregate – von Aminosäuren über Peptide bis hin zu Proteinen – mittels IM-MS und IR-Spektroskopie in der Gasphase untersucht. Zunächst wurde eine neuartige Hydrophobizitätsskala für Aminosäuren entwickelt, die im Gegensatz zu vorherigen Skalen entropische Effekte des Lösungsmittels ausschließt und damit die wahre Natur der Seitenkettenhydrophobizität widerspiegelt. Zudem kann die neue Skala ohne Weiteres für nicht natürliche Derivate erweitert werden, wie dies beispielsweise für fluorierte Aminosäureanaloga gezeigt wurde.

Anschließend wurden anhand eines helikalen Modellpeptides weitere, für die Peptid- und Proteinfaltung entscheidende Effekte, insbesondere Ladungs-Dipol-Wechselwirkungen und Rückgrat- als auch Seitenketten-Wasserstoffbrückenbindungen, untersucht. Es wurde gezeigt, dass die Ladungs-Dipol Wechselwirkung am wichtigsten für die Stabilisierung der helikalen Struktur in der Gasphase ist, während das Entfernen einer einzelnen Wasserstoffbrückenbindung die Feinstruktur nur geringfügig ändert.

Im nächsten Schritt wurde untersucht, inwieweit die native Struktur von zwei prototypischen helikalen und β-Faltblatt-reichen Proteinen in Abwesenheit des Lösungsmittels erhalten bleibt und welchen Einfluss die Ladung auf die Proteinfaltung in der Gasphase hat. Sobald Proteine unter milden Bedingungen aus gepufferter Lösung in die Gasphase überführt werden, können Spezies in niedrigen Ladungszuständen beobachtet werden, deren Form und Gestalt vergleichbar ist mit der jeweiligen nativen räumlichen Struktur in Lösung. Des Weiteren wird eine bemerkenswerte Übereinstimmung zwischen den IR-Spektren aus der Gasphase und den Fourier-Transform-Infrarot (FT-IR) Spektren als auch Zirkulardichroismus (CD) Spektren aus kondensierter Phase beobachtet. Damit verdeutlicht diese Studie, dass Aspekte der nativen Sekundär- als auch Tertiärstruktur von mittelgroßen Proteinen nach dem Transfer in die Gasphase bewahrt werden können. Steigt jedoch die Anzahl der Ladungen am Protein an, wird dessen native Faltung destabilisiert und folglich kommt es zu einer Umfaltung in artifizielle Helices. Bei sehr hohen Ladungszuständen entwinden sich selbst diese helikalen Strukturen, sodass letztlich gestreckte Konformationen vorliegen, die sich durch Coulombabstoßung und C_5-Wasserstoffbrückenbindungen auszeichnen. Die Bildung gestreckter Strukturen ist unabhängig von der ursprünglichen Sekundärstruktur des Proteins und kann deshalb als eine neue universelle Sekundärstruktur angesehen werden, die jedes Protein bei sehr hohen Ladungszuständen in der Gasphase annimmt.

Abschließend konnten erstmalig Änderungen in den Sekundärstrukturen einzelner, kurzlebiger, amyloider Oligomere, die von den Peptidsequenzen VEALYL und NFGAIL gebildet werden, direkt nachverfolgt werden. Oligomere, die nur aus vier bis neun Untereinheiten bestehen, enthalten bereits einen signifikanten Anteil an β-Faltblattstrukturen. Demzufolge erfolgt der charakteristische Übergang von ungeordneten zu β-Faltblatt-reichen Strukturen bereits im frühen Stadium des Aggregationsprozesses amyloider Peptide und Proteine. Diese Studien repräsentieren die erste direkte Sekundärstrukturanalyse von individuellen amyloiden Oligomeren und verdeutlichen das Potential von MS-basierten Techniken im Bereich der Amyloidforschung.

Abstract

Proteins are one of the most important biomolecules and are involved in a vast number of vital processes inside living organisms. Their structure is closely related to their function and therefore protein misfolding can not only impede crucial biological processes but may also trigger a variety of human disorders such as Alzheimer´s and Parkinson´s disease or type II diabetes. The involved proteins assemble and undergo a characteristic transition from usually unordered conformations into mainly β-sheet rich, insoluble amyloid fibrils. Recent investigations, however, indicate that not the mature fibrils but rather transient, polydisperse and polymorph folded intermediates represent the toxic species in the afore-mentioned diseases. Thus, understanding key factors such as inter- and intramolecular interactions that stabilize the native protein structure or either manipulate or hinder the amyloid formation process is essential for future drug development.

Traditional condensed-phase methods are of limited use for the structural characterization of amyloid intermediates, because they only provide ensemble-averages rather than information on individual oligomeric states. In the last decade, ion mobility-mass spectrometry (IM-MS) has emerged as a powerful alternative to investigate amyloid forming systems. IM-MS separates ions not only based on mass-to-charge (m/z) but also on differences in size, shape and charge. More importantly, it can be combined with other orthogonal techniques such as gas-phase infrared (IR) spectroscopy, providing complementary and highly diagnostic information on inter- and intramolecular interactions of m/z- and shape-selected ions. As a result, this combination yields crucial insights on secondary- as well as tertiary-structural transitions of individual oligomeric states.

In this thesis, a variety of biomolecular assemblies, ranging from amino acids over peptides to proteins, have been studied using IM-MS with gas-phase IR spectroscopy. First, a novel hydrophobicity scale for amino acids has been developed that in contrast to previous scales now excludes entropic effects of the solvent and therefore more accurately represents the true nature of the side-chain hydrophobicity. In addition, the new scale can be readily extended for non-natural versions, as exemplarily shown for fluorinated amino acid analogues.

Subsequently, other effects crucial for peptide and protein folding, in particular charge-dipole interactions and backbone as well as side-chain hydrogen bonding have been studied on a helical model peptide. It is shown that the charge-dipole interaction is most critical for stabilizing the helical structure in the gas phase, while the deletion of a single hydrogen bond only marginally alters the fine structure.

Furthermore, it is demonstrated in how far the native structure of two prototypical helical and β-sheet-rich proteins can be maintained in the absence of solvent, and which influence charge has on protein folding in the gas phase. When the proteins are carefully transferred form buffered solution into the solvent-free environment of the gas phase, usually lowly-charged species with comparable sizes to the respective native-structure are observed. In addition, their gas-phase IR spectra show an incredible agreement with their condensed-phase Fourier-transform infrared (FT-IR) and circular dichroism (CD) spectra. Thus, this study unambiguously demonstrates that aspects of the native secondary- as well as tertiary-structure of medium-sized proteins can be preserved in the absence of solvent. When the charges on the protein, however, increase, the native fold is destabilized and the protein refolds into artificial, gas-phase helices. At very high charge states these helices then unravel to form string-like structures that are governed by Coulomb-repulsion and C_5-hydrogen bonds. The formation of these string-like structures is independent of the initial protein conformation and therefore can be seen as a new and universal secondary-structure that all proteins can access at very high charge states in the gas phase.

Finally, for the first time secondary structural transitions for individual, transient and toxic amyloid intermediates formed by the two peptide sequences VEALYL and NFGAIL have been successfully identified. The data show, that oligomers with as little as four to nine subunits can already contain a significant amount of β-sheet structure. Thus, the characteristic transition from unordered to β-sheet structures already occurs very early in the assembly process of amyloidogenic peptides and proteins. These studies represent the first direct secondary structure analysis of individual amyloid intermediates and highlight the potential of MS-based techniques for the amyloid research.

This work is based on the following publications[1]

Novel Ion Mobility-Mass Spectrometry based Hydrophobicity Scale for Amino Acids

W. Hoffmann, J. Langenhan, S. Huhmann, J. Moschner, M. Accorsi, J. Seo, J. Rademann, B. Koksch, M. T. Bowers, G. von Helden, K. Pagel, **submitted for review**.

Assessing the Stability of Alanine-Based Helices by Conformer-Selective IR Spectroscopy

W. Hoffmann, M. Marianski, S. Warnke, J. Seo, C. Baldauf, G. von Helden, K. Pagel, *Phys. Chem. Chem. Phys.* **2016**, 18, 19950-19954.

doi: 10.1039/C6CP03827A

Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study

J. Seo, **W. Hoffmann**, S. Warnke, M. T. Bowers, K. Pagel, G. von Helden, *Angew. Chem. Int. Ed.* **2016**, 55, 14173-14176.

doi: 10.1039/C6CP03827A

An Infrared Spectroscopy Approach to follow β-sheet Formation in Peptide Amyloid Assemblies

J. Seo[‡], **W. Hoffmann[‡]**, S. Warnke, X. Huang, S. Gewinner, W. Schöllkopf, M. T. Bowers, G. von Helden, K. Pagel, *Nat. Chem.* **2017**, 9, 39-44.

doi: 10.1038/nchem.2615

NFGAIL Amyloid Oligomers: The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation

W. Hoffmann, K. Folmert, J. Moschner, X. Huang, H. von Berlepsch, B. Koksch, M. T. Bowers, G. von Helden, K. Pagel, *J. Am. Chem. Soc.* **2018**, 140, 244-249.

doi: 10.1021/jacs.7b09510

Ion Mobility-Mass Spectrometry and Orthogonal Gas-Phase Techniques to study Amyloid Formation and Inhibition

W. Hoffmann, G. von Helden, K. Pagel, *Curr. Opin. Struct. Biol.* **2017**, 46, 7-15.

doi: 10.1016/j.sbi.2017.03.002

---

[1] For further publications obtained during this thesis see **List of Publications**.

[‡] J. Seo and W.H. contributed equally to this work.

Contents

1 Introduction

The development of soft ionization techniques[1] that are capable of producing gaseous ions of large biomolecules initiated a new golden age in which biological macromolecules have been studied in the gas phase. From these initial studies,[2-4] this field of research tremendously evolved and demonstrated that valuable insight could not only be obtained for monomeric systems but also for a variety of non-covalently associated biomolecular complexes.[5-10] Recently, the combination of ion mobility-mass spectrometry (IM-MS) with gas-phase infrared (IR) spectroscopy further showed that information of the mass-to-charge (m/z) ratio as well as on inter- and intramolecular interactions can simultaneously be obtained in short time frames using very low sample amounts. This combination allows the isolation of a single species that can be directly probed out of its heterogenous environment without changing the underlying equilibrium. Such sensitivity and selectivity cannot be achieved by other traditional methods in solution and turned out to be of exceptional use for the structural investigation of a variety of biomolecules ranging from nucleobases,[11-12] carbohydrates,[13-14] amino acids,[6, 15] peptides[16-17] and even large biomolecular machineries of proteins.[18]

Probably one of the most relevant examples for systems forming complex mixtures are amyloidogenic peptides and proteins that are involved in roughly 20 human disorders such as Alzheimer´s and Parkinson´s disease or type II diabetes.[7, 19-20] These proteins aggregate and undergo a conformational transition from a soluble, usually unordered or helical conformation into mainly β-sheet rich, insoluble amyloid fibrils. Depending on the type of the disease, these deposits can be found in the brain, in skeletal tissue or in other organs with amounts ranging from hardly detectable quantities to kilograms. Nevertheless, the underlying assembly pathways and involved structures are still poorly understood to date. Increasing evidence indicates, however, that not mature fibrils but rather oligomeric folding intermediates that are present during the transformation process are the toxic species in the afore-mentioned diseases. Thus, elucidating molecular interactions that occur during the structural transformation and understanding the consecutive formation of amyloids on a molecular level is critical for improving the treatment of these diseases and the development of effective drugs in the future.

1.1 Motivation and Outline of the Thesis

Neurodegenerative disorders such as Alzheimer´s disease are the second most important widespread disease in industrialized countries as shown by a worldwide study of the *World Health Organization* (WHO).[21] As a results of the steadily increasing life expectancy – on average from 47 years (1900-1911) to nowadays over 80 years – more and more people reach an age, where neurodegenerative disorders are common and therefore affect the society. The amyloid fibril formation is a universal feature of all proteins, but certain amino acids exhibit a higher propensity towards fibril formation than others.[22] The fibril formation rate of peptides and proteins was shown to correlate with high level of amino acid side-chain hydrophobicity, high amino acid propensity to participate in a β-sheet structure, small number of charged residues and the presence of alternating hydrophilic and hydrophobic residues in the aggregation core. This thesis is dedicated to evaluate the potential and utility of combining IM-MS with gas-phase IR spectroscopy for understanding properties and the structural evolution of amyloid intermediates. In the future, this information could help to stabilize the native protein structure or manipulate or hinder the amyloid formation process. **Chapter 2** covers fundamentals of protein folding in solution and in the gas phase as well as misfolding, which trigger amyloid fibril formation. Furthermore, applied methods and their current applications for the structural characterization of amyloid intermediates in the gas phase are described.

Amino acid side-chain hydrophobicity influences peptide and protein oligomerization as well as the amyloid fibril formation propensity. However, a consensus about classifying amino acids according to their side-chain hydrophobicity and about their relative ranking of hydrophobicity values is still not obtained because current condensed-phase approaches depend on unique interactions between the environment and the amino acid. In **Chapter 3**, a novel concept for an IM-MS-based hydrophobicity scale is introduced. It is shown that the polarity of the amino acid side-chain highly influences the respective cluster formation in the gas phase. This effect can be used to generate a novel hydrophobicity scale for amino acids which in contrast to previous scales now excludes entropic effects of the solvent and therefore more accurately reflects the intrinsic side-chain hydrophobicity. Furthermore, the presented approach can be readily extended for classifying non-natural amino acids.

Besides the side-chain hydrophobicity, other inter- and intramolecular interactions are important for peptide and protein folding. In particular, charge-dipole interactions, backbone and side-chain hydrogen bonding are crucial for the formation of helical structures in the gas phase. In **Chapter 4**, the stabilizing impact of these interactions within the model peptide

$[Ac\text{-}Ala_{10}Lys + H]^+$ is studied in detail using the combination of IM-MS with gas-phase IR spectroscopy and theoretical calculations. Alanine-to-lactic acid (amide-to-ester) substitutions have been used to selectively delete a single backbone hydrogen bond, while side-chain hydrogen bonds were inhibited by the non-covalent attachment of a crown ether.

Electrostatic interactions are enhanced in the gas phase and charged residues that are usually solvated in solution can destabilize the native protein structure due to Coulomb repulsion. For medium-sized proteins it is therefore not clear, in how far their native structure can be maintained in the solvent-free environment of the gas phase. **Chapter 5** is dedicated to study the influence of charge on the native protein structure of myoglobin and β-lactoglobulin. It is shown that structural aspects of the native protein structure can be preserved in the absence of solvent when these proteins are carefully transferred from buffered solution into the gas phase. In addition, the influence of an increasing amount of charge on gas-phase protein folding is demonstrated.

Finally, in **Chapter 6** and **Chapter 7** the combination of IM-MS and gas-phase IR spectroscopy is used to pre-select individual amyloidogenic oligomers in order to study their transitions in secondary as well as tertiary structure during their early assembly pathway. These studies successfully identified for the first time characteristic transitions from an unordered to a mainly β-sheet structure for early amyloidogenic oligomers which consists of as little as four to nine peptide strands. In addition, it shows that the combination of IM-MS with gas-phase IR spectroscopy is a versatile and highly valuable analytical tool for the structural investigation of amyloid oligomers.

Chapter 8 concludes the thesis and gives future prospects for instrumentalization and possible applications of IM-MS with gas-phase IR spectroscopy for the structural characterization of other biological systems.

2 Fundamentals*

2.1 Proteins and their Structural Elements

Proteins represent one of the most important classes of biological molecules in living organisms. They are composed of 20 α-L-amino acids, which are connected to each other in a linear fashion *via* covalent peptide bonds (**Figure 2.1**). This sequence, often referred to as the primary structure, determines how the protein backbone folds into a specific three-dimensional shape. Under physiological conditions the protein adopts a unique, thermodynamically most stable and kinetically accessible native structure (Anfinsen´s dogma).[23-24] This native conformation represents the global minimum in a steep funnel-shaped free energy surface, which effectively helps the protein to fold into its favored structure and thereby solving the Levinthal paradox.[25] The folding is mainly influenced by intra- and intermolecular interactions such as hydrogen bonds, van der Waals forces, hydrophobic and electrostatic interactions as well as the interplay of the protein with the surrounding environment. These interactions lead to local, three-dimensional arrangements described as the secondary structure of the protein. The three main secondary-structural elements are helices, β-sheets and turns, and their further arrangement is described as the tertiary structure. In addition, multiple proteins can subsequently assemble into a non-covalently bound complex which then is referred to as the quaternary structure.[26] The final three-dimensional shape of a protein is closely related to its biological function and misfolding therefore may not only vanish the functionality but can, under certain circumstances, also tremendously impede biological processes, thus triggering various pathological conditions.

* This chapter is in part based on the work published in W. Hoffmann, G. von Helden and K. Pagel, *Curr. Opin. Struct. Biol.* **2017**, *46*, 7-15, DOI: 10.1016/j.sbi.2017.03.002. Figures and content adapted with permission. Copyright 2017 Elsevier Ltd..

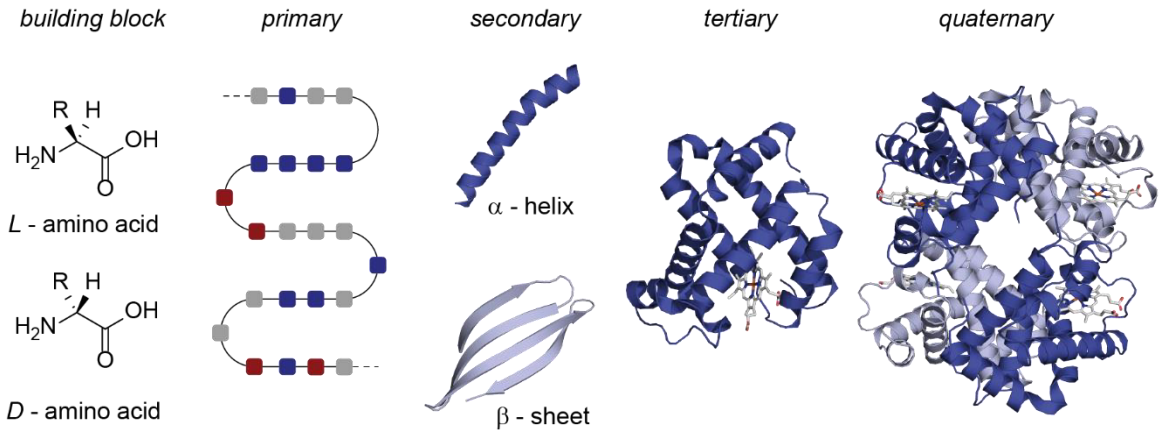


Figure 2.1: Schematic representation of protein structural elements. Amino acids (with a given side chain *R*) are covalently linked *via* a peptide bond into a linear sequence, referred to as the primary structure. This primary structure determines the folding into local, secondary structural elements, which can then arrange to the tertiary structure. Moreover, multiple proteins can assemble into a non-covalently bound multi-subunit complex, the so-called quaternary structure. Figure adapted with permission from Melanie Göth, *Investigation of Protein-Ligand Complexes by Native Mass Spectrometry and Ion Mobility-Mass Spectrometry*, Inaugural-Dissertation, **2017**, Freie Universität Berlin.

2.2 Protein Misfolding – Amyloid Structure

Although the primary structure determines a favored three-dimensional conformation in which the respective peptides and proteins maintain their soluble states, the protein backbone can under certain circumstances misfold and convert into nonfunctional and potentially damaging aggregates associated with various disorders such as Alzheimer´s and Parkinson´s disease or type II diabetes.[19-20, 27] The formation of insoluble deposits, often referred to as amyloid plaques, was shown to be an universal phenomenon of all peptides and proteins independent of a defined sequence homology.[28] They all form amyloid fibrils featuring a threadlike structure with typically 7-13 nm in diameter and several microns in length as shown by transmission electron (TEM) and atomic force microscopy (AFM). These mature fibrils have a high level of structural order and are usually composed of 2-8 protofilaments (each approximately 2-7 nm in diameter) twisted around each other or associated laterally as flat ribbons (**Figure 2.2**a).[29] Extensive research in the last 25 years has revealed the detailed structural information of mature fibrils formed by a couple of smaller peptides[30-31] up to fibrils of the full-length Aβ peptide[32-33]. These fibrils share a common fibrillar cross-β structure[34-35], in which β-strands are aligned perpendicularly to the fibril axis and are assembled into a repeating pattern of either parallel or antiparallel β-sheets[31] as initially detected using X-Ray

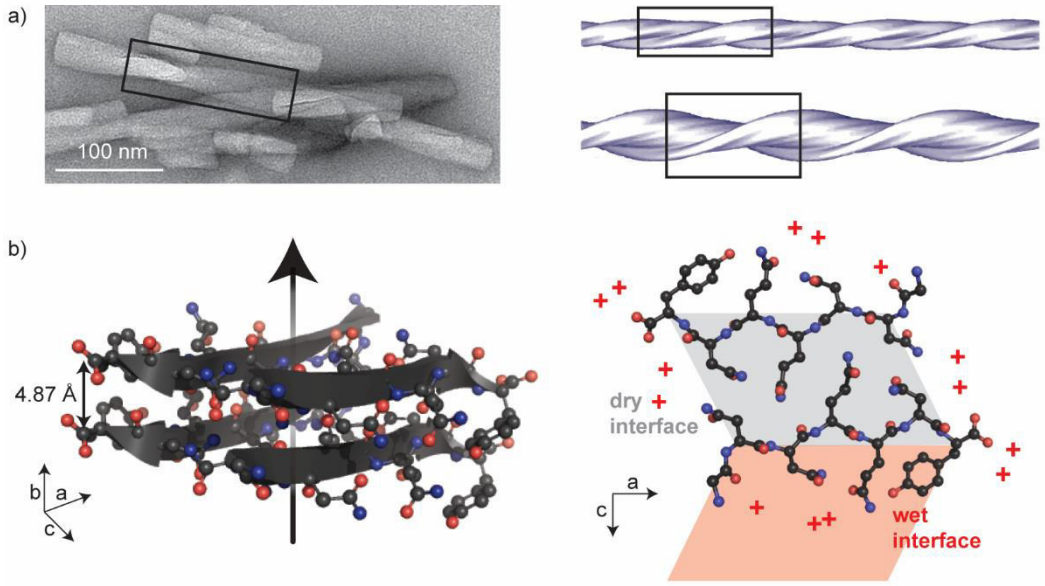


Figure 2.2: Three-dimensional structures and structural models of amyloid fibrils. a) The NFGAIL hexapeptide forms *in vitro* fibrillar aggregates with anisometric cross section, a mean apparent diameter of 20-30 nm (left). Similar mature fibrils are usually formed by a pair of protofibrils (2-8 in total with each 2-7 nm in diameter) twisting around each other (right). b) Structure of an amyloid fibril formed by the peptide sequence GNNQQNY (PDB: 1YJP). Carbon atoms are colored in black, oxygen in red, nitrogen in blue, and water molecules are indicated as red +. The peptide assembles into a repeating pattern of stacked β-sheets along the fibril axis. The protruding side-chains interdigitate like a steric-zipper, forming a dry interface, which excludes water molecules from the fibril interior.

diffraction[36] and later supported by Fourier Transform-Infrared (FT-IR)[37-38] spectroscopy, solid state nuclear magnetic resonance (ssNMR) spectroscopy[29] as well as X-Ray diffraction on small peptide fibrillar microcrystals[27, 30-31] (**Figure 2.2**b). Furthermore, the protruding side-chains of each β-strand interdigitate like a steric zipper, forming a dry, tightly self-complementing interface. This in-register alignment optimizes intermolecular interactions and maximizes the number of hydrogen bonds and hydrophobic interactions along the fibril axis.

Mature fibrils are, however, not static oligomers, but rather represent a dynamic assembly at the ends of which monomer units dynamically dissociate and associate.[39] This "molecular recycling" takes place in biologically relevant time scales [40] and mature fibrils may therefore serve as a potential source of oligomers.[41]

2.3 Mechanism of Amyloid Formation

Even though the proteins involved in amyloid formation may differ significantly in their size and sequence, they share similar assembly characteristics.[42] The pathway usually relies on a

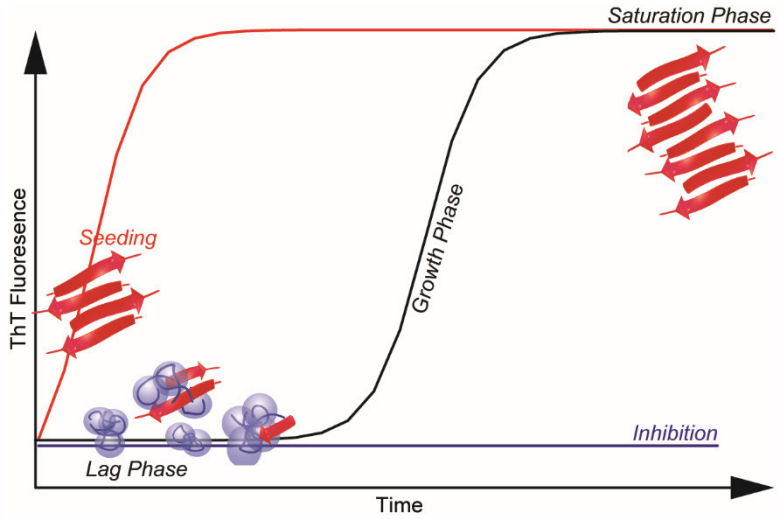


Figure 2.3: Characteristics of amyloid formation. The typical amyloid formation pathway can be divided into three stages, the lag-, growth- and saturation phase (black trace). During the lag phase, transient and highly interconverting intermediates are found, but no fibrils are present. Once a so-called nucleus is formed, fibril formation is initiated (growth phase), and an autocatalytic growth phase leads to mature fibrils that can be observed at the saturation phase. The addition of nuclei in the form of seeds eliminates the lag phase and directly initiates the growth phase (red trace). The presence of small inhibitors suppresses the formation of fibrils (blue trace).

nucleation-dependent mechanism, which can be divided into the lag-, growth- and saturation-phase (**Figure 2.3**). Dye molecules such as Congo red or thioflavin T (ThT) are commonly used to monitor the kinetics of fibril formation in real-time. Both dyes bind to amyloid fibrils *via* intercalation, which results in change of the structure of the dye and increased fluorescence of the dye.[43] It is believed that protein aggregation is initiated by an unfolding/misfolding event of the monomer, generating species capable of undergoing oligomerization (lag phase).[42] These oligomers can further exchange subunits or undergo conformational transitions, but once a so-called nucleus is formed, fibril formation is facilitated *via* an autocatalytic growth (growth phase) up to the presence of mature fibrils (saturation phase). The autocatalytic growth can be further accelerated by fibril fragmentation, providing additional fibril ends for monomer association, whereas end-joining of two fibrils may reduce the rate of mature fibril formation.[44]

The duration of the lag phase is highly dependent on environmental conditions[28] or the presence of small molecules that can inhibit fibril formation (**Figure 2.3**, blue trace).[45] The dynamics of the oligomers can also lead to different behavior, as for example demonstrated for beta-2 microglobulin and its H51A alloform.[46] Both proteins form similar oligomers, but the lag phase of the H51A alloform is significantly longer than that of the wild type beta-2

microglobulin. Here, real-time mass spectrometry (MS) using ^{15}N-enriched proteins was performed to monitor the exchange of subunits at the end of the individual lag phases. H51A oligomers were shown to undergo no, or a barely measurable, exchange signifying that these oligomers are kinetically trapped species. The wild type oligomers, on the other hand, exhibited a rapid subunit exchange, and this dynamic behavior accounts for the short lag phase of the wild type beta-2 microglobulin.[46]

In the presence of pre-formed nuclei that can act as fibrillar seeds, the entire lag phase can be eliminated (**Figure 2.3**, red trace).[42] The fibrillar seeds presumably act as a template for monomers to refold and assemble into amyloid fibrils. For example, fibrillar amyloid-β40 (Aβ40) can cross-seed the human islet amyloid polypeptide (hIAPP). The reverse, however, does not occur, *i.e.* hIAPP seeds are not capable of cross-seeding Aβ40 monomers. The cross-seeding phenomenon is therefore, in many cases, sequence-dependent. If both peptides are co-incubated another effect termed co-polymerization occurs, where mixed Aβ40/hIAPP-oligomers and mixed fibrils are observed. Therefore, it is essential to distinguish the cross-seeding phenomenon from co-polymerization.[47]

Recent investigations, however, indicate that not the mature fibrils but rather transient intermediates that arise during the transformation process (lag phase) are the toxic species in neurodegenerative disorders.[28, 48] A detailed structural knowledge of these oligomers remains elusive due to the inherent difficulty of detection and quantification of such systems. Amyloid intermediates are polydisperse, meaning they occur in a wide distribution of aggregation states with complex structural transitions. In addition, they are also often polymorph, *i.e.* more than one conformation coexist for one oligomeric state. As a consequence, X-Ray diffraction experiments exclusively deliver artificial crystal structures of amyloid oligomers,[49-50] while other traditional condensed-phase methods like NMR or circular dichroism (CD) spectroscopy only provide averaged properties of the assembly in solution. Thus, very little is known about the structure of individual oligomers and it is one of the most challenging issues to directly investigate these early oligomers.

2.4 Ion Mobility-Mass Spectrometry

2.4.1 General Principles

In contrast to the above-mentioned limitations of traditional condensed-phase techniques, gas-phase methods enable the isolation and characterization of individual species in the presence of many others without affecting the equilibrium of the ensemble. Especially, ion

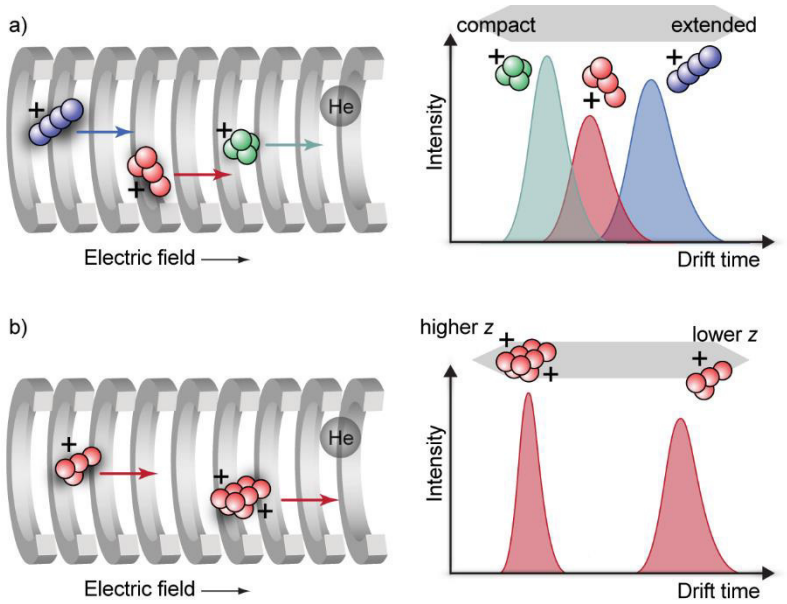


Figure 2.4: Schematic principle of an ion mobility separation. Ions are guided by a weak electric field through a drift cell that is filled with an inert buffer gas such as helium. During their migration ions of the same m/z are separated based on differences in their (a) size and shape as well as (b) charge. The drift time of each ion package is depicted as an arrival time distribution (ATD).

mobility-mass spectrometry (IM-MS) has recently emerged as a valuable tool for the structural characterization of amyloid intermediates.[7] While MS separates ions based on their mass-to-charge ratio (m/z), ion mobility spectrometry distinguishes between ions based on differences in size, shape and charge. As both methods operate on different time scales, they are compatible for an orthogonal analysis within a single instrument.[51]

An IM-MS analysis for amyloid intermediates typically starts with the preparation of the sample under conditions where fibril formation occurs. Soluble oligomers are then carefully transferred into the gas phase using soft ionization techniques such as electrospray ionization (ESI),[9, 16, 52] matrix-assisted laser desorption (MALDI)[53] or laser-induced liquid bead ion desorption (LILBID)[54-55]. Packets of these ions are further injected into an ion mobility cell filled with an inert buffer gas, through which they migrate due to a weak electric field that is applied. During this migration, compact ions (green) undergo fewer collisions with the buffer gas than extended species (blue) and therefore traverse the cell faster, that is with a shorter drift time (**Figure 2.4**a). Furthermore, highly charged ions experience a greater force of the electric field than ions with less charges and will consequently have a higher mobility through the drift cell, thus resulting with a shorter drift time (**Figure 2.4**b). This analysis based on charge, size and shape enables the analysis of species, which exhibit an identical m/z but differ in their oligomer size or conformation.

The drift time of an ion is further influenced by various instrumental parameters (such as pressure and drift voltage) similar to the retention time obtained in liquid chromatography. However, the drift times measured in IM-MS experiments can be converted into an instrument-independent parameter, the so-called collision cross-section (CCS).

2.4.2 Collision Cross-Section

The CCS is a molecular property that represents the rotationally-averaged area of an ion which interacts with the buffer gas. It is independent from instrumental parameters, universal comparable and can also be theoretically derived from X-Ray, NMR or model structures *via* established techniques.[56-58] Thus, the comparison of experimental and theoretical CCS values often allows a detailed structural assignment for the investigated species.[59-60]

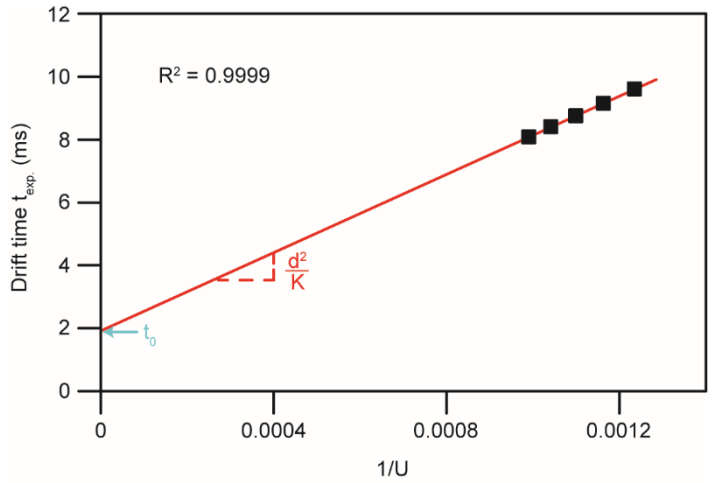
The CCS of an ion can be obtained under so-called low field conditions by which the weak electric field induces a directed motion of the ion that is slower than its thermal motion.[61-62] Under this condition, and at constant pressure p and temperature T, the ion velocity v of an ion is constant and can be described classically as the drift time t_D that an ion needs to traverse the drift cell of length d. Furthermore, the velocity v is proportional to the mobility K of the ion and the applied electric field, $E = U/d$ (equation 2.4.1).

$$v = \frac{d}{t_D} = K \cdot E \tag{2.4.1}$$

The measured drift time $t_{exp.}$ is, however, a sum of the time t_D that an ion needs to travel through the drift cell d and the time t_0 that an ion spends outside of the drift cell until it reaches the detector. Thus, equation 2.4.1 can be expressed as follows:

$$t_{exp.} = \frac{d}{K \cdot E} + t_0 = \frac{d^2}{K \cdot U} + t_o \tag{2.4.2}$$

The resulting linear equation 2.4.2 shows that $t_{exp.}$ is proportional to the inverse mobility ($1/K$) as well as the inverse of the applied drift voltage ($1/U$). The mobility K of an ion can be extracted from the slope of the linear regression (**Figure 2.5**). The intercept of the fit represents the time t_o that an ion spends outside the drift cell (for *e.g.* ion funnels, ion guides, quadrupoles). At known mobility K, temperature T and pressure p the CCS Ω of an ion can be calculated using the Mason-Schamp equation[61-62] 2.4.3, where $q = ez$ is the total charge



$$\Omega = \frac{3q}{16N}\sqrt{\frac{2\pi}{\mu k_B T}}\frac{1}{K}\frac{p_0}{p}\frac{T}{T_0} \tag{2.4.3}$$

Figure 2.5: Determination of the mobility *K* of an ion from drift-tube measurements. The experimental drift time $t_{exp.}$ is plotted as a function of the inverse drift voltage $1/U$. The resulting linear regression follows equation 2.4.2 and allows to extract *K* from the slope.

of the ion, N the drift gas number density, μ the reduced mass of the ion and the drift gas, k_B the Boltzmann constant, p_0 the standard pressure and T_0 standard temperature.

2.4.3 A Collision Cross-Section Model for Self-Assembly

The CCS can also provide information regarding the structural evolution of amyloid intermediates during their self-assembly. For a fibril, stacked β-sheet-dominated self-assembly, where the association of single monomer units proceeds in a highly directional manner along the fibril axis, the increase in CCS Ω_n, with growing size n, follows a linear pattern (equation 2.4.4).[63] A scheme of the fibril self-assembly and its correlation to the CCS is shown in **Figure 2.6**. The slope a represents the change in Ω_n, which arises by the addition or removal of one monomer unit along the fibril axis. For a finite-sized fiber, the intercept k describes the CCS of fiber caps. This model was already applied on a variety of amyloidogenic peptides such as VEALYL (derived from the B-chain of insulin), NNQQNY (derived from the yeast prion protein Sup35) or SSTNVG (derived from the human islet amyloid polypeptide) and the linear evolution of experimental CCSs for more extended oligomers formed by these sequences showed a good agreement with theoretical CCSs derived from their fibril X-Ray structures.[63]

$$\Omega_n = a \cdot n + k \tag{2.4.4}$$

When the early self-assembly of peptides and proteins, however, proceeds non-specifically in an unordered way, the oligomer CCS values scale with $n^{2/3}$. This is

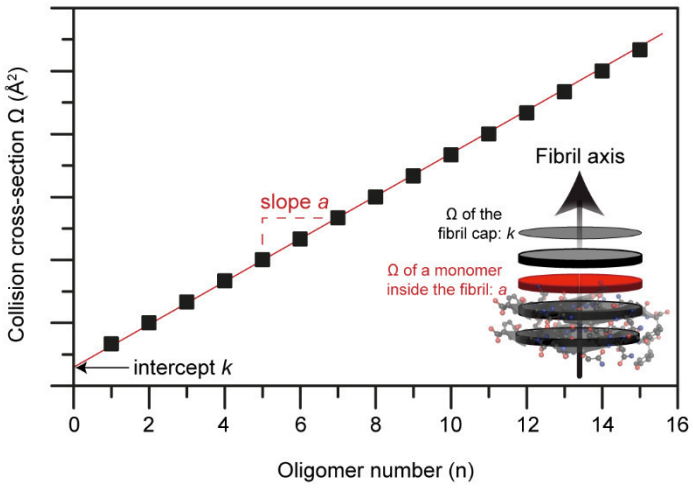


Figure 2.6: Theoretical model for the CCS Ω evolution of a fibril self-assembly. The CCS Ω of any intermediate scales linearly with the oligomer size n (see equation 2.4.4). The slope of the linear growth trend reflects the change in CCS generated by adding or removing one monomer unit along the fibril axis, whereas the intercept k stands for the CCS of the fiber caps.

mathematically described as a spatially isotropic self-assembly according to equation 2.4.5.[63] This description is based on the assumption that the density of two spatially isotropic oligomers with cluster sizes n and m is not changed. The CCS Ω_n for a spatially isotropic peptide cluster with an arbitrary oligomer unit n is connected to any reference points σ_m with size m. Typically, the CCS σ_m of the monomer $m = 1$ is used as a reference point, so that it holds $\sigma_m^* = \sigma_1$. Experimental CCSs of early, compact oligomers formed by Leu-enkephalin (YGGFL) follow this model[63], which is in good agreement with the isotropic crystal lattice[64] in the solid state.

$$\Omega_n = \frac{\sigma_m}{m^{2/3}} \cdot n^{\frac{2}{3}} = \sigma_m^* \cdot n^{\frac{2}{3}} = \sigma_1 \cdot n^{2/3} \tag{2.4.5}$$

The oligomer formation is usually accompanied by an increasing amount of charge, which can cause an unfolding into more extended species (higher CCS values) due to Coulomb repulsion. It is therefore often not clear if the presence of an extended species already reflects the transition from an unordered (compact) into an extended and ordered (β-sheet rich or helical) oligomer or is just an artefact of Coulomb repulsion in the gas phase. Here, infrared (IR) spectroscopy can yield complementarily, highly diagnostic information of the secondary structural content in peptide and protein assemblies by probing their intra- and intermolecular vibrations.[65-66]

2.5 Infrared Spectroscopy

2.5.1 General principles

The vibrations of a molecule or a molecular complex strongly depend on the local environment of the respective oscillators. Interactions such as hydrogen bonds and salt bridges vary the strength of chemical bonds and thereby yield diagnostic shifts in the absorption bands towards lower or higher wavenumbers and may increase or decrease relative intensities of the vibrational modes. Utilizing these shifts in the position and intensity of absorption bands, IR spectroscopy has been used extensively for the structural characterization of proteins and peptides.[65-66] The protein backbone yields three characteristic vibrational modes: the amide I, II and III bands, corresponding to the C=O stretching ($\tilde{\nu}$ = 1600-1700 cm^{-1}), the N-H bending ($\tilde{\nu} \approx$ 1550 cm^{-1}) and the C-N stretching mode ($\tilde{\nu}$ = 1220-1320 cm^{-1}), respectively. Upon strong hydrogen bonding, the amide I band shifts to lower wavenumbers (red-shifts), whereas the N-H bending mode appears at higher wavenumbers (blue-shifts). The N-H bending and C-N stretching vibrations are, however, highly coupled, whereas the amide I band is a rather pure mode and therefore highly sensitive to changes in the secondary structure. As such, the amide I band is often deconvoluted within condensed-phase IR spectra to derive the individual secondary structural content in peptides and proteins. Typical amide I positions characteristic for the respective secondary structural elements are summarized in **Table 2.1**.

Table 2.1: Correlations between common protein structures and amide I frequency[65-66]

Secondary structure	Amide I band (cm^{-1})
aggregated strands	1610-1628
β-sheet	1610-1640
random coil	1640-1648
a-Helix	1648-1660
3_{10}-Helix, type II *β*-turn	1660-1685
antiparallel *β*-sheet / aggregated strands	1675-1695

Classically, an IR spectrum is obtained by measuring the attenuation of a beam of light passing through a cuvette of length *l*, filled with molecules at a number density of *n*. The intensity of the transmitted light *I(ν)* is described by the Lambert-Beer law,

$$I(\nu) = I_0 \cdot e^{-\sigma(\nu) l n} \tag{2.5.1}$$

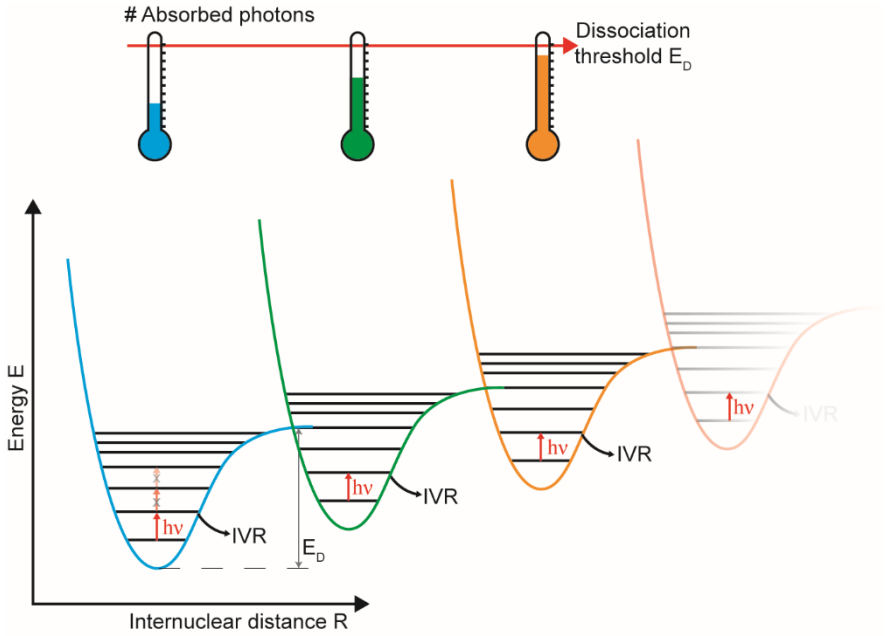


Figure 2.7: Simplified schematic of the infrared multiple photon dissociation (IRMPD) mechanism. The anharmonic nature of the electronic potential causes an internal vibrational redistribution (IVR) of the energy input of absorbed, resonant photons. The successive photon absorption therefore increases the internal energy of the molecule until the dissociation threshold E_D is reached and fragmentation occurs.

where I_0 is the intensity of the light before it passes through the cuvette filled with molecules and $\sigma(\nu)$ is the absorption cross-section of the molecules. However, for ions in the gas phase, the maximum ionic number density of $n_{SCL} \approx 10^6$ cm^{-1} at the space charge limit[67] is not sufficient to effectively attenuate the incoming light for direct absorption experiments. Instead, so-called action spectroscopy is employed, where the influence of the light on the molecular ion is monitored. In infrared multiple photon dissociation (IRMPD), a common methodology for action spectroscopy of gas-phase ions, the absorption of multiple resonant IR photons leads to fragmentation of the molecule. This fragmentation is not the result of a consecutive absorption process ($v = 0 \rightarrow v = 1 \rightarrow \ldots$), as the anharmonic nature of the electronic potential effectively shifts higher vibrational quantum numbers v out of resonance at a fixed wavelength of the incoming light (anharmonicity bottleneck) (**Figure 2.7**). Instead, the energy input of a single photon is immediately redistributed over vibrational background states, leading to a depletion of the excited vibrational mode. This intramolecular vibrational redistribution[68] (IVR) occurs on the sub-ns timescale, and therefore a successive absorption of multiple photons with energies in the range of 100 meV, followed by the redistribution of energy over the entire molecule, leads to a gradual increase in the internal energy up to the dissociation threshold E_D of the molecule. As the internal energy is statistically redistributed

over the entire molecule, the dissociation occurs through a low-energy pathway, *i.e.* fragmentation of the weakest bond. In the case of biomolecular complexes, the lowest-energy pathway often corresponds to the loss of a non-covalent subunit. Thus, in contrast to the fragmentation of a covalent bond in molecules, the dissociation of a molecular complex requires fewer photons and the resulting gas-phase IR spectrum is less influenced by anharmonic effects, which can cause shifts in the absorption bands or changes in relative intensities. A more detailed description of the involved processes can be found in the literature.[67]

The number of precursor ions after irradiation $N_p(\nu)$ can be described in analogue to the Lambert-Beer law,

$$N_p(\nu) = N_0 \cdot e^{-\sigma(\nu)F(\nu)} \tag{2.5.2}$$

where N_0 is the number of precursor ions before irradiation and $F(\nu)$ is the photon fluence. In contrast to classical absorption spectroscopy in the condensed phase, where a high number density of molecules is needed, action spectroscopy requires a high flux of photons. Throughout the thesis, the Fritz-Haber-Institut (FHI) free electron laser[69] (FEL) has been used to generate IR radiation of high intensity and therefore high photon fluence $F(\nu)$.

2.5.2 Fritz-Haber-Institut Free Electron Laser

The FHI-FEL[69] provides intense and tunable radiation in the mid-IR range (3 μm to 50 μm) with pulse energies of up to 200 mJ and a spectral width (FWHM) of ~ 0.5%. The laser medium of the FHI-FEL is a relativistic beam of electrons, which is generated in pulses by an electron gun and subsequently accelerated to energies of several MeV (in this thesis typically 37 MeV) by two linear accelerators. These relativistic electrons are further steered into the mid-IR undulator (**Figure 2.8**), which consists of an array of alternating pairs of oppositely poled permanent magnets, which forces the electrons into an oscillatory motion perpendicular to the orientation of the magnetic fields (due to the Lorentz force). This radial acceleration of relativistic electrons causes the emission of synchrotron radiation, which is stored inside the cavity by two highly-reflecting gold-plated mirrors at a distance equal to a multiple of the generated wavelength. This spontaneous emission is, however, usually very weak and incoherent, but it is increased by a factor of 10^6 - 10^8 due to the interaction of electron bunches from different passages. The interaction is based on an exchange of energy from the optical wave and the electron bunches, leading to a modulation of the electron density, which phases the emission and restores coherence. A hole inside one of the cavity mirrors allows for

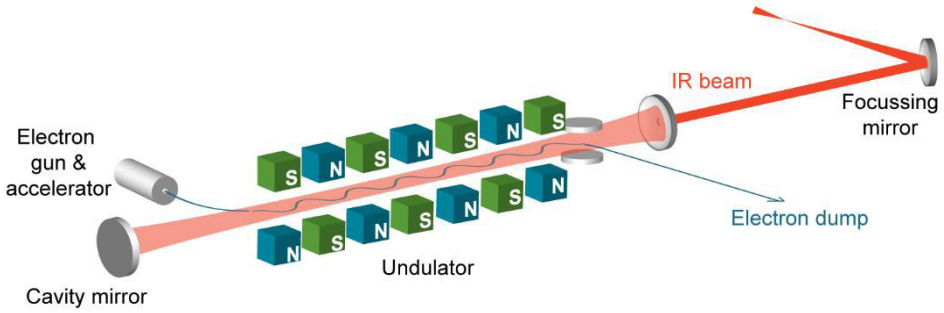


Figure 2.8: Schematic representation of an free electron laser (FEL). Electrons are generated by an electron gun and subsequently accelerated to energies of several MeV by linear accelerators. These relativistic electrons are further brought on oscillatory wiggle motion by an array of alternating pairs of oppositely poled permanent magnets. The emitted synchrotron radiation constructively interferes with the electron beam, leading to electron bunching and amplification of the emitted light. A small hole inside one of the cavities allows the outcoupling of the light, which is then brought to the experiments *via* an evacuated beamline. After generating IR photons, electrons are steered into an electron dump.

outcoupling of the light, which is then transported *via* an evacuated beamline to the experiments. The wavelength is typically manipulated by either changing the electron energy or by altering the strength of the magnetic field. The latter can easily be achieved by changing the distance gap between the oppositely poled permanent magnets. Thus, a frequency scan is accomplished by scanning the undulator gap.

2.6 Instrument and Data Acquisition

2.6.1 Linear Drift Tube Instrument

In this thesis, IM-MS and gas phase IR spectroscopy experiments have been performed on a home-built linear drift-tube instrument (iMob)[70] (**Figure 2.9**). The instrument is equipped with an offline *nano*-electrospray ionization (*n*ESI) source that produces ions. *n*ESI is known as a very soft ionization technique that allows the intact transfer of large and fragile biomolecules or complexes for which even weak non-covalent interactions can be preserved in the gas phase.[1, 71-72] Typically, 2-7 µL of sample is loaded into in-house prepared Pd/Pt-coated borosilicate capillaries, and the sample is further ionized by applying a high voltage (0.4-1.2 kV). After ionization, ions are radially confined and stored over several tens of milliseconds using a radio-frequency (RF) ion funnel (entrance funnel). From there, ions are pulsed into the 80.55 cm-long drift tube (resolution R ~ 40), where they travel under the influence of a weak electric field (10-20 V/cm) through helium buffer gas (~ 5 mbar). The

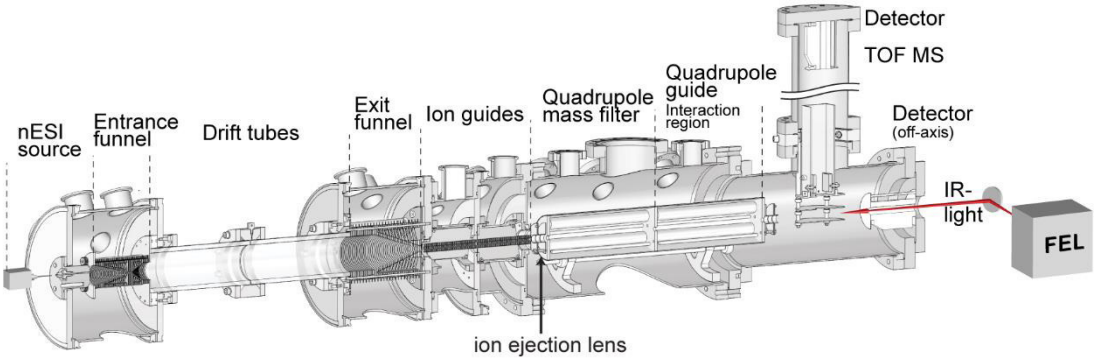


Figure 2.9: Schematic representation of the linear drift-tube ion mobility-mass spectrometer (*i*Mob) connected to the beam line of the FHI-FEL. Figure adapted with permission. Copyright 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

drift tube is constructed from four similar segments of conductive glass tubes (Photonis, USA), which feature a high and homogenous electric resistance over their entire length of 20.0 cm per tube. When a difference in electric potential is applied between both ends of the drift tube, a homogenous electric field inside the tubes is established and therefore allows the direct measurement of rotationally-averaged absolute CCSs. After exiting the drift tube, the diffusionally enlarged ion cloud is again radially confined by a second RF ion funnel (exit funnel) and further transported through two differentially pumped stages into the high vacuum of the instrument (cylindrical ring electrode ion guides operating at $p_1 = 10^{-2}$ mbar and $p_2 = 10^{-5}$ mbar). An electrostatic einzel lens then focuses the ions into a quadrupole mass filter (m/z range up to 4000) followed by a second quadrupole guide of identical dimensions, a second einzel lens stack and an additional set of steering lenses. Subsequently, ions are either detected by an off-axis electron multiplier detector (ETP Ion Detect, Australia, model AF632) or pulsed in a perpendicular direction into a time-of-flight (ToF) mass spectrometer with a Wiley-McLaren type setup (resolution R ~ 500). There, ions are also detected by an ETP electron multiplier (model AF882).

For IR experiments on drift-time and m/z-selected ions, optical access to the ion-light interaction region (quadrupole guide) is provided *via* a KRS-5 (Thallium Bromine-Iodide) window with an optical transparency of about 70% (5-10 μm). The tunable IR light is provided by the FHI-FEL (described in chapter 2.5.2)[69] and is transported from the FEL to the experimental setup through an evacuated beamline. For the accessibility of optical components, however, the last ~ 3 meters of the beamline are flushed with dry nitrogen in order to avoid the absorption of IR light by water and CO_2 in the ambient air.

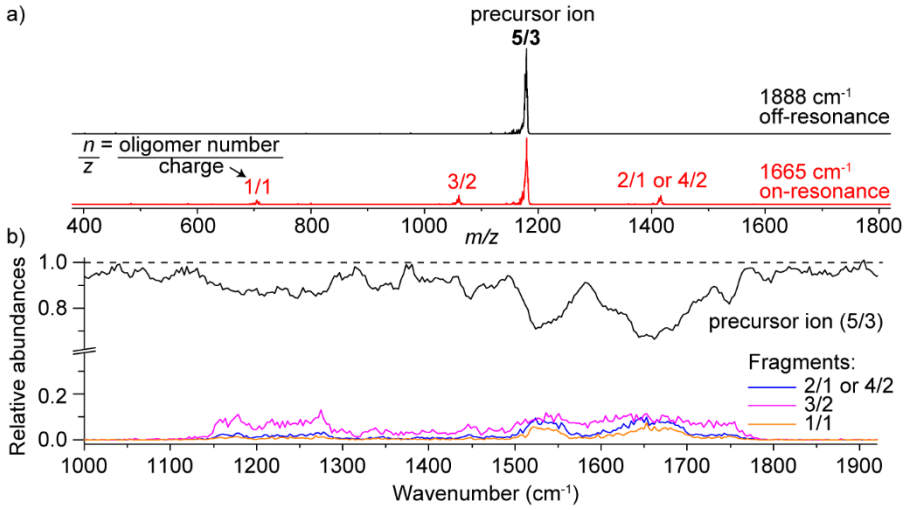


Figure 2.10: Data acquisition of a gas-phase IR spectrum for a triply protonated peptide pentamer (n/z = 5/3). a) Photofragmentation mass spectra irradiated with IR off-resonant photons at 1888 cm^{-1} (upper panel, black) and on-resonant IR photons at 1665 cm^{-1} (lower panel, red). b) Relative abundances of precursor and fragment ions as a function of irradiated IR wavenumber. The solid, black line depicts the depletion of the precursor ion, whereas colored lines represent the appearance of fragments.

2.6.2 Data Acquisition and Analysis

Manual data analysis was performed with OriginPro 2016G (64-bit) Sr2 (OriginLab Corporation) software. The linear drift-tube ion mobility-mass spectrometer (iMob) is controlled by a LabVIEW-based software and is able to record mass spectra, arrival-time distributions (ATDs) and gas-phase IR spectra. For each experiment, the apparatus is running in a different instrumental mode, which are described below:

Mass Spectra: In this thesis, the mass spectra of an electrosprayed sample are recorded by pulsing the ions into the ToF mass spectrometer and measuring their m/z- dependent flight time.

Arrival-Time Distribution: For the acquisition of ATDs of a specific species, the ions are first accumulated and stored within a trap and subsequently released into the drift tube by 150 µs-long pulses at a repetition rate of 10 or 20 Hz. After exiting the drift tube, ions are m/z-selected by the quadrupole mass filter, and their ATDs are recorded by measuring the time-dependent ion current of the m/z-selected species. From these ATDs, absolute collision cross sections ($^{DT}CCS_{He}$) of a particular ion species can be determined using the Mason-Schamp equation[62] as described in chapter 2.4.2.

Gas-Phase IR Spectra: For the acquisition of gas-phase IR spectra, a narrow drift-time window (80-200 µs width) is selected using electrostatic deflection at the first einzel lens

prior to mass selection by the quadrupole mass filter. This m/z- and ion mobility-selected ion cloud is then further irradiated by an intense (10-100 mJ) 10 μs pulse of IR photons within the interaction region of the instrument (quadrupole guide). When the IR wavelength is resonant with an allowed vibrational transition, the absorption of multiple photons leads to an increase in internal ion energy and can induce fragmentation. Photofragments are detected by the ToF mass analyzer. This is exemplarily shown in **Figure 2.10**a, where the ToF mass spectra of a small peptide oligomer that can form amyloid-like fibrils (YVEALL, see chapter 6) is depicted without (upper panel, black line) and with irradiation (lower panel, red line) with resonant IR photons.

The wavenumber of the FEL is typically scanned in steps of 3 cm^{-1} and 20-120 photofragment mass spectra are averaged at each step. The wavenumber-dependent abundances of precursor and fragment ions (I_{parent} and $I_{fragment}$) (**Figure 2.10**b) are then used to determine the fragmentation yield P at the specific wavenumber with equation 2.6.1. Division by the laser fluence $F(\tilde{\nu})$ (*i.e.* the laser power) accounts in first-order approximation for intensity variations of the FEL over the entire wavenumber range. In addition, the exact wavelength position of the FEL radiation depends strongly on the instrumental tuning and was therefore calibrated after each scan. The final IR spectrum is constructed by plotting the fragmentation yield P as a function of the wavelength-corrected wavenumber and averaging at least two individual scans.[67, 73]

$$P = -\log\left(\frac{I_{parent}}{I_{parent} + \sum I_{fragment}}\right)\frac{1}{F(\tilde{\nu})} \tag{2.6.1}$$

2.7 Protein Structure in the Gas Phase

Proteins have evolved over billions of years to yield a functional structure that is unique and easily accessible under physiological conditions. The folding into the native structure is supported by multiple factors such as the hydrophobic effect, in which non-polar residues associate in the protein interior while they simultaneously exclude water molecules. However, this effect is absent in the gas phase and hydrophobic residues can thus expose themselves to the non-polar environment of the gas phase, forming the so-called inside-out structures. Consequently, if an *in vacuo* equivalent to the native structure exists, it will be somewhat altered to account for the new conditions of the vacuum environment. In addition, repulsive Coulomb forces can occur when charged residues are not solvated, whereas stabilizing

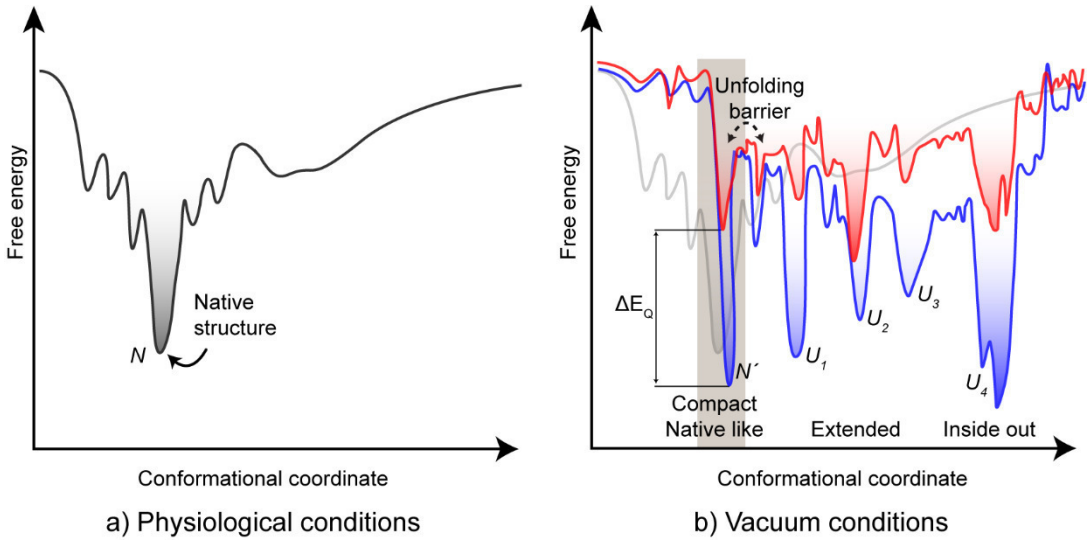


Figure 2.11: Schematic representation of free energy surfaces that reflect protein folding under a) physiological and b) vacuum conditions. The native, solution structure *N* will be altered in the absence of solvent, leading to an *in vacuo* native-like structure *N´*. For highly charged species (red) the Coulomb repulsion (ΔE_Q) destabilizes the native fold *N´* and drives highly charged ions into unfolded, extended states (U_1, U_2, etc.). Low-charge ions (blue) are separated by higher energy barriers from other local energy minima and are therefore often entropically trapped. The global energy minimum of a protein in the gas phase may represent a hypothetical inside-out conformation.

electrostatic interactions, such as hydrogen bonds or salt bridges, are enhanced in the gas phase. While the free energy surface of a protein follows a steep funnel with low energy barriers into a single and unique native fold in solution (**Figure 2.11**a), multiple local minima coexist in the gas phase[10, 74] (**Figure 2.11**b). Despite these fundamental differences between solution and gas phase conditions, there is evidence that native structural elements are maintained in the gas phase at least for very large systems such as viruses,[75-76] proteasomes,[77-78] membrane proteins,[79-80] or very complex molecular machines such the V-type ATP synthase[81]. The solution-phase conformation of small peptides can also be preserved when the respective activation barrier for an unfolding event is high as in the case of the *cis-trans* isomerization of a proline peptide bond.[82-83] For medium-sized proteins the situation is not clear, but compact structures can be stabilized in the gas phase by the non-covalent attachment of crown ethers to protonated lysine side-chains. The crown ether micro solvates the charge and thereby oppresses their destabilizing backfolding on the protein backbone.[84-85]

It is assumed that *in vacuo* native-like structures are kinetically trapped in the gas phase but they can, especially when highly charged, easily overcome the unfolding barrier due to repulsive Coulomb destabilization and yield extended structures. Thus, the charge state

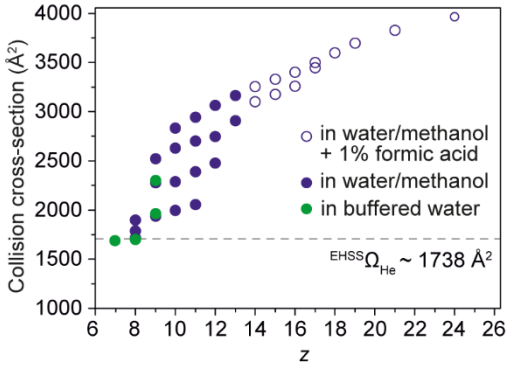


Figure 2.12: CCS of myoglobin as a function of charge. Solid green circles represent ions sprayed from buffered aqueous solution, solid blue circles represent ions from water/methanol solution, and open blue circles represent ions from water/methanol solution with 1% formic acid.

distribution of a protein already indicates the preservation of structural, solution-phase features. When a protein, such as myoglobin, is transferred from buffered conditions into the gas phase, a narrow and low charge state distribution (+7 to +9) of the protein holo form is observed. These species usually adopt a single and compact conformation with sizes expected for the native structure (**Figure 2.12**). In contrast, when the same protein is sprayed from denaturing conditions (water/methanol) a broad charge state envelope ranging from +8 to +13 is observed along with the presence of multiple conformations (partially unfolded) and a large jump in CCSs due to Coulomb repulsion. At very high charge states (+14 to +24, sprayed from acidic water/methanol) the protein further undergoes a gradual unfolding with increasing charge. Thus, the underlying conformation of all proteins highly depends on the amount of charges the system bears.[86]

IR spectroscopy revealed that proteins carrying multiple charges typically undergo a Coulomb-induced unfolding to adopt artificial gas-phase helical structures.[87-89] In contrast to β-sheets, these helices are significantly more stable in the gas phase[90-91] because they can maintain most of their hydrogen bonds while maximizing the distances between equal charges. However, even these helices unravel at the highest charge states and form extended string-like structures that are governed by Coulomb repulsion and a C_5-hydrogen bonding network.[87-88]

2.8 Amyloid Intermediates in the Gas Phase

In the clean-room environment of the gas phase not only native-like structures can be preserved but also a single species can be isolated and individually characterized in the presence of many others without affecting the underlying equilibrium. These inimitable advantages make gas-phase techniques an ideal tool to study the aggregation process of amyloidogenic peptides and proteins as exemplarily shown for the hIAPP hormone, which is linked to the etiology of type II diabetes[92], and the non-toxic rat IAPP (rIAPP) isoform, which only differs by six amino acids[93]. IM-MS data show that the hIAPP monomer $n/z = 1/4$ (n = oligomeric number, z = charge) exists in two major conformations, whereas the rIAPP

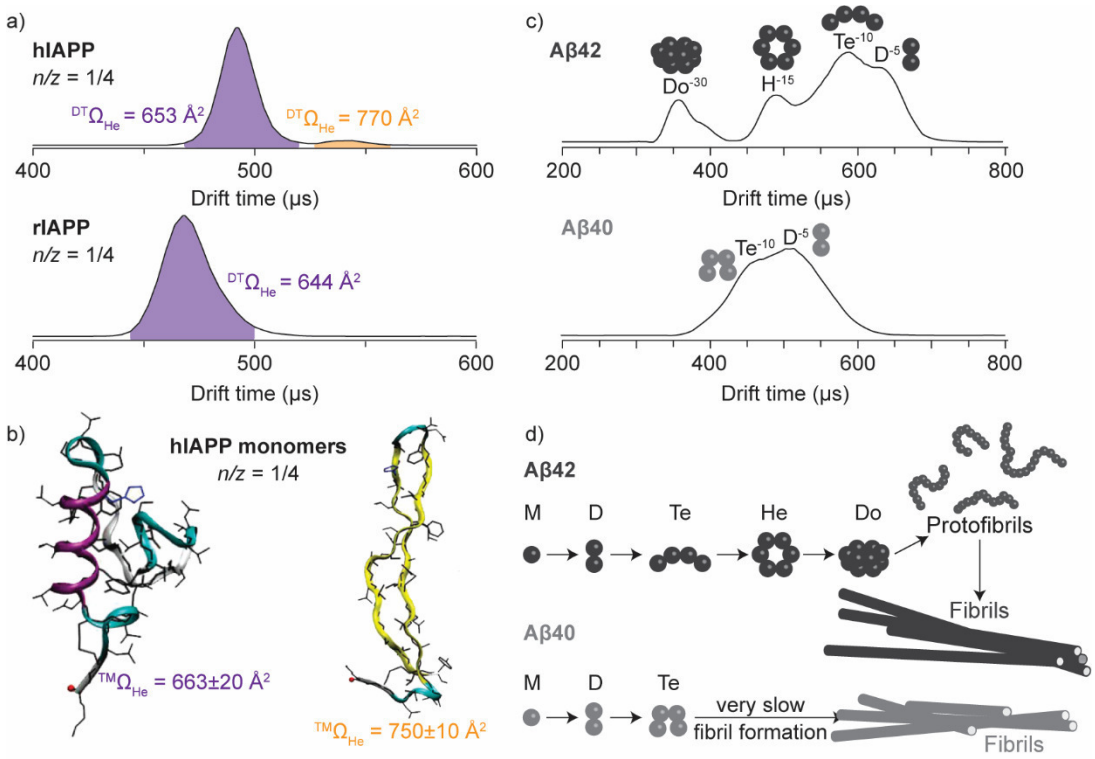


Figure 2.13: Proposed assembly pathway of IAPP and the Aβ peptide. a) The ATD of the quadruply charged hIAPP monomer (n/z = 1/4) shows two major conformations, whereas only one compact structure is observed for the non-fibril forming rIAPP sequence. b) Theoretical calculations have shown that the experimental $^{DT}CCS_{He}$ of the compact hIAPP monomer is in good agreement with a helix-coil structure, whereas the extended conformation represents a β-hairpin structure. Since the β-hairpin is only present for the fibril forming hIAPP it might be an amyloidogenic precursor. c) The ATD of the Aβ42 peptide at m/z = 1807 shows four features whereas for the Aβ40 peptide only two features are observed. d) From this result it was concluded that Aβ42 forms an "open" tetramer, where a dimer can easily attach to form a hexamer, which then can stack together to form dodecamers. Higher Aβ42-oligomers are not observed, and the dodecamer is suggested to seed the formation of protofibrils and fibrils. The Aβ40 peptide, however, forms a "closed" tetramer, where no further oligomers can easily attach to. As a consequence the Aβ40-fibril formation proceeds very slowly. (Adapted with permission from *J. Am. Chem. Soc.* **2009**, *131*, 18283. Copyright 2009 American Chemical Society. Adapted by permission from Macmillan Publishers Ltd.: *Nature Chemistry* **2009**, **1**, 326, copyright 2009.)

monomer only forms one compact structure (**Figure 2.13**a).[94] Theoretical calculations suggest a helix-coil conformation for the compact hIAPP monomer whereas an extended β-hairpin structure was proposed for the extended hIAPP monomer (**Figure 2.13**b). Both conformations of hIAPP are highly pH-dependent, with the extended β-hairpin conformation being favored at high pH. CD spectroscopy indicates that faster transition to β-sheet-rich structures occurs at high pH level, and TEM further confirmed the formation of fibrils for hIAPP.[95] Therefore,

it was suggested that the extended β-hairpin hIAPP monomer may serve as an amyloidogenic precursor from which the formation of a β-sheet rich dimer is initiated.[96]

Early Aβ oligomers, which are thought to be responsible for the neurodegeneration of Alzheimer´s disease (AD) patients, have also been analyzed in the gas phase.[97] Generally, full-length Aβ42 aggregates faster and is more cytotoxic than the shorter Aβ40 variant.[98] These differences have been explained in detail on basis of IM-MS-derived structural data.[52] Aβ40 assembles into a compact tetramer (**Figure 2.13c** and d), where further peptide monomers cannot easily attach to and, in consequence, no higher oligomers are formed. Aβ42, on the other hand, forms an "open" tetramer, where a dimer can attach to form a hexamer. Two of these hexamers in turn can stack together to form the presumably toxic species in AD, the dodecamer [99]. This assembly pathway was first proposed on basis of IM-MS data measured in negative ion mode, but an unambiguous verification based on the isotopic pattern is still missing due to the low mass resolution. The assignment is, however, further supported by data obtained from photo-induced cross-linking experiments of unmodified Aβ42, size exclusion chromatography and dynamic light scattering.[100] Recently, AFM was further used to directly monitor the assembly of Aβ42. These data indicate that dodecamers are on-pathway oligomers, which seed fibrillar aggregates.[101] Furthermore, the neurotoxicity of Aβ-oligomers was shown to correlate well with their H/D exchange (HDX) pattern.[102] Aβ40 and Aβ42 oligomers which persist HDX, *i.e.* they are at least partially folded and therefore experience slower HDX, are more cytotoxic than aggregates, which rapidly undergo HDX. In addition, a hydroxyl radical-based, fast photochemical oxidation approach indicates that Aβ42-fibril formation is predominantly driven by the central hydrophobic core (residues 16-27) as well as the C-terminus (residues 28-42), whereas the N-terminus remains almost unstructured.[103]

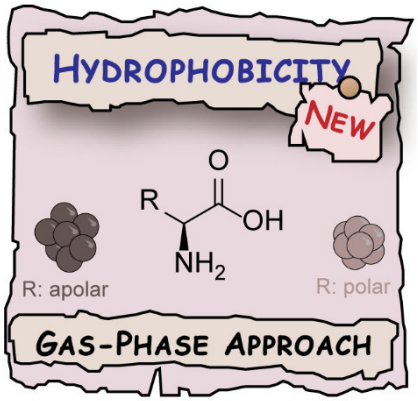
The formation of Aβ40- and Aβ42-oligomers has also been studied on basis of IM-MS data measured in positive ion mode. At high concentrations, the formation of Aβ40 oligomers up to the 13-mer can be observed.[104] Aβ42 on the other hand, predominantly forms dimers and trimers, which, in the presence of sodium dodecyl sulfate, can further evolve into higher oligomers such as pentamers and hexamers.[105] The variance to the above-described negative ion data may be influenced by diverse factors such as different ion polarity and instrumentation, as well as variations in the sample preparation methodologies. Therefore, further studies are necessary to verify the altered assembly behavior.

In addition, gas-phase methods can screen for potential drugs which inhibit fibril formation, and can as well simultaneously provide further valuable information regarding the

binding modes. IM-MS showed that small heat-shock proteins can specifically prevent the aggregation of Aβ42 and reduce the toxicity in a dose-dependent manner.[106] Molecular tweezers such as CLR01 use electrostatic interactions to bind to lysine 16 and 28 in Aβ42.[107] This binding affects the structure of the "open" Aβ42-tetramer, altering it to a more compact conformation akin to the "closed" Aβ40 tetramer. As a consequence, the formation of higher-order oligomers is inhibited.[108] Other IM-MS-based studies showed that Cu(II) ions bind to the histidine 18 residue of hIAPP, leading to a stabilization of off-pathway oligomers hence suppressing the hIAPP fibril formation.[109-111] A conformer-specific inhibition for hIAPP aggregation was furthermore reported for the molecules silibinin[9] and insulin[112]. Silibinin exclusively binds to the extended β-hairpin hIAPP monomer (amyloidogenic precursor), and as a result the oligomerization is arrested. In contrast, the helical-rich insulin monomer is proposed to stabilize hIAPP *in vivo* by interacting with the compact hIAPP monomer. The interaction shifts the hIAPP monomer equilibrium to the compact helix-coil conformation and thereby reduces the abundance of the amyloidogenic precursor.

Most studies on amyloid intermediates are based on IM-MS data and only a few use orthogonal methods such as gas-phase IR spectroscopy,[16-17] which is capable of providing additional, complementary and therefore more unambiguous information on the secondary structure. In this thesis, both techniques have been combined to not only study biomolecular assemblies of small amino acids, but also aggregation of short peptides into amyloid fibrils.

3 Novel Hydrophobicity Scale for Amino Acids*



3.1 Introduction

The accurate determination of the intrinsic hydrophobicity of each amino acid is crucial for understanding key aspects of biology. Side-chain hydrophobicity affects many fundamental biological processes, including the folding,[113-114] stability[115-116] and oligomerization/aggregation[117-118] of proteins as well as ligand-protein interactions[119-120]. Thus far, more than 100 hydrophobicity scales[121-122] have been established, with most derived from condensed-phase methods such as partitioning (1),[123-124] accessible surface area calculations (2),[125] the direct measurements of physical properties[126] (3) and chromatographic techniques[127] (4). However, there are significant differences among these scales as they utilize markedly different techniques and/or vary the type of species investigated, *e.g.* peptides/proteins or amino acids and their derivatives. As a result, a comprehensive ranking of the relative hydrophobicity of amino acids is precluded.[125, 128]

A more detailed assessment of hydrophobicity measurements reveals the limitations of current approaches. Scales based on partitioning (1) use organic solvents such as octanol to mimic the protein interior and rank tryptophan as the most hydrophobic amino acid.[123-124] Organic solvents are, however, often slightly miscible with water, and consequently the characteristics of both phases change. This effect makes it difficult to obtain an intrinsic

* This chapter is based on the work published in W. Hoffmann, J. Langenhan, S. Huhmann, J. Moschner, M. Accorsi, J. Seo, J. Rademann, B. Koksch, M. T. Bowers, G. von Helden and K. Pagel, *Angew. Chem. Int. Ed.*, **submitted for review**.

hydrophobicity scale. On the other hand, surface area calculations (2) utilize a database of protein crystal structures and define the hydrophobicity as the tendency of a residue to be found inside of a protein rather than on its surface.[125] Here, cysteine is ranked as the most hydrophobic amino acid, because its thiol group can form disulfide bonds that are frequently present inside a globular structure. In contrast, proline participates in turns located at the protein surface and is therefore classified as very hydrophilic. The most popular scale based on physical properties (3) was developed by measuring the surface tension values of amino acid solutions in reference to a glycine solution.[126] Here, leucine is reported as the most hydrophobic amino acid because it yields the largest decrease in surface tension. On the other hand, proline, arginine, and lysine adopt a different ionic configuration at their isoelectric points compared to glycine. This introduces further discrepancies compared to hydrophobicity scales based on other methods. In addition, the precise measurement of the surface tension values depends on the solubility of the respective amino acids and therefore only an extrapolated value is provided for tyrosine. Chromatographic techniques[127] (4), on the other hand, use amino acid derivatives or model peptides to define the hydrophobicity as a change in retention time relative to the glycine-substituted analog. In the case of the model peptide approach, a change in peptide sequence,[127, 129] peptide length[129] and substitution position[130] can have a dramatic impact on the hydrophobicity values. Upon proline substitution, a secondary structural change is often induced, and the value of proline is therefore not comparable to hydrophobicity values of other amino acids.[127-128, 130] In addition, the choice of pore diameter, aqueous buffer pH and temperature, or bonding density of the stationary reverse phase alkyl chains can influence the hydrophobicity scale.[122] Thus, it is suggested that more than 100 peptides should be sampled in order to obtain a reliable scale.[131-132]

Overall, the most common hydrophobicity scales do not allow a universal comparison and classification of amino acids because they include contributions from solvent effects, and the implementation of new amino-acid derivatives can be very time-consuming. In this work, this problem is overcome by studying the pairwise interaction of all canonical amino acids in the "clean-room" environment of the gas phase, which has a relative permittivity ($\varepsilon_r = 1$) suitable to mimic the protein interior[133] (ε_r = 6-7). The rotationally-averaged CCS Ω of an ion as measured by IM-MS experiments (see chapter 2.4) is a direct molecular property that is largely independent from instrumental parameters such as chromatographic settings, temperature, pressure or solvent conditions. Thus, in this work, amino-acid clusters are

generated utilizing *n*ESI and the relationship between amino-acid cluster size and side-chain hydrophobicity is explored.

3.2 Experimental Details

Materials: Fluorinated amino acid derivatives were kindly provided by Susanne Huhmann, Johann Moschner and Matteo Accorsi (Freie Universität Berlin, Berlin). Other amino acids and solvents are purchased from Sigma Aldrich (Taufkirchen, Germany) and used without further purification.

Ion Mobility-Mass Spectrometry: For IM-MS experiments, amino acids were dissolved in water to obtain a 5 mM solution (except for tyrosine: 1 mM) and the experiments have been performed on a home-built drift-tube mass spectrometer, which is described in detail in chapter 2.6.[70] $^{DT}CCS_{He}$ have been calculated according to the procedure described in chapter 2.4.2.

Spatially Isotropic Self-Assembly with Changing Density: The theoretical CCS evolution Ω_n for a spatially isotropic peptide cluster with oligomer size n is described in chapter 2.4.3 and is based under the assumption of a constant cluster density.[63] However, for a spherical growth with changing cluster density a correction factor $a^{2/3}$ can be introduced:

$$\Omega_n = \sigma_1 \cdot [n\alpha]^{\frac{2}{3}} \tag{3.1}$$

The size of the amino acid meta cluster highly depends on the hydrophobicity of the respective amino acid. Hydrophobic residues, like isoleucine or phenylalanine, form extended oligomers which are larger in size ($a > 1$) than expected for an idealized spherical growth (for $a = 1$) of the oligomers. Hydrophilic amino acids (serine, asparagine), on the other hand, assemble into more compact oligomers ($a < 1$). Thus, a scale for the hydrophobicity a of a single amino acid can be obtained by fitting equation 3.1 to the experimental ion mobility data of the meta clusters. In order to keep the fitting consistent for each amino acid and to avoid gas-phase artifacts due to charge repulsion, only singly- and doubly protonated meta clusters have been included in the fit.

3.3 Results and Discussion

Figure 3.1a, b shows *n*ESI mass spectra of leucine and arginine sprayed from pure water. Leucine assembles into a dimer $n/z = 2/1$ along with larger clusters starting from an octamer up to a 36-mer $n/z = 36/4$, where n stands for the number of leucine units in the cluster and z for the charge. The more polar arginine, which carries a guanidine moiety, behaves differently – it assembles into cluster in a more consecutive manner, where clusters up to the 24-mer are observed. Other amino acids assemble in a similar fashion (see Appendix A and literature[134-136]). They predominantly follow a smooth intensity distribution with cluster size, forming spherical-like clusters.

The CCSs as a function of the oligomer number n as measured by IM-MS are shown in **Figure 3.1** for leucine (c) and arginine (d). The uncertainties in the measured CCSs are smaller than the actual size of the corresponding symbol. The black solid line represents theoretical isotropic growth[63] (chapter 2.4.3), *i.e.* growth of an idealized spherical assembly. From a visual inspection, it appears that leucine forms more extended clusters than predicted by theoretical isotropic growth, whereas the more polar arginine assemble into more compact oligomers. Their packing efficiency seems not to depend on the size of the side-chain (Ω_{Leu} = 66 Å^2 vs. Ω_{Arg} = 72 Å^2). This finding points towards the fact that cluster formation is strongly influenced by the polarity of the side chain. A similar dependence of cluster growth on side-chain polarity was recently observed[136] for serine, tryptophan, phenylalanine, asparagine as well as isoleucine and has been more systematically confirmed for all canonical amino acids in this work (Appendix A). From this data set, it is clear that hydrophobic side chains result in larger clusters as they orient themselves towards the low permittivity of the gas phase, whereas polar side chains adopt more compact structures as they seek to maximize intramolecular interactions.

Now a new correction factor a is introduced to account for the deviation of cluster growth from the theoretical isotropic growth according to the equation 3.1. This a value provides a measure for the packing efficiency and is correlated to the hydrophobicity of each side chain. Values for $\alpha > 1$ represent hydrophobic amino acids, whereas $\alpha < 1$ stands for more hydrophilic side-chains. As such, it is used as a basis to generate a new intrinsic hydrophobicity scale for amino acids. The charge states 1+ and 2+ are used to derive a, whereas higher charge states are excluded in order to circumvent the influence of Coulomb repulsion on the packing efficiency and to ensure comparable datasets for all amino acids.

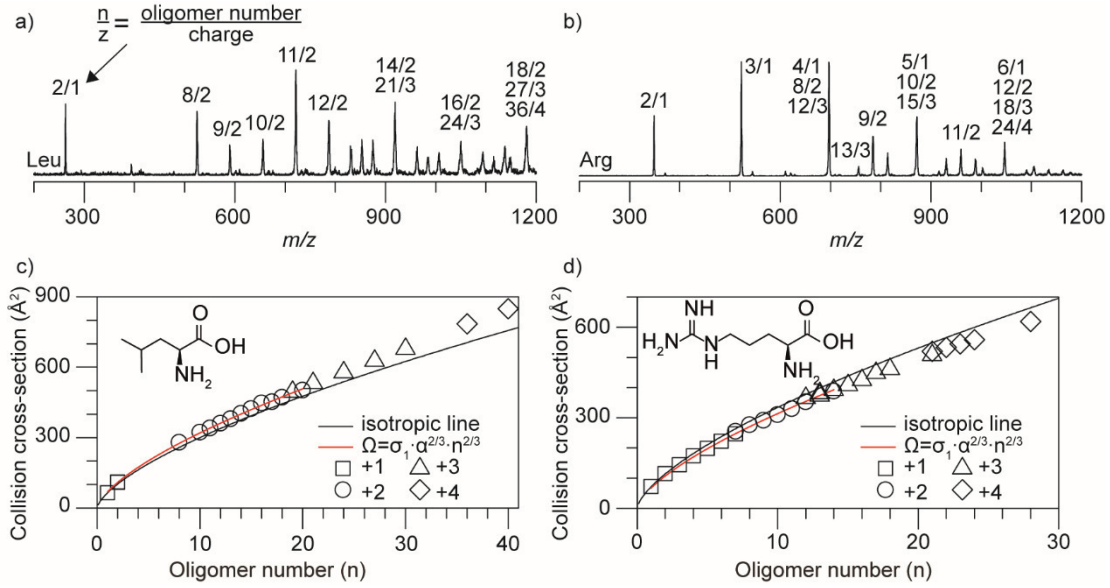


Figure 3.1: Mass spectra and CCSs Ω for leucine and arginine. a, b) Mass spectra obtained when spraying aqueous 5 mM concentrated amino acid solutions. Prominent peaks are labeled with their n/z ratio, where n represents the number of amino acid units in the cluster and z the charge. c, d) CCSs as a function of the oligomer number n. The solid black line represents their theoretical isotropic growth, *i.e.* growth of an idealized spherical assembly, whereas the red line shows the fit to derive the respective hydrophobicity value α.

A summary of the α values as a function of the size of an amino acid is given in **Figure 3.2** and **Table A34**. The here-introduced hydrophobicity scale ranks the natural amino acids leucine > valine ≈ methionine > isoleucine > phenylalanine as most hydrophobic, which is in good qualitative agreement with previous scales[123, 127] classifying these also as very hydrophobic. In addition, the new scale indicates a plausible relative ranking of amino acids: 1) Phenylalanine is more hydrophobic than tyrosine, which carries one additional hydroxyl group at the phenyl group. 2) Serine contains a primary alcohol and is therefore more hydrophilic than threonine, which carries a secondary alcohol. 3) Glutamine is more hydrophobic than asparagine due to the longer aliphatic chain, whereas both glutamine and asparagine are more hydrophobic than their corresponding carboxylic acid analogues (glutamic acid and aspartic acid). 4) Lysine and arginine have a long aliphatic chain with a polar functional group at its end in common, either a guanidine group or a primary amine. The guanidine group is, however, more polar, and consequently arginine has a lower α value than lysine. Interestingly, lysine shows neither a very polar nor a hydrophobic character within the here-presented scale. This result contradicts previous condensed-phase scales[123, 125, 127] which rank lysine as one of the most polar amino acids. In solution the lysine side-chain is predominantly protonated, whereas

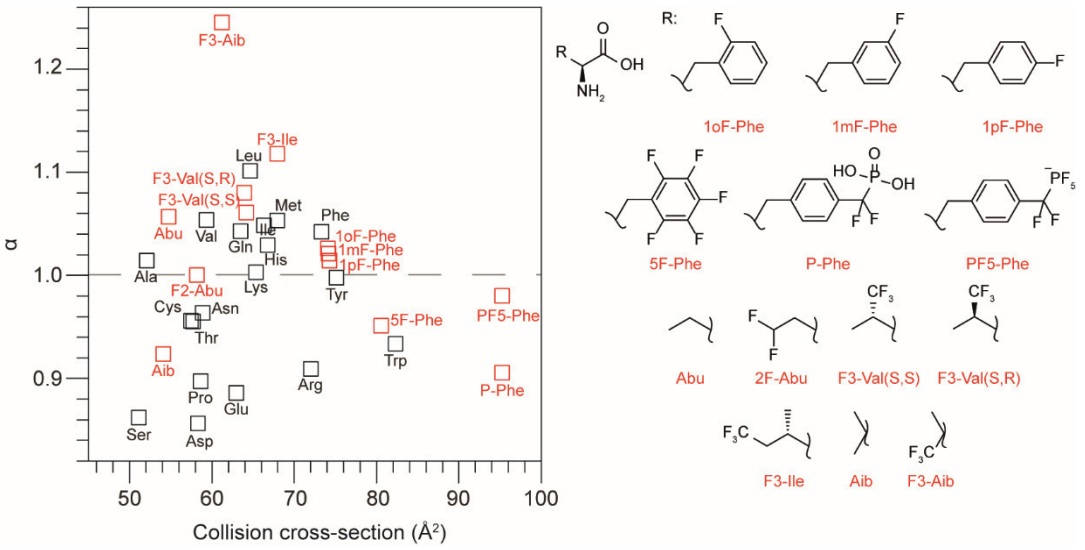


Figure 3.2: Relative hydrophobicity scale for amino acids. The hydrophobicity α as a function of amino-acid size (given as the monomer CCS **Ω**). Values for α > 1 represent more hydrophobic amino acids, whereas α < 1 stands for hydrophilic side chains. Non-natural variants are shown in red.

in the here-presented gas-phase approach the intrinsic hydrophobicity of the neutral side-chain is examined. Thus, the long aliphatic chain counteracts the hydrophilic character of the neutral amine and thereby yields an α value of ~ 1 for lysine. This relative ranking for lysine depicts more accurately the intrinsic nature of the side-chain.

The quantitative comparison of the here-presented hydrophobicity scale relying on *a* with scales based on condensed-phase approaches are shown in **Figure 3.3**, where Pearson correlation coefficients R are displayed as a heat map. A coefficient R value of 1 (red) or -1 (blue) indicates a linear correlation where all data points lie on a line, whereas a R value of 0 (green) implies no linear correlation between two hydrophobicity scales. A very high ($|R| > 0.6$) linear correlation between *a* and other hydrophobicity scales is not observed, as they are based on vastly different approaches (gas-phase vs. condensed-phase). The contribution of solvent effects, the dependence on the type of the investigated species (*e.g.* different peptides) and dependence on parameters such as pH, chromatographic equipment as well as solubility influence hydrophobicity values within condensed-phase approaches and thereby contribute to the pattern of the correlation matrix. Nevertheless, the data indicate a positive relationship ($|R| > 0.4$) between the relative ranking of hydrophobicity values based on *a* and other scales, and therefore further support that the hydrophobicity of an amino-acid side chain predominantly dictates the packing efficiency in clusters. The ion polarity (protonated or deprotonated amino acid clusters), however, slightly influences the fit outcome of α as exemplarily shown for arginine, glutamic acid, isoleucine, phenylalanine and serine (see

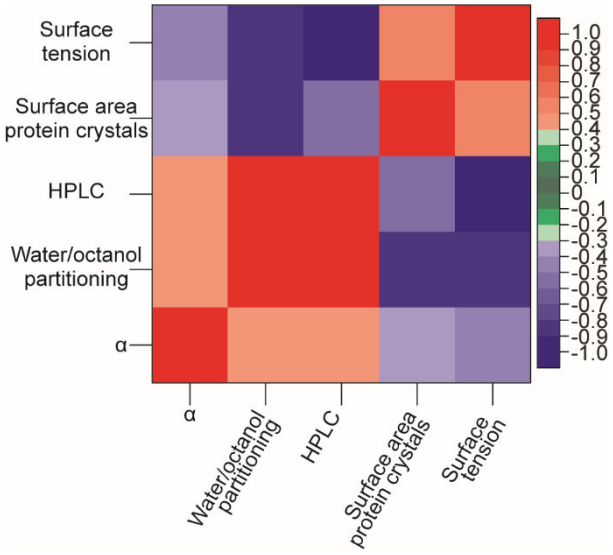


Figure 3.3: Heat map of the Pearson correlation coefficients R obtained by comparing the here-presented hydrophobicity scale α and scales based on condensed-phase approaches. A positive or negative linear correlation is shown in red and blue, respectively. No linear correlation is indicated in green.

Table A34). In general, marginal higher α values are observed for deprotonated clusters. Thus, the respective α values of protonated and deprotonated amino acid clusters should be not merged to guarantee an accurate relative hydrophobicity scale for each amino acid.

The robustness of the new approach for the implementation and classification of non-natural fluorinated amino-acid derivatives was further evaluated. Fluorine substitution is a standard strategy to modulate the properties of pharmaceuticals[137] and peptides/proteins[138]. Its impact on folding properties is a subtle interplay of a new interaction partner, changes in hydrophobicity and size.[139-140] Hydrophobicity values for a small set of fluorinated amino acid analogous of isoleucine, leucine and phenylalanine as well as of 2-aminobutyric acid (Abu) and 2-aminoisobutyric acid (Aib) are shown in red in **Figure 3.2**. In general, CF_3-fluorination in aliphatic side-chains increases the hydrophobicity compared to the analogues (isoleucine, valine and Aib). The CF_3-substutition in isoleucine, however, only marginally alters the overall size (Ω_{Ile} = 66 Å^2 vs. $\Omega_{F3\text{-}Ile}$ = 68 Å^2), whereas for valine and Aib an increase of ~ 8-12% in CCS is observed (Ω_{Val} = 59 Å^2 vs. $\Omega_{F3\text{-}Val}$ = 64 Å^2 and Ω_{Aib} = 54 Å^2 vs. $\Omega_{F3\text{-}Aib}$ = 61 Å^2). Interestingly, the fluorinated diastereomers of F3-Val yield different hydrophobicity values. F3-Val(S,S) is less hydrophobic than F3-Val(S,R) but both are more hydrophobic than valine. This observation is in good agreement with a previous result obtained by HPLC[118] as well as theory[141], and it indicates that the here-presented approach is sensitive to small variations within a given structure.

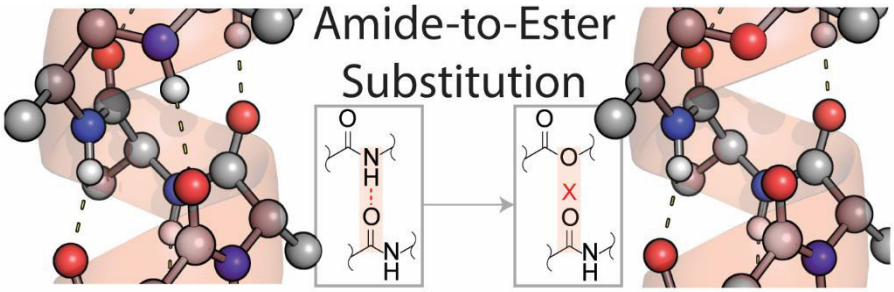
A CF_2-fluorination, on the other hand, lead to a different result – 2F-Abu exhibits a smaller *a* value than its non-fluorinated analogue. Thus, partial fluorination in aliphatic side-chain can decrease the hydrophobicity for a given amino acid.[140] The prediction of such hydrophobicity properties upon fluorination is not trivial, but amino acids can be readily classified using the here-presented approach.

The incorporation of fluorine into phenyl rings differs from the results obtained for aliphatic side chains. Here, the H/F substitution reduces the hydrophobicity in following order: phenylalanine > 1mF-Phe > 1oF-Phe > 1pF-Phe > 5F-Phe. The substitution alters the electronic properties of the phenyl ring and the introduced dipole moment leads to an increase in polarity. This effect is even more pronounced when a phosphonate group (R–CF_2–$PO(OH)_2$) is attached to the phenyl ring (phenylalanine vs. P-Phe). However, when the phosphonate group (P-Phe) is perfluorinated to yield a hypervalent R–P^-F_5 group that carries one permanent negative charge on the side-chain (PF5-Phe), an increase in hydrophobicity is observed compared to the neutral phosphonate group in P-Phe.

3.4 Conclusions

In summary, the packing efficiency of amino acid clusters in the gas phase strongly depends on the polarity of the respective side-chain. Hydrophobic residues form more extended clusters in which their apolar side-chains reach into the "clean-room" environment of the gas phase. Polar amino acids, on the other hand, form more compact clusters than predicted for an idealized spherical growth. Therefore a a new correction factor a was introduced that takes the deviation from the theoretical isotropic growth $\Omega = \sigma_1 \cdot n^{2/3} \cdot a^{2/3}$ into account, and this factor is used to establish a novel IM-MS-based hydrophobicity scale for amino acids. This new scale provides an easy way to assess the hydrophobicity of non-natural amino acid derivatives in the absence of solvent, allowing a more reliable and universal comparison and classification of amino acids. In addition, a good qualitative and quantitative agreement between the here-introduced scale and previous condensed-phase based scales is observed.

4 Assessing the Stability of Alanine-based Helices*



4.1 Introduction

Several decades of protein-structure investigation have led to an in-depth understanding of the effects that stabilize α-helices in the condensed phase.[142-145] Also in the "clean-room" environment of the gas phase helical peptides have been studied extensively.[146-148] In particular, polyalanine-based peptides turned out to be highly suitable systems due to their simplicity and strong tendency to adopt helical structures. It has been shown that if the Ala_n peptide chain is C-terminated by a lysine, peptides as short as $[Ac\text{-}Ala_nLys + H]^+$ ($n > 7$) are able to form stable α-helices in the gas phase.[149] Furthermore, water adsorption measurements,[150] IM-MS,[151] IRMPD spectroscopy[152] as well as theoretical methods[149, 152-153] showed that the protonated peptide $[Ac\text{-}Ala_{10}Lys + H]^+$ exists solely as an α-helix in the gas phase.

Three factors stabilize $[Ac\text{-}Ala_{10}Lys + H]^+$ in the α-helical conformation (**Figure 4.1**): (i) the attractive interaction between the protonated lysine and the helix macro-dipole, (ii) the capping of the protonated lysine residue to dangling carbonyls at the C-terminal end of the helix by protonated lysine, and (iii) the cooperative network of H-bonds in the helix. Multiple studies have been designed to quantify these[90, 149-151, 153-154] and other[91, 152, 155-157] factors. In particular, Jarrold and co-workers showed that simply by shifting the lysine residue from C- to N-terminus eliminates the attractive interaction between the charge and the helix macro-dipole, which substantially destabilizes the α-helix.[90, 151, 154] As a result, these peptides adopt more compact structures, which exhibit significantly shorter drift times in IM-MS

* This chapter is based on the work published in W. Hoffmann, M. Marianski, S. Warnke, J. Seo, C. Baldauf, G. von Helden and K. Pagel, *Phys. Chem. Chem. Phys.*, **2016**, *18*, 19950-19954, DOI: 10.1039/c6cp03827a. Figures and content adapted with permission. Copyright 2016 Royal Society of Chemistry.

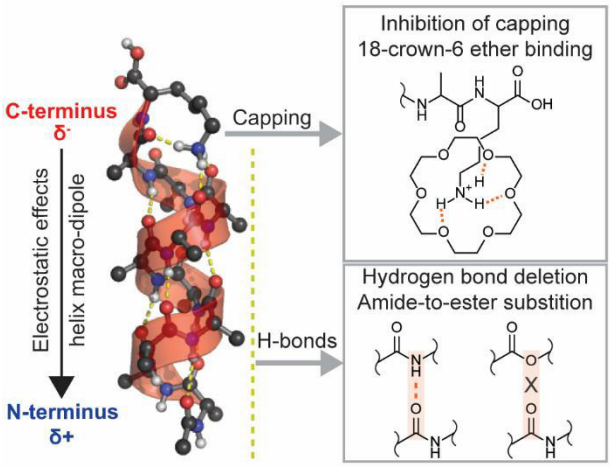


Figure 4.1: Three factors that stabilize $[Ac\text{-}Ala_{10}Lys + H]^+$ in the α-helical conformation in the gas phase. Capping of the lysine side-chain residue to carbonyl groups can be turned off by attaching 18-crown-6 ethers. H-bonding can be selectively deleted *via* amide-to-ester substitutions.

experiments. In another study, Kim et al. coordinated the protonated lysine side chain with 18-crown-6 ether in order to inhibit its H-bonding to the dangling carbonyl groups near the C-terminus.[154] Surprisingly, it is observed that the isolation of the charge enhances the peptide helicity, presumably due to the elimination of geometry constraints posed by the H-bonds with the lysine amine group. The influence of the third factor, however, the network of backbone H-bonds on the stability of the helix, so far remains less understood. In this work, the impact of individual H-bonds in $[Ac\text{-}Ala_{10}Lys + H]^+$ (AK00) on the conformational stability was studied by introducing amide-to-ester substitutions at specific positions within the sequence.[156, 158-160] The replacement of an alanine building block by an isostructural lactic acid unit results in formation of an ester bond (often referred to as a depsi peptide bond, **Figure 4.1**). This replacement of an N-H donor group by an oxygen atom selectively turns off one specific hydrogen bond while the dipole moment associated with the peptide bond remains largely unaltered (1.7 D vs. 3.5 D for the ester and regular amide bonds, respectively).[161] The position of the depsi peptide bond is systematically varied from the C- (AK01) to the N-terminus (AK10), with the exception of the peptide bond between the Lys and the Ala residue. The resulting peptides are characterized by a combination of experimental (IM-MS coupled to IRMPD spectroscopy) and theoretical methods (enhanced sampling molecular dynamics, MD). This approach allows probing the molecular structure on complementary levels, ranging from the overall molecular shape *via* IM-MS, through secondary-structure sensitive gas-phase IR spectroscopy, up to atomistic resolution MD simulations.

4.2 Experimental Details

Materials: The peptide $Ac\text{-}Ala_{10}Lys\text{-}OH$ (AK00) and the corresponding depsi-modified (AK01-AK10) peptides were purchased from KareBay Biochemicals (Rochester, New York)

and used without further purification. Samples were dissolved at 50 μM in a 1:1 H_2O:MeCN solution with 1% trifluoroacetic acid (TFA).

IM-MS and Drift-Time Selective IR Spectroscopy: IM-MS and gas phase IR spectroscopy were performed using an in-house built drift-tube instrument coupled to the FHI-FEL[69]. A detailed description of the instrumental procedure and data analysis can be found in chapter 2.6.

Simulation details: Molecular dynamics (MD) simulations were performed using the gromacs-4.5.5 package.[162] The OPLS-AA force field[163-164] was augmented with parameters for the depsi peptide bond taken from a previous work on polylactic acids (**Figure B.1** and **Figure B.2**).[165-166] The AK00-AK10 peptides were first thermalized at 300 K for 1 ns, followed by 100 ns production runs with 1 fs time step. The temperature of 300 K was maintained during the simulations using stochastic velocity rescaling thermostat[167] with a coupling constant of 0.05 ps. To enhance the sampling of the conformational space, replica-exchange MD (REMD) simulations were performed for AK00 (2.0 μs accumulated time of 4 replicas at 300-380 K) and AK03-AK06 (8.0 μs accumulated time of 16 replicas at 300-915 K). For each system, snapshots in 25 ps intervals were extracted from the 300 K trajectory. The recorded geometries were clustered with a single-linkage algorithm and the cutoff of 0.1 nm. Clustering yielded up to two different clusters. A summary is given in **Table 4.1**.

Table 4.1: Summary of total simulation time at T = 300 K of a) MD simulations, b) REMD simulations using 4 replicas, c) REMD simulations using 16 replicas.

Protonated peptides	*Total simulation time, T = 300 K*	*Sum of central structures used for DFT calculations*
$[AK00+H]^+$	500 ns [b)]	1
$[AK01+H]^+$	100 ns [a)]	1
$[AK02+H]^+$	100 ns [a)]	1
$[AK03+H]^+$	500 ns [c)]	1
$[AK04+H]^+$	500 ns [c)]	2
$[AK05+H]^+$	500 ns [c)]	2
$[AK06+H]^+$	500 ns [c)]	1
$[AK07+H]^+$	100 ns [a)]	1
$[AK08+H]^+$	100 ns [a)]	1
$[AK09+H]^+$	100 ns [a)]	1
$[AK10+H]^+$	100 ns [a)]	1

The central structures of most abundant clusters were optimized using the non-local hybrid B3LYP[168] density-functional and the cc-pVDZ basis set using the Gaussian09 package.[169] B3LYP shows an excellent performance in electrostatic-dominated systems like the peptides studied here.[170] The geometry relaxations at the B3LYP level introduced only minor

structural changes. The calculated RMSD values between the initial (from MD) and optimized structures are typically smaller than 1 Å. Only for AK00, AK03, AK04 and AK08 were somewhat larger RMSD values observed, which essentially reflect rearrangements in the lysine side chain. Nevertheless, these changes should be attributed to the different meaning of the arbitrarily chosen of the central structure, which represents a cluster, and the 0 K DFT optimized energetic minimum. The overlays of each pair also show that there is no significant rearrangement (**Figure B.3**). Harmonic vibrational frequencies were calculated for the relaxed structures. Predicted vibrations were then scaled by 0.971 and convoluted with Gaussian functions (bandwidth 10 cm^{-1}).

Atomic partial charges (ESP charges) for the equilibrium structures were derived by fitting the electrostatic potential with the Merz-Singh-Kollman sheme.[171] These ESP charges are used in the calculation of theoretical $^{TM}CCS_{He}$ using the trajectory method[57] as implemented in the MOBCAL program.[56] In the trajectory method, each atom of the molecule is represented by a 12-6-4 potential (Lennard-Jones term plus ion-induced dipole interactions) and the effective potential of an ion is built by summation of the potentials of all atoms. The trajectories of the incoming buffer gas (He) are simulated under the effective potential, then the scattering angle is determined as a function of the impact parameter for a given collision geometry. The scattering angle is related to the $^{TM}CCS_{He}$ and an average $^{TM}CCS_{He}$ is calculated by integrating over all possible collision geometries. A summary of experimental $^{DT}CCS_{He}$ and theoretical $^{TM}CCS_{He}$ is given in **Table B.1**.

4.3 Results and Discussion

The IM-MS ATDs of AK00 and its depsi-peptide analogues (AK01-AK10) exhibit narrow, single peaks, indicating the presence of a single conformer (**Figure 4.2**a and **Figure B.4**a). The drift time of AK00 corresponds to an overall $^{DT}CCS_{He}$ of 237 $Å^2$. This is in good agreement with the previously reported 235 $Å^2$, which has been assigned to an α-helical structure.[149-151]

The NH groups close to the N-terminus are not involved in H-bonds in helical structures and typically point towards the N-terminus. As a result, amide-to-ester substitution at these dangling N-H groups (peptides AK08-AK10) does not delete any H-bonds and induces only minor structural changes (AK08 = 234 $Å^2$, AK09 = 236 $Å^2$, AK10 = 237 $Å^2$, **Figure 4.2**a and **Figure B.4**a). Intuitively, the deletion of an NH donor group close to the

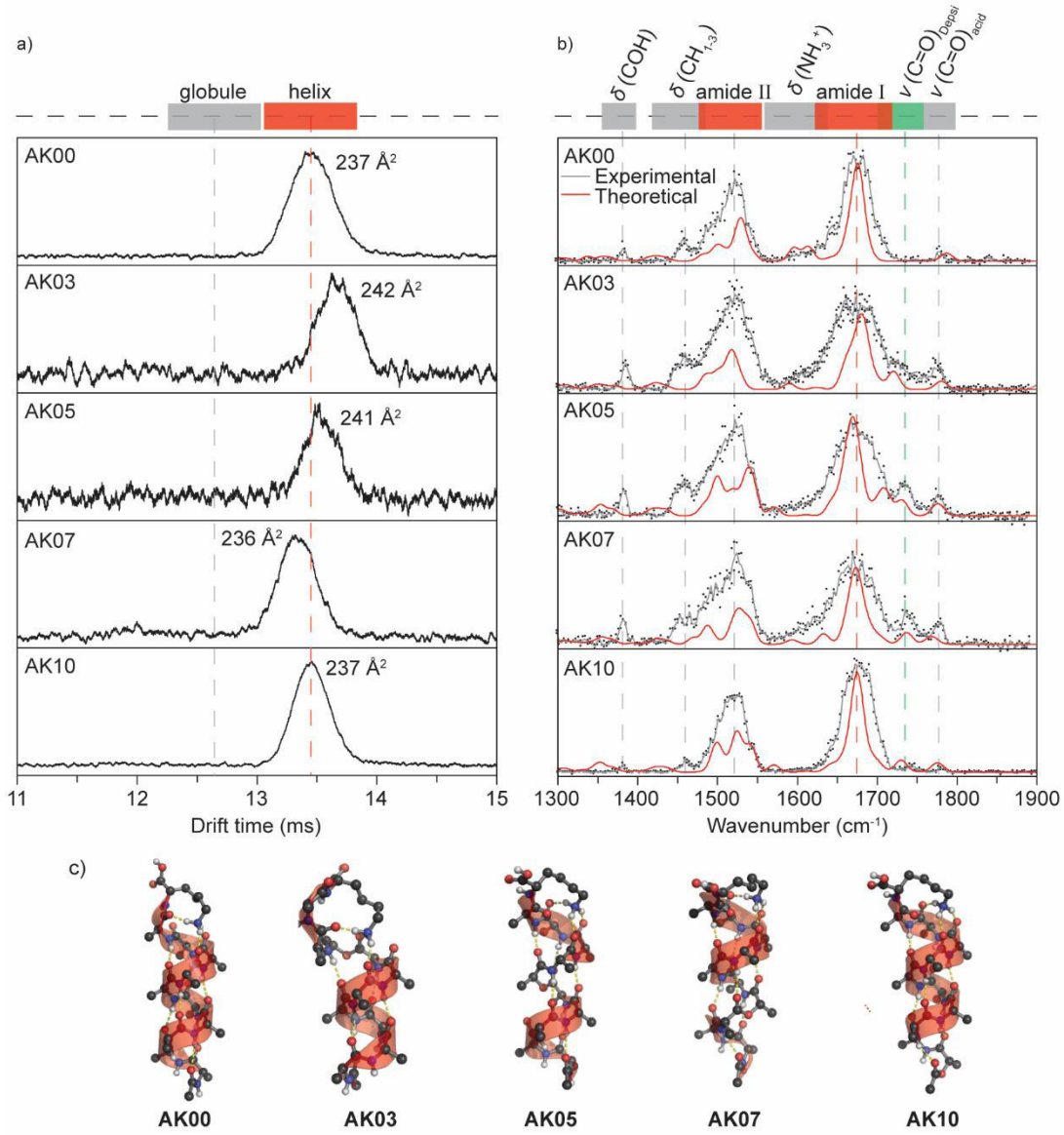


Figure 4.2: (a) ATDs of selected depsi peptides as well as their corresponding $^{DT}CCS_{He}$ are shown in black. Dashed red and gray lines represent drift-time positions of calculated $^{DT}CCS_{He}$ reported for helical and globular conformations of $[Ac\text{-}Ala_{10}Lys + H]^+$, respectively. (b) Gas-phase IR spectra of ion mobility- and m/z-selected conformations. Data from two individual scans per species are depicted as black dots. The solid gray lines represent an average of these two individual scans. Theoretical vibrational spectra of geometry-optimized structures (c) are displayed as solid red lines. Band assignment is schematically shown in the upper panel.

C-terminus or in the middle of the helix may have a larger impact on the H-bond network and therefore on the conformational stability, since these residues are involved in the formation of stabilizing H-bonds.[156] However, IM-MS analysis of peptides AK01-AK07 reveals drift times and $^{DT}CCS_{He}$ that correspond to conformations of similar (AK07 = 236 Å² and

AK0*n* = 237 Å^2 with *n* = 0, 2, 4, and 6) or slightly extended size (AK01 = 240 Å^2, AK03 = 242 Å^2 and AK05 = 241 Å^2). Remarkably, none of these peptides adopts a more compact, globular structure, as observed for protonated $[Ac\text{-}LysAla_{10} + H]^+$ (reported $^{DT}CCS_{He}$ = 219 Å^2).[149, 151, 154] This observation therefore indicates that for the molecules studied here, the helical scaffold remains largely unaltered.

To further investigate the structural details of the depsi-peptide analogues the shape and *m*/*z*-selected ions were probed using IRMPD spectroscopy. The resulting gas-phase IR spectra shown in **Figure 4.2b** and **Figure B.4b**, exhibit two characteristic bands at 1600-1700 cm^{-1} (amide I band) and 1450-1550 cm^{-1} (amide II band). The IR features above 1700 cm^{-1} originate from carbonyl groups of free or only weakly-interacting carboxylic acids (~1780 cm^{-1}) and the decoupled carbonyl stretching mode of the introduced ester group (1720-1750 cm^{-1}). The dominant amide I band of the α-helical AK00 appears at 1675 cm^{-1} and the position of the respective band in the amide-to-ester substituted peptides shows little variation. The amide I bands of AK01-AK07 are somewhat broader than the respective band in the unsubstituted AK00 and AK08-AK10 peptides. This may be due to the deletion of hydrogen bonds positioned in the middle of the H-bond chain, which decouples in-phase stretching modes of carbonyl groups and effectively broadens the amide I band.[155] Theoretical spectra of the α-helical peptides (red line in **Figure 4.2b** and **Figure B.4b**) confirm that the substitution in the middle of the helix leads to such a broadening of the amide I band.

In order to better understand the structure of the peptides on the atomistic level, a conformational search with enhanced sampling MD or plain MD techniques (see simulations details) were performed. While the starting geometry determines the fraction of the phase space that is sampled, and insufficient sampling might lead to trapping of a system in a meta-stable-state[172], an idealized α-helix is presumably still a reasonable choice starting point for MD simulations. Previous theoretical studies used extensive sampling techniques to show that lysine C-terminated polyalanine peptides form α-helices in the gas phase,[153] while N-terminated ones form other motives.[152] Moreover, Jarrold et al. have shown that the protonated AK00 peptide exists solely as an α-helix in the gas phase.[149] Our results show that the amide-to-ester substitution in AK08-AK10 does not alter the overall α-helical structure (**Figure 4.2c** and **Figure B.4c**) whereas the deletion of a hydrogen bond in AK06 and AK07 is compensated by minor rearrangements between 3_{10} and α-helices in the vicinity of the lactic acid. Such transitions have been observed before in amide-to-ester substituted helices in crystallized peptides.[173-174] Substitution in the middle of the helix (AK05) yields two abundant

conformers. While one maintains the α-helical character ($^{TM}CCS_{He}$ = 238 Å^2), the other conformer adopts a rather compact structure ($^{TM}CCS_{He}$ = 230 Å^2). In the latter, amide-to-ester substitution terminates the helix (**Figure B.3**), which renders part of the peptide unstructured. The $^{TM}CCS_{He}$ of this more compact structure, however, does not agree with the $^{DT}CCS_{He}$ of 241 Å^2. Finally, the replacement of an amide bond by an ester bond close to the C-terminus of the helix (AK01-AK04) results in the formation of a π-turn, combined with a partial unfolding of the helix induced by binding of the charged lysine to the dangling carbonyls. More importantly, however, the helical structure largely prevails here as well. Such unfolding at the C-terminus justifies structures with larger $^{DT}CCS_{He}$ in IM-MS and the broadening of the amide I absorption band within the IRMPD spectra shown in **Figure 4.2b** and **Figure B.4b**.

4.4 Conclusions

The protonated peptide [Ac-Ala$_{10}$Lys + H]$^+$ (AK00) is known to adopt an α-helical structure in the gas phase. This structure is stabilized by three factors: (i) the attraction between the helix macro-dipole and the charge, (ii) the capping to dangling carbonyls by the protonated lysine side chain, and (iii) the cooperative H-bonding along the helix (**Figure 4.1**).[149-151] In this work, we complement previous studies focused on the first two factors by evaluating the impact of individual backbone H-bonds on the stability of the α-helix scaffold.

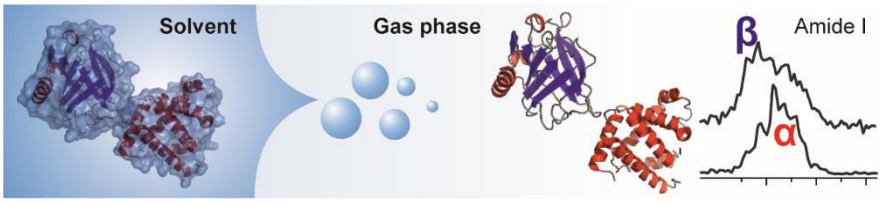
IM-MS experiments reveal that regardless of the position within the sequence, amide-to-ester substitution of one peptide bond only results in minor structural changes and conformations that, if altered at all, are slightly more extended than the unmodified peptide. This is confirmed by IRMPD spectroscopy experiments, which indicate that the predominantly helical conformation of the unmodified peptide AK00 is still present in the depsi analogues. Molecular modelling shows that the observed subtle changes in structure are largely dependent on the sequence position at which the H-bond is deleted. Close to the N-terminus (AK08, AK09 and AK10), the removal of a single H-bond leaves the α-helix essentially unaffected. Further towards the helix centre (AK06 and AK07), H-bond deletion introduces small structural rearrangements to bifurcated 3_{10}/α-helices. A substitution close to the C-terminus (AK01-AK05) leads to unfolding of the helix and formation of a π-turn induced by the protonated lysine. Nevertheless, the α-helix scaffold remains the defining structural motif.

Our findings are in good agreement to previous work of Schultz and co-workers who investigated the effect of amide-to-ester substitution in protein secondary structures in the

condensed phase. They found that a single ester group destabilizes α-helices by a mere 0.7-1.7 kcal mol^{-1}.[175] This surprisingly small energetic penalty might be an effect of increased solvent accessibility to the backbone and/or an electrostatic effect associated with the backbone.[176] In the gas phase, the solvent is absent and electrostatic effects are substantially stronger. As a result, the energetic penalty of the amide-to-ester substitution in the depsi-peptides studied in this work might even be smaller than 0.7-1.7 kcal mol^{-1}, and therefore too small to disturb the α-helix significantly.

Previous studies have shown that the attractive interaction between the helix macro-dipole and the charge shapes the energy landscape of polyalanine peptides, whereas capping to dangling carbonyl groups by the protonated lysine side-chain has only a minor influence on the stability of the α-helix.[90, 152, 154] The removal of a single backbone H-bond has also only a small impact on the stability of the helix in the gas phase. Therefore, the interplay of the charge and the helix macro-dipole is by far the most crucial parameter for the formation of the α-helical conformation in AK00. Capping of the protonated lysine side-chain to dangling carbonyls can slightly disturb the helicity, while individual backbone H-bonds have only minor influence on the overall stability.

5 Retention of Native Protein Structures in the Gas Phase*



5.1 Introduction

Owing to its accuracy, sensitivity and speed, MS coupled to fragmentation techniques is the method of choice for determining the primary structure of peptides and proteins.[177] To obtain information on higher-order structures, condensed-phase spectroscopic or scattering techniques are usually employed. Often they have limited sensitivity and therefore require high amounts of sample but they can provide a wealth of information under close to physiological conditions. An important question is whether native solution structures can be maintained when a molecule is transported into the solvent-free environment, and if so, over what protein size range. If this transfer can be done with the native structure remaining intact, highly sensitive MS techniques can be employed to interrogate these structures. For very large species, such as protein complexes[178-179] or even an entire virus[75], there is unambiguous evidence that native structures can be retained in the absence of solvent. For smaller species, however, the situation is not so clear, and this is the region which is examined here.

To investigate higher-order structures of peptides and proteins in isolation, MS can be coupled to HDX measurements,[180] non-thermal dissociation techniques that rely on the absorption of UV photons or the capture of electrons,[181] IM-MS,[182-184] or optical/IR-spectroscopy.[185-187] Of these, IM-MS and IR-spectroscopy provide the most direct information on higher-order structure. These methods have shown that when the charge increases, the molecules undergo Coulomb-induced unfolding to adopt extended helical structures[86, 188] and ultimately string-like extended structures at very high charge states.[87, 189] At low charge states,

* This chapter is based on the work published in J. Seo, W. Hoffmann, S. Warnke, M. T. Bowers, K. Pagel and G. von Helden, *Angew. Chem. Int. Ed.*, **2016**, *55*, 14173-14176, DOI: 10.1002/anie.201606029. This work was highlighted with a back cover. Figures and content adapted with permission. Copyright 2016. Published by Wiley-VCH Verlag GmbH & Co. KGaA.

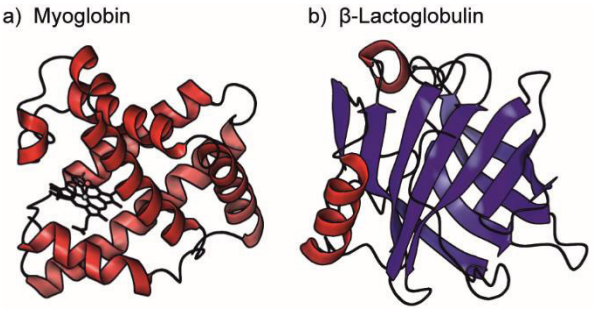


Figure 5.1: Condensed-phase structures of a) myoglobin, a 153 amino acid protein with a native secondary structure that is mostly α-helical and which contains a non-covalently attached heme group (PDB ID: 1MBN) and b) β-lactoglobulin, a 162 residue protein that has abundant β-sheet secondary structure (PDB ID: 3BLG).

however, the proteins usually remain compact, with sizes that are expected for condensed-phase native structures. [86, 188] However, IM-MS is not directly sensitive to protein secondary structure, and observing a compact structure does not guarantee that the secondary structure accurately reflects that of the native form. IR spectroscopy, on the other hand, is sensitive to secondary structure.

IRMPD spectroscopy[67] has been applied to a large number of small peptides or peptide complexes[185-186] in the gas phase. When combined with quantum chemical calculations, detailed structural information can be obtained. For very small systems containing only a few amino acids, gas-phase structures that are distinctively different from those in the condensed phase are usually observed. However, when large energy barriers – such as in case of the *cis-trans* isomerization of proline – are involved, condensed-phase structural motifs can be retained.[82-83] In any case, these gas-phase structures are still very useful because they can be used to calibrate theory and to test our general understanding of intrinsic intramolecular interactions in biological molecules.

There are only a few IR studies on isolated proteins in the absence of solvent.[89, 190-192] In mid-IR spectra of larger species, individual oscillators cannot be spectrally resolved and, as for proteins in the condensed phase, amide-I and -II envelopes are observed.[87, 89, 190-191] The band positions in these studies suggest predominantly helical structures. However, in these early studies, the conditions employed and the relatively high charge states[87, 89, 190-191] investigated were not favorable for retaining native solution structures.

In this work, IM-MS is used to select proteins based on their global shape, followed by IR spectroscopy to probe their secondary structure. Experiments were performed on the two proteins myoglobin and β-lactoglobulin. Myoglobin is a 153 amino acid protein with a native secondary structure that is mostly (ca. 85%) α-helical, while β-lactoglobulin is a 162 residue protein that has a significant (ca. 60%) proportion of β-sheet secondary structure. The crystal structures of both proteins are shown in **Figure 5.1**.

5.2 Experimental Details

Materials: All chemicals were purchased from Sigma-Aldrich (Taufkirchen, Germany). Proteins were dissolved at a concentration of 50 µmol/L in either a) pure water with ammonium acetate (10 mmol/L) as a buffer, b) 1/1 water/methanol or c) 1/1 water/methanol with the addition of 1% formic acid. To all solutions containing β-lactoglobulin, 50 µmol/L palmitic acid was added as well.

IM-MS and Drift-Time Selective IR Spectroscopy: The experiments and data analysis were performed using the apparatus described in chapter 2.6.

Myoglobin has a non-covalently bound heme group attached which is easily lost for high charge states in the gas phase or when being activated. β-Lactoglobulin is able to bind hydrophobic molecules such as fatty acids.[193-194] To record some of the IR spectra shown, the dissociation of the ligand (heme group or palmitic acid) from the protein is used as a marker for the absorption of (multiple) photons. Compared to standard IRMPD experiments,[67] a smaller number of absorbed photons are needed to induce dissociation, thereby reducing the influence of anharmonicities and non-linearities on the resulting spectra.

In the here-presented experiments, palmitic acid is observed to bind to β-lactoglobulin when being sprayed from aqueous solution. The corresponding mass peaks are marked in the mass spectra of **Figure 5.2**b as 9+´, 8+´ and 7+´. For the charge states 8+ and 9+, $^{DT}CCS_{He}$ have been measured for both the bare ions and the corresponding ions complexed with palmitic acid. Their respective differences in $^{DT}CCS_{He}$ are small and within the dot sizes of **Figure 5.2**d. When being sprayed from water/methanol, no palmitic acid adduct is observed. For myoglobin, the heme-containing holo form is observed when being sprayed from aqueous solutions as well as water/methanol, however, not when 1% formic acid is added, in which case the apo form of myoglobin is observed. Mass spectra of β-lactoglobulin and myoglobin ions when being sprayed from water/methanol with 1% formic acid added are shown in **Figure C.2**. The mass spectra shown here are similar to those reported by others.[193, 195]

The IR spectra of myoglobin 8+ and 10+ (**Figure 5.3**) are recorded by monitoring the loss of the heme group as a function of excitation wavelength. For the 10+ charge state, several conformers are present (**Figure C.3**), all having $^{DT}CCS_{He}$ larger than expected for a native protein. The spectrum shown in **Figure 5.3**b results from the low $^{DT}CCS_{He}$ conformer II, however, the spectra for all other myoglobin conformers in this charge states are essentially identical (**Figure C.3**). To record the IR spectrum of the 18+ charge state, dissociation of the molecule is monitored. For β-lactoglobulin, the IR spectrum of the 8+ charge is recorded by

measuring the loss of the palmitic acid ligand. For higher charge states, fragmentation of the molecule is monitored.

Condensed-Phase Circular Dichroism Spectroscopy: CD spectra were measured at 20 °C using a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD, USA) equipped with a Jasco PTC-423S Peltier temperature control system and a HAAKE WKL water recirculator (Thermo Electron GmbH, Karlsruhe, Germany). Proteins are dissolved at a concentration of 50 μmol/L in either a) 10 mmol/L ammonium acetate buffer (pH ~ 7) or b) 1/1 water/methanol, identical conditions as used for the gas-phase experiments. A quartz glass cuvette with 1-mm path length was used to record CD spectra at 190-240 nm with 2 nm spectral resolution. Solvent spectrum is recorded under identical conditions and used for the background-correction. CD spectra for β-lactoglobulin were measured with and without the addition of 50 μmol/L of palmitic acid.

5.3 Results and Discussion

Figure 5.2a, b shows *n*ESI mass spectra of myoglobin and β-lactoglobulin sprayed from solvents of different compositions. The lower spectra in green are obtained with spraying from aqueous solutions with ammonium acetate buffer at a concentration of 10 mmol/L. Under these conditions, narrow charge state distributions with maxima at 8+ are observed. For β-lactoglobulin, additional mass peaks resulting from adducts of palmitic acid are found and marked in **Figure 5.2**b as 9+', 8+' and 7+'. Changing the solvent to water/methanol (1/1) results in the mass spectra shown in blue. The charge-state distributions are broader and shifted to higher charge states. For myoglobin, all labeled peaks in the two spectra correspond to the heme-containing holo form. The mass spectra shown here are similar to those reported by others.[193, 195]

The resulting $^{DT}CCS_{He}$ as a function of charge z measured by IM-MS are shown in **Figure 5.2**c, d. The uncertainties in the $^{DT}CCS_{He}$ are in all cases smaller than the size of the corresponding symbol. Green circles correspond to ions sprayed from aqueous solution while blue circles correspond to ions sprayed from water/methanol solutions. In the case of myoglobin, multiple conformers are observed for most charge states and, correspondingly, more than one $^{DT}CCS_{He}$ value per charge state is shown. Theoretical $^{EHSS}CCS_{He}$ expected for condensed-phase structures (**Figure 5.1**) have been calculated using the EHSS algorithm[57] and the values are shown as dotted lines in **Figure 5.2**c, d.

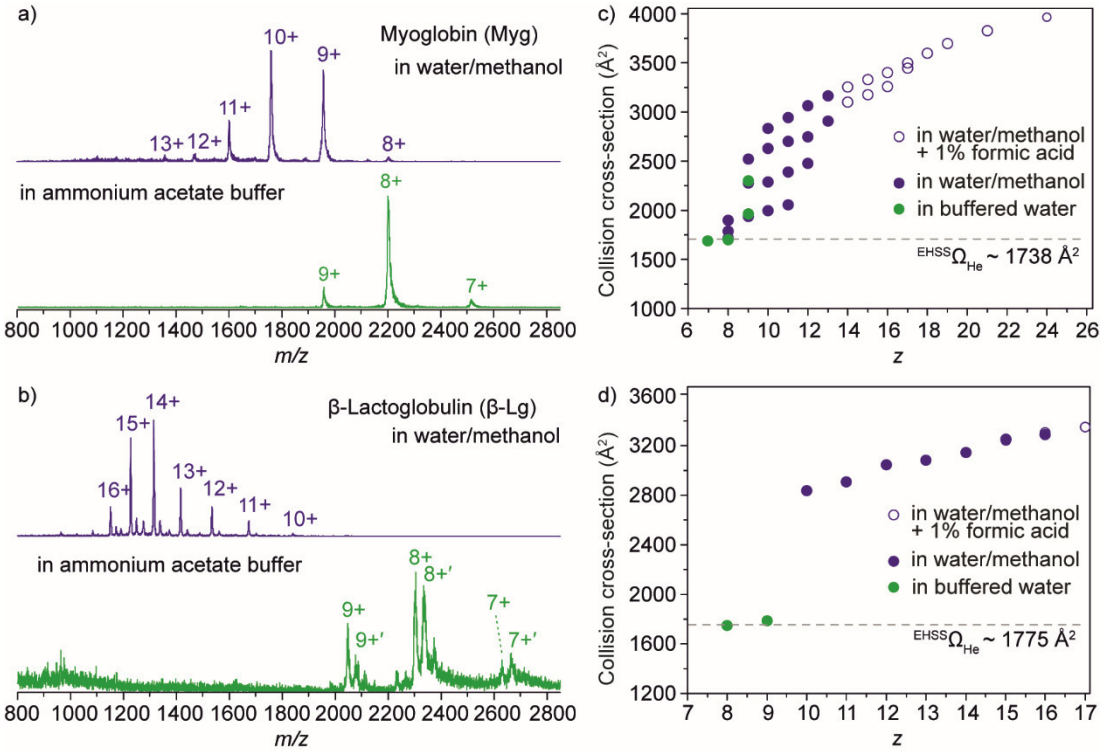


Figure 5.2: Mass spectra and $^{DT}CCS_{He}$ for myoglobin and β-lactoglobulin. a, b) Mass spectra obtained when spraying from buffered (ammonium acetate, pH ~ 7) aqueous or water/methanol solutions (green and blue, respectively). Prominent peaks are labeled, and in the case of myoglobin, they correspond to the holo form. For low charge states of β-lactoglobulin, adducts with palmitic acid can be observed (labeled 7+´, 8+´ and 9+´). c, d) $^{DT}CCS_{He}$ as a function of charge. Solid green circles represent ions sprayed from buffered aqueous solution, solid blue circles represent ions from water/methanol solution, and open blue circles represent ions from water/methanol solution with 1% formic acid.

As expected, the $^{DT}CCS_{He}$ increase with increasing charge. At low charge states, compact ions are observed that have $^{DT}CCS_{He}$ consistent with those expected for native structures. When the charge increases, the $^{DT}CCS_{He}$ increase, thus indicating unfolding of the protein. For myoglobin several conformers with different degrees of unfolding coexist for all but the lowest and highest charge states.[196-197] β-Lactoglobulin shows a different behavior. For each charge state, only a single conformer is observed. Furthermore, a large jump in $^{DT}CCS_{He}$ occurs when the solvent is changed to water/methanol, and the minimum observable charge state increases to 10+.

In order to probe the secondary structure, IR spectroscopy experiments were performed using the FHI-FEL.[69] **Figure 5.3** shows IR spectra for several charge states for both proteins. In all of the spectra, amide-I and -II bands can be observed. In **Figure 5.3**c (green lines) IR-spectra for ions with a low $^{DT}CCS_{He}$ and low charge state are shown.

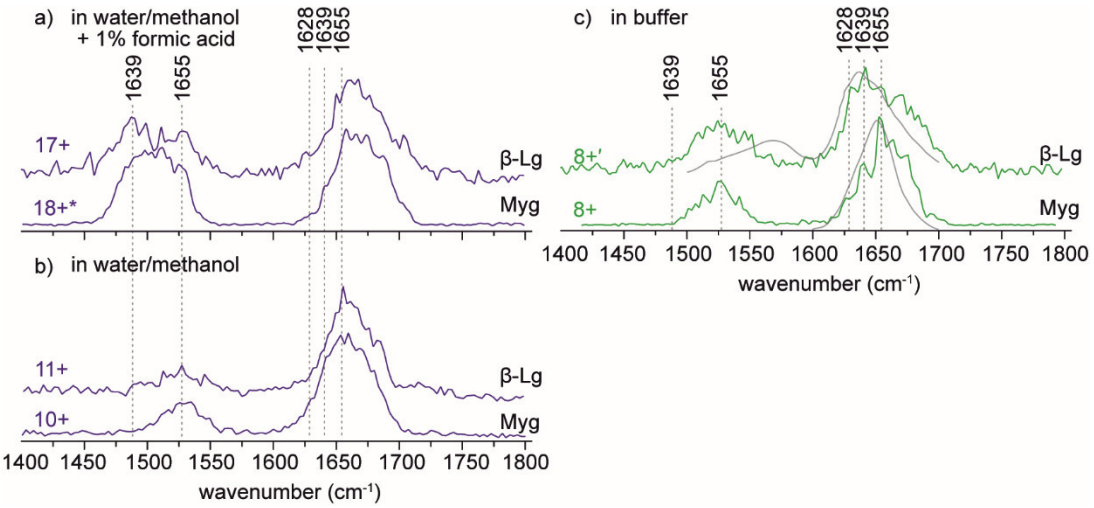


Figure 5.3: IR spectra for gas-phase and condensed-phase myoglobin and β-lactoglobulin. a, b) Spectra for ions sprayed from water/methanol solutions. The positions and shapes of the amide-I bands around 1655 cm^{-1} indicate helical secondary structures. c) Spectra for samples sprayed from aqueous solutions (green lines) and solution phase FT-IR spectra (gray lines; data reproduced from literature). The amide-I band for myoglobin 8+ indicates helical secondary structure, while that of β-lactoglobulin shows a clear β-sheet signature.

For both 8+ ions, the amide-II bands are found around 1525 cm^{-1}, a value that is typical for proteins in the condensed phase.[65] The amide-I band for myoglobin 8+ is roughly symmetrical and has a maximum at 1655 cm^{-1}, which is typical for a helical secondary structure.[65] The IR spectrum of β-lactoglobulin in the 8+ charge state, on the other hand, exhibits a rather different signature. There, the amide-I band is asymmetric, with a maximum around 1639 cm^{-1}, typical for β-sheet-rich proteins.[65]

Figure 5.3a, b shows IR-spectra for higher charge states. For myoglobin in the 10+ charge state, the spectrum shown in **Figure 5.3**b results from a low $^{DT}CCS_{He}$ conformer (see **Figure 5.2** for the $^{DT}CCS_{He}$ distribution), however the spectra for all myoglobin conformers in this charge state are essentially identical (**Figure C.3**). The positions of the amide-I and -II bands are almost unchanged for the 8+ and 10+ charge states, and only the width of the amide-I band increases slightly. For β-lactoglobulin, on the other hand, the shape and position of the amide-I band changes dramatically between the 8+ and 11+ charge states. This transition is accompanied by a very large change in $^{DT}CCS_{He}$ between these two charge states (**Figure 5.2**). For the unfolded and higher charged 11+ conformer, the amide-I band is similar in shape and position to that of 10+ myoglobin, which indicates a helical structure for β-lactoglobulin 11+ ions as well.

When the charge increases further to 17+ for β-lactoglobulin and 18+ for myoglobin, the shapes of the amide-I bands do not change and only show a slight shift to higher wavenumbers. The amide-II bands, on the other hand, gain in relative intensity and shift significantly to lower wavenumbers. This observation can be attributed to the Coulomb-driven unraveling of helices to form extended string-like structures.[87, 189]

From a visual inspection of the spectra in **Figure 5.3**, it appears that gas-phase β-lactoglobulin in the 8+ charge state has a predominantly β-sheet secondary structure, while myoglobin has an essentially helical secondary structure, which is consistent with the native solution-phase structures of both proteins. As charge increases, however, both proteins exhibit helical structures, and finally at the highest charge states, even these helices unravel. It is interesting to compare these spectra with results from the condensed phase. CD spectra were measured for both proteins using the same solvents as used for the spectroscopy experiments on the isolated species. All of the CD spectra for myoglobin and the CD spectrum of β-lactoglobulin in water/methanol have signatures that are typical for helical structures[198] (**Figure C.1**). On the other hand, the CD spectrum of β-lactoglobulin in buffered aqueous solution is quite different and clearly indicates a β-sheet structure.

In **Figure 5.3c**, the gas phase spectra (green lines) are compared to the condensed phase FT-IR spectra (grey lines) for β-lactoglobulin[199] and myoglobin[200] reproduced from previous reports. The FT-IR amide-I band of β-lactoglobulin shows the typical shape expected for a β-sheet-rich protein, while that of myoglobin is rather symmetrical and is found in a position typical for helical species. Clearly, in both cases, the match between the corresponding gas-phase and condensed-phase spectra in the amide-I region is very good, thus giving further evidence that, for low charge states, the condensed-phase structure is at least in part preserved after transfer to the gas phase.

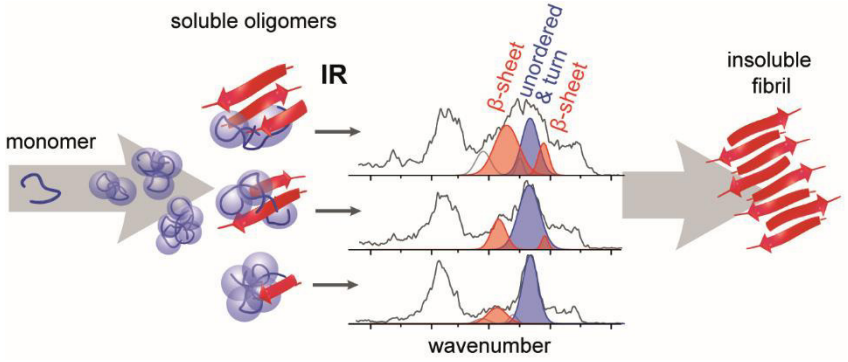
IR spectroscopy directly probes local secondary structure and can differentiate between helices and β-sheets. However, it does not provide direct information about the tertiary structure in which those secondary structure elements are embedded. IR spectroscopy is thus a good complement to IM-MS, which on its own is blind to structural details and only probes the overall shape and size. The combination of IR spectroscopy and IM-MS shown here gives a clear picture on the structural evolution of isolated proteins as their charge increases. Low charge states can retain the native structure. When Coulomb repulsion increases and charged side chains coordinate to the protein backbone,[84] the native fold begins to lose stability, which is indicated by an increase in $^{DT}CCS_{He}$. When the original native

structure is predominantly helical, IR spectroscopy shows that the unfolded species remain helical. This is to be expected, since extended helical structures can retain most of their hydrogen bonds while maximizing the distances between equal charges. However, the situation is different for native β-sheet proteins. In these, the individual β-sheet strands are embedded in a hydrogen bond network, and when an increase in charge causes unfolding, preservation of the β-sheet strands would result in many disrupted hydrogen bonds. Hence, for an initial β-sheet dominated structure, the unfolded protein will adopt a helix-rich structure as well. Finally, at very high charges, Coulomb repulsion will cause helix unzipping in both cases.[87, 189]

5.4 Conclusions

In summary, the presented data clearly show that for β-lactoglobulin and myoglobin, the condensed-phase secondary as well as tertiary structure can be conserved when solvent is completely removed. Methods based on gas-phase MS provide extremely high sensitivity. When MS is combined with IM-MS, additionally selectivity can be gained through the selection of individual charge states, conformations, or aggregation states, which then can be investigated using spectroscopy. Such sensitivity and selectivity cannot be achieved using condensed-phase methods. However, care must be taken to be sure the journey from the condensed phase to the solvent-free gas phase is very gentle to avoid possible refolding and structure changes.

6 β-Sheet Formation in Peptide Amyloid Assemblies*



6.1 Introduction

The assembly of peptides and proteins into insoluble, β-sheet-rich fibrils is a common hallmark of all amyloid diseases.[24, 201-202] However, small soluble oligomers rather than mature fibrils have been suggested as the symptom-inducing toxic species. Oligomers of amyloidogenic peptides, such as the β-amyloid peptide,[99, 203-204] α-synuclein,[205] islet amyloid polypeptide,[206-207] insulin,[20, 208] and several others are suspected to be responsible for Alzheimer's, Pakinson's, type II diabetes and other amyloid diseases. Exploring the structure of these oligomers is not only crucial to develop a better understanding of the underlying molecular details of amyloid assembly, but is also key to the development of novel diagnostic and therapeutic methods.

Extensive structural studies using X-ray diffraction, nuclear magnetic resonance spectroscopy (NMR) and electron microscopy (EM) have revealed that highly ordered β-sheets maintained by cross-β spines are dominant in mature fibrils.[30-31, 208-210] Moreover, CD and IR spectroscopy studies revealed that large secondary structural transitions from partially helical or unordered structures into repeating β-sheets occur during the assembly process.[211-213] Thus, the determination of the secondary structure of pre-fibrillar oligomers and the identification of the specific states responsible for the secondary structural transitions to a β-sheet should be the first steps in investigating the assembly of amyloid peptides and

* This chapter is based on the work published in J. Seo, W. Hoffmann, S. Warnke, X. Huang, S. Gewinner, W. Schöllkopf, M. T. Bowers, G. von Helden and K. Pagel, *Nat. Chem.*, **2017**, *9*, 39-44, DOI: 10.1038/nchem.2615. J. Seo and W. Hoffmann contributed equally to this work. Figures and content adapted with permission from Macmillan Publishers Ltd., copyright 2017.

proteins. However, the polydisperse, polymorph, and transient nature of the involved oligomers makes structural investigations inherently challenging, because oligomers usually exist in a distribution of distinct stoichiometries and conformational states. Traditional condensed-phase methods are of limited use for structural studies of these systems because they only yield ensemble-averaged results. To address this experimental difficulty, several methods based on MS and IM-MS have been applied to study amyloid assembly.[9, 45, 52, 63, 96, 108, 214-217] With the unique capability to isolate and detect molecular ions of a specific m/z ratio and measure CCSs, recent IM-MS studies successfully isolated size- and shape-specific oligomer states of amyloidogenic peptides and extracted information on structure as well as oligomer distribution.

The peptide VEALYL, which corresponds to segment 12-17 of the insulin B chain, was recently shown to form amyloid fibrils that later evolve into microcrystals in solution.[31, 63, 210] X-ray studies have revealed that the peptides assemble into a so-called steric zipper – a repeating pattern of antiparallel β-sheets with densely packed, interdigitated side-chains.[30-31, 210] IM-MS methods found that a structural transition from a compact to a more extended structure appears early in the assembly process.[63] However, IM-MS is not sensitive to the secondary structure itself; it only provides information on the overall shape of a molecule and relies on comparison with modeling studies to obtain structural details. To address this limitation, IR spectroscopy with IM-MS were combined to probe the structure of shape- and m/z-selected oligomers of VEALYL and its sequence-scrambled peptides VELYAL and YVEALL.

6.2 Experimental Details

Solid Phase Peptide Synthesis: All peptides were synthesized according to standard Fmoc (9-fluorenyl-methoxycarbonyl) solid phase peptide synthesis (SPPS)[218] by using preloaded Fmoc–L-Leu Sasrin® resin ester (0.64 mmol g^{-1}) and the following Fmoc-protected amino acid derivatives: Fmoc–L-Ala–OH, Fmoc–L-Glu(O-*tert*-Bu)–OH, Fmoc–L-Leu–OH, Fmoc–L-Val–OH, Fmoc–L-Tyr(*tert*-Bu)–OH. All solvents and reagents were used without further purification.

SPPS was carried out on a fully automatic peptide synthesizer: Activo-P11 (Activotec, Cambridge, United Kingdom) or Syro XP-1 (Multi-Syn Tech GmbH, Witten, Germany). Peptides were cleaved and preparative, reverse phase-high performance liquid chromatography (RP-HPLC) was performed on a Knauer high-pressure-gradient system

equipped with a LUNA™ C8(2) column (10 μm, 250 × 21.2 mm, Phenomenex®, Torrance, CA, USA). The flow rate was 20 mL min^{-1}, and UV-detection occurred at $\lambda_{abs.}$ = 220 nm. Analytical RP-HPLC was performed on a computer-controlled high-pressure-gradient LaChrom-ELITE-HPLC-System (VWR, International Hitachi) using a LUNA™ C8(2) column (5 μm, 250 × 4.6 mm, Phenomenex®, Torrance, CA, USA) or a Kinetex C18 column (5 μm, 250 × 4.6 mm, Phenomenex®, Torrance, CA, USA). The flow rate was 1 mL min^{-1}, and UV-detection occurred at $\lambda_{abs.}$ = 220 nm. All runs used water and acetonitrile (ACN) as eluents, both containing 0.1% trifluoroacetic acid (TFA). High-resolution *n*ESI mass spectra were measured at a Synapt G2-S HDMS mass spectrometer (Waters Corporation, Massachusetts, USA).

SPPS using Activo-P11: The preloaded resin was swollen twice in 2 mL of dimethylformamide (DMF) for 15 minutes. Fmoc-deprotection was carried out by the addition of piperidine and 1,8-diazabicyclo[5.4.0]-undec-7-en (DBU). For amino acid coupling, 10 equivalents (eq.) of the Fmoc-protected amino acid derivative were pre-activated with 10 eq. *N*,*N*-diisopropylcarbodiimide (DIC), 10 eq. 1-hydroxy-7-azabenzoriazole (HOAt) in DMF. The solution also contained 5 eq. sodium perchlorate to prevent aggregation on the resin during synthesis. After each coupling or Fmoc-deprotection step, the resin was washed with DMF multiple times.

Table 6.1: SPPS protocol of the automatic peptide synthesizer Activo-P11

Process	Reagents	Reaction time
Swelling	2.0 mL DMF	2 x 15 min
+ Fmoc-deprotection	2.0 mL 2% piperidine, 2% DBU in DMF	2 x 5 min 2 x 10 min
+ Washing	2.0 mL DMF	5 x 1 min
Coupling	10 eq. Fmoc-X-OH, 10 eq. HOAt, 10 eq. DIC in 2.0 mL DMF	2 x 60 min
Washing	2.0 mL DMF	4 x 1 min
Fmoc-deprotection	2.0 mL 2% piperidine, 2% DBU in DMF	2 x 5 min, 2 x 10 min
Washing	2.0 mL DMF	5 x 1 min

SPPS using Syro XP-1: The preloaded resin was swollen twice in 4 mL DMF for 15 minutes. Fmoc-deprotection was carried out by the addition of piperidine and DBU. For amino acid coupling 5 eq. of the Fmoc-protected amino acid derivative were pre-activated with 5 eq. 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethylurnium tetrafluoroborate (TBTU), 5 eq. 1-hydroxy-benzotriazole (HOBt) in DMF and 10 eq. *N*,*N*-diisopropylethylamine (DIPEA) in N-methyl-2-pyrrolidinone (NMP). The solution also contained 5 eq. sodium perchlorate to

prevent aggregation on the resin during synthesis. After each coupling- or Fmoc-deprotection step, the resin was washed with DMF multiple times.

Table 6.2: SPPS protocol of the automatic peptide synthesizer Syro XP-1

Process	Reagents	Reaction time
Swelling	4.0 mL DMF	2 x 15 min
+ Fmoc-deprotection	2% piperidine, 2% DBU in DMF	4x 5 min
+ Washing	2.5 mL DMF	6 x 1 min
Coupling	5 eq. Fmoc-X-OH and HOBt in DMF, 5 eq. TBTU in DMF, 10 eq. DIPEA in NMP	60 min
Washing	2.5 mL DMF	1 min
Coupling	5 eq. Fmoc-X-OH and HOBt in DMF, 5 eq. TBTU in DMF, 10 eq. DIPEA in NMP	60 min
Washing	2.5 mL DMF	3 x 1 min
Fmoc-deprotection	2% piperidine, 2% DBU in DMF	4x 5 min
Washing	2.5 mL DMF	6 x 1 min

Peptide recovery and purification: Cleavage from the resin and side-chain deprotection was carried out by treatment with a solution of 2.67 mL TFA (89%), 300 µL triisopropylsilane (TIS, 10%) and 30 µL water (1%) for 2-4 hours at room temperature. The peptide was successively precipitated using ice-cold and dried diethylether, isolated by centrifugation and dissolved in water/ACN (19/1-9/1, v/v, containing 0.1% TFA) prior to analysis *via* RP-HPLC and *n*ESI-MS. The peptide of interest was purified by preparative RP-HPLC using water (solvent A) and ACN (solvent B) as eluents, both containing 0.1% TFA. A 5-70% solvent B gradient was applied over 30 minutes. HPLC chromatograms and *n*ESI mass spectra of purified peptides are shown in **Figure D.1**. Fractions judged to be pure by analytical RP-HPLC and *n*ESI-MS were combined and lyophilized.

Transmission Electron Microscopy Characterization: TEM images were recorded by Dr. Xing Huang (Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin). Briefly, Samples for TEM imaging were dissolved in water/methanol (1/1, v/v) to yield an acidic (pH ~ 3) solution of 2 mM peptide. The peptide concentration was estimated from the molar weight as TFA-salt and the pH-value was determined by using a WTW pH526 pH meter equipped with an InLab® Micro electrode (Mettler Toledo). VELYAL and YVEALL were incubated under stirring at 37 °C for 24 hours and were further incubated at room temperature for 29 days before TEM analysis. Precipitation for the YVEALL sequence was

observed during incubation. VEALYL was dissolved and used for TEM analysis after only several hours. For TEM images, aliquots of same peptide sample solutions were drop cast onto a carbon coated Cu grid. After being dried in ambient air, they were put into an FEI aberration-corrected Titan 80-300 microscope operated at 300 kV for morphology characterization.

Solid Fourier Transform IR Spectroscopy. Samples for solid FT-IR spectroscopy were dissolved in water/methanol (1/1, v/v) to yield 2 mM solutions, which were incubated under stirring at 37 °C for 24 hours. Subsequently samples were centrifuged for 10 minutes, the precipitations that were formed were freeze-dried. The background-corrected FT-IR spectra of the precipitations (650-4000 cm^{-1}) were measured using a JASCO FT/IR-4100 spectrometer (JASCO Inc., Easton, MD, USA) in transmission mode. Solid samples for the non-fibril forming sequence VELYAL were prepared by quick-freezing the incubated sample solution and further freeze-drying. Solid FT-IR spectra of freshly dissolved sample solutions were prepared as described for the incubated VELYAL sample solution.

Time-dependent UV and FT-IR Spectroscopy in Solution. The UV absorbance at 280 nm (characteristic absorption of the tyrosine side chain) was used to determine the relative amount of precipitation as a function of incubation time. The solution FT-IR was used to monitor the formation of β-sheet-rich intermediates in solution. Peptides were dissolved in D_2O/CH_3OD (1/1, v/v) to a final concentration of 2 mM and incubated at room temperature. Instead of H_2O/CH_3OH, deuterated solvents were used because IR features of H_2O overlap with amide I modes. For measuring solution FT-IR spectra as a function of incubation time, the 10-30 μL of peptide solution was placed between two IR-transparent circular CaF_2 windows separated by a spacer of 0.015 mm thickness, and mounted on the sample holder. The background-corrected IR spectrum was obtained every ~ 30 min using the Bruker IFS 66v spectrometer in the wavenumber range 400-4000 cm^{-1} with a 4 cm^{-1} step. In the amide I region (1600-1700 cm^{-1}), the fraction of IR features at 1600-1630 cm^{-1} (β-sheet content) was determined as a function of incubation time. It should be noted that the amide I feature in the deuterated solvent (as in the case here) shifts slightly to red compared to the values given in reference[65-66]. For measuring the degree of precipitation as a function of incubation time, the supernatant of centrifuged peptide solution was transferred into a disposable cuvette of 1 cm path length. To completely separate precipitate and supernatant, more than 10 min of centrifugation with 12,000 G was applied to the sample solution. The UV absorption spectrum of the supernatant was measured using a SPECORD® 40

spectrophotometer (Analytik Jena AG) in the wavelength range 250-315 nm at a scan rate of 60 nm min^{-1}. The measured supernatants were backfilled to the respective sample solutions for further fibril formation.

IM-MS and Drift-Time Selective IR Spectroscopy. IM-MS and drift-time selective IR spectroscopy as well as data analysis were performed according to the procedures described in chapter 2.6. All peptide samples for IM-MS and IRMPD spectroscopy were dissolved in water/methanol (1/1, v/v) to give solutions at concentrations in the range of 0.15-2 mM.

IR features at 1700-1800 cm^{-1} were fitted with three Gaussians (not shown in figures). These features are expected to arise from the carboxylic acid moieties. The remaining amide I features at 1600-1700 cm^{-1} were fitted with six Gaussians (1-6) using constraints in the peak center (x_C) and width (σ) as shown in **Table 6.3**. The constraints were set based on the known IR band positions for the secondary structures shown in **Table 2.1**. The one additional Gaussian at 1560-1600 cm^{-1} (0) is used to correct the baseline below 1600 cm^{-1}.

Table 6.3: Constraints in the peak center and width for the fitting amide I with multiple Gaussians

n	Peak center (cm^{-1})	Peak width σ (cm^{-1})
0	1560-1600	0-30
1	1600-1630	0-30
2	1620-1640	0-30
3	1640-1660	0-30
4	1650-1680	0-30
5	1660-1690	0-30
6	1680-1710	0-30

R-squared values (R^2) for all fits are larger than 0.989. The fraction of β-sheet IR signature $F(\beta)$ in the amide I region is determined from the areas of fitted Gaussians (1-6) as equation 6.1, where $A(n)$ denotes the area of *n*-th Gaussian.

$$F(\beta) = \frac{A(1) + A(2) + A(6)}{\sum_{n=1}^{6} A(n)} \tag{6.1}$$

6.3 Results and Discussion

The hexapeptide VEALYL is well known to form amyloid fibrils.[30-31, 208, 210] This manifests itself in the formation of an insoluble deposition after the incubation of VEALYL in water/methanol at room temperature. **Figure 6.1**a shows a TEM of amyloid fibrils that are formed from the VEALYL sample. The secondary structure of non-soluble VEALYL fibrils was previously found to be dominated by β-sheets.[30-31, 208-210] Solid-state FT-IR spectroscopy, especially in the amide I region where C=O vibrations are probed, is sensitive to the secondary structure. Highly ordered antiparallel β-sheets are expected to yield amide I bands at 1610-1640 cm^{-1} and ~1690 cm^{-1}, whereas signals at 1640-1650 cm^{-1} and 1660-1685 cm^{-1} indicate random coils and several types of turn structures, respectively.[65-66] A comparison between FT-IR spectra of freshly dissolved VEALYL and the precipitate is shown in **Figure 6.1**b. For the freshly dissolved sample (black curve), the maximum of the amide I band at 1641 cm^{-1} together with several broad features up to 1700 cm^{-1} indicate the presence of unordered and turn-like structures. In contrast, the IR spectrum of the VEALYL precipitate (red curve) shows a narrower and red-shifted amide I band centered at ~1627 cm^{-1}. This suggests that the VEALYL precipitate predominantly consists of highly ordered β-sheets. In addition, UV spectroscopy and FT-IR experiments in solution (**Figure D.2**) show that aggregation occurs after a certain lag phase, which is characteristic for nucleation-dependent fibril formation.[219]

A disadvantage of condensed-phase methods is that in solution they can only provide ensemble-averaged results of all aggregation states. Structural details of the individual pre-fibrillar oligomers are, thus, impossible to obtain with these methods. To address this limitation, IM-MS-based methods were used in the present work. These methods allow a clear separation of oligomer number (n) and conformation and provide absolute $^{DT}CCS_{He}$. When a freshly dissolved VEALYL sample is electrosprayed, a wide distribution of multiply charged peptide oligomers is observed (**Figure D.3**a, b and **Figure D.4**). The $^{DT}CCS_{He}$ values of each individual VEALYL conformer were measured, and the results are shown in **Figure 6.1**c. The dashed line represents an ideal isotropic (that is, globular) growth behavior (see chapter 2.4.3).[63] This $^{DT}CCS_{He}$ plot clearly shows the polydisperse and polymorph nature of VEALYL oligomers, and features oligomers with various stoichiometries ranging from the singly protonated monomer (n/z = 1/1) to the quadruply protonated decamer (n/z = 10/4). The inset in **Figure 6.1**c shows an ATD of the triply protonated tetramer (n/z = 4/3) for which multiple species with different $^{DT}CCS_{He}$ were observed. Interestingly, VEALYL oligomers

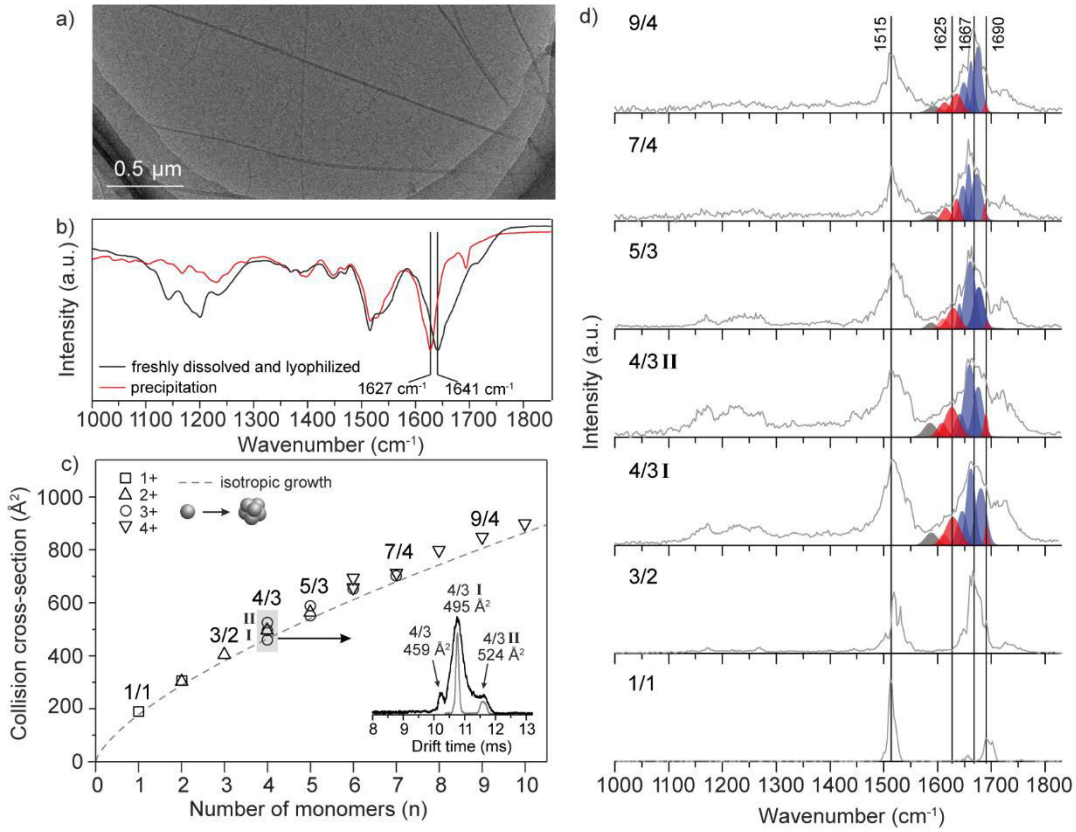


Figure 6.1: Structural analysis of VEALYL oligomers and fibrils in the condensed phase and in the gas phase. a) TEM of an incubated VEALYL solution. b) Solid FT-IR spectra of the VEALYL precipitate obtained after two days of incubation at room temperature and of the freshly dissolved VEALYL without precipitation after lyophilization. c) Measured $^{DT}CCS_{He}$ values of VEALYL oligomers as a function of the number of monomers. The dashed line denotes the expected $^{DT}CCS_{He}$ assuming isotropic growth. The inset shows a typical ATD of the triply charged tetramer ($n/z = 4/3$) in which multiple conformers with distinct $^{DT}CCS_{He}$ values are observed. The statistical error of the $^{DT}CCS_{He}$ measurements is less than 1% and smaller than the size of the symbols. d) IRMPD spectra of m/z- and drift-time-selected VEALYL oligomer ions from the singly protonated monomer ($n/z = 1/1$) up to the quadruply protonated nonamer ($n/z = 9/4$) measured in the wavenumber range 1000-1850 cm^{-1}. Fractions of the amide I band that are representative for β-sheets (1610-1640 cm^{-1} and 1690 cm^{-1}) and unordered or turn-like (1660-1685 cm^{-1}) structures are represented in red and blue, respectively.

larger than the trimer ($n > 3$) exhibit $^{DT}CCS_{He}$ values that are larger than predicted by the isotropic model. This deviation from the isotropic growth line may be indicative of the presence of soluble β-sheet-rich intermediates on the way to fibril formation, as suggested by previous studies.[31, 63]

Although ion-mobility data contain information about the overall size of oligomer ions, they cannot provide clear evidence for a specific secondary structure. To obtain information about structural details of individual oligomers, the IM-MS instrument was used

as a preparative tool to pre-select ions of a specific size and m/z, followed by irradiation with intense IR laser pulses generated by the FHI-FEL.[69] An IRMPD spectrum can then be generated from the wavelength-dependent dissociation of the pre-selected ions. The resulting IRMPD spectra of individual VEALYL oligomers up to the quadruply protonated nonamer (n/z = 9/4) measured in a wavenumber range of 1000-1850 cm^{-1} are shown in **Figure 6.1**d. The spectra consist of two major bands at 1600-1700 cm^{-1} and 1450-1550 cm^{-1}, which can be assigned as amide I and II vibrations, respectively. IR features above 1700 cm^{-1} originate from carboxylic acid vibrations either from the C-terminus or from the side chain of glutamic acid. Weaker features in the wavenumber range 1100-1300 cm^{-1} can be assigned to Tyr and amide III vibrations.[66, 220] For the singly protonated monomer (n/z = 1/1), well-separated features can be observed in the amide I region at 1655, and 1695-1705 cm^{-1}, as well as a very weak feature at 1750 cm^{-1}. Compared with the amide I bands of higher oligomers, these monomer features are found at higher wavenumbers. This result is not surprising because the extent of hydrogen bonding inside the hexapeptide monomer is expected to be rather low and, thus, provide little perturbation of the individual C=O oscillators. As the size of the oligomer increases, amide I features are shifted towards lower wavenumbers. The spectra of higher oligomers (n = 3-9) feature amide I bands that are centered around 1660-1670 cm^{-1}, which is indicative of turn-like structures. Most interesting, however, is the observation of typical β-sheet IR bands in the amide I region (1610-1640 cm^{-1}) for the tetramer and higher oligomers, which clearly have larger $^{DT}CCS_{He}$ values than predicted by the isotropic growth model. To ascertain qualitatively the relative abundances of β-sheets and turn-like structures, the amide I region (1600-1700 cm^{-1}) of these oligomers (n = 4-9) was fitted with multiple Gaussian curves (**Figure 6.1**d, for details see **Table D.1** and **Table D.2**). Each of these Gaussians represents a specific motif: β-sheets at 1610-1640 cm^{-1} and ~ 1690 cm^{-1} (**Figure 6.1**d, red curve) and small turns at 1660-1685 cm^{-1} (blue curve). This analysis clearly indicates a significant proportion of β-sheet in higher VEALYL oligomers of $n \geq 4$.

Compared with the clear evidence for β-sheets in the solid FT-IR spectra of the mature VEALYL fibrils (**Figure 6.1**b), however, β-sheet signatures in the IRMPD spectra of VEALYL oligomers (**Figure 6.1**d) are weaker and exhibit little variation over the different oligomer sizes. This makes it difficult to correlate the IRMPD results with the observed $^{DT}CCS_{He}$ values and to pinpoint specific oligomers that are involved in structural transitions. To generate comparable peptides that may differ in their amyloid formation propensity, the VEALYL-sequence variants VELYAL and YVEALL were synthesized and investigated.

Keeping the overall amino acid composition unchanged helps to minimize differences in the IR spectra that arise due to different amino acid side-chain absorptions and greatly facilitates the comparison between different sets of IR data.

As shown in the TEM image in **Figure 6.2**a, VELYAL does not appear to form amyloid fibrils. Similarly, no visible precipitate was formed after several days of incubation, which was further confirmed by UV spectroscopy (**Figure D.5**a, b). Time-resolved FT-IR measurement of a VELYAL solution (**Figure D.5**c, d) did not indicate a structural transition into β-sheets. Similarly, the solid-phase FT-IR spectra of fresh and incubated VELYAL samples both exhibit an amide I peak at 1638 cm^{-1} and broad features up to 1750 cm^{-1}, which indicates that mostly unordered structures are present and no β-sheets are formed (**Figure 6.2**b). Taken together, these results show that VELYAL is a non-amyloid-forming sequence, which makes it an ideal reference for further comparison. VELYAL does form oligomers that can be investigated further by IM-MS and IRMPD (**Figure D.3**c, d and **Figure D.6**). However, their $^{DT}CCS_{He}$ values more closely follow the isotropic growth curve than those of either of the other two peptides, which indicates a compact and rather globular set of conformations (**Figure 6.2**c). In addition, the IRMPD spectra show almost no sign of β-sheet formation and point to a predominantly unordered or turn-like conformation (**Figure 6.2**d and **Table D.4**). These results emphasize that the increased β-sheet content observed for VEALYL is a real effect and not just an artifact of the IRMPD measurements.

In stark contrast to the non-amyloidogenic variant VELYAL, the TEM of YVEALL clearly point to the formation of amyloid fibrils (**Figure 6.3**a). This is further supported by the FT-IR spectra of the lyophilized precipitate, which exhibit a sharp IR band at 1627 cm^{-1}, indicating the presence of highly ordered β-sheets (**Figure 6.3**b). In addition, time dependent UV spectroscopy (**Figure D.7**a, b) and solution-phase FT-IR experiments (**Figure D.7**c, d) of YVEALL reveal a typical nucleation-dependent fibril growth, which is, however, slower than that of the original sequence VEALYL. This makes YVEALL an ideal candidate for the spectroscopic investigation of amyloid oligomers.

In IM-MS a wide distribution of oligomers of different size and charge are observed for YVEALL (**Figure D.8**). The $^{DT}CCS_{He}$ evolution of different oligomers as a function of the oligomer number n is shown in **Figure 6.3**c. For species up to the tetramer (n = 1-4), only

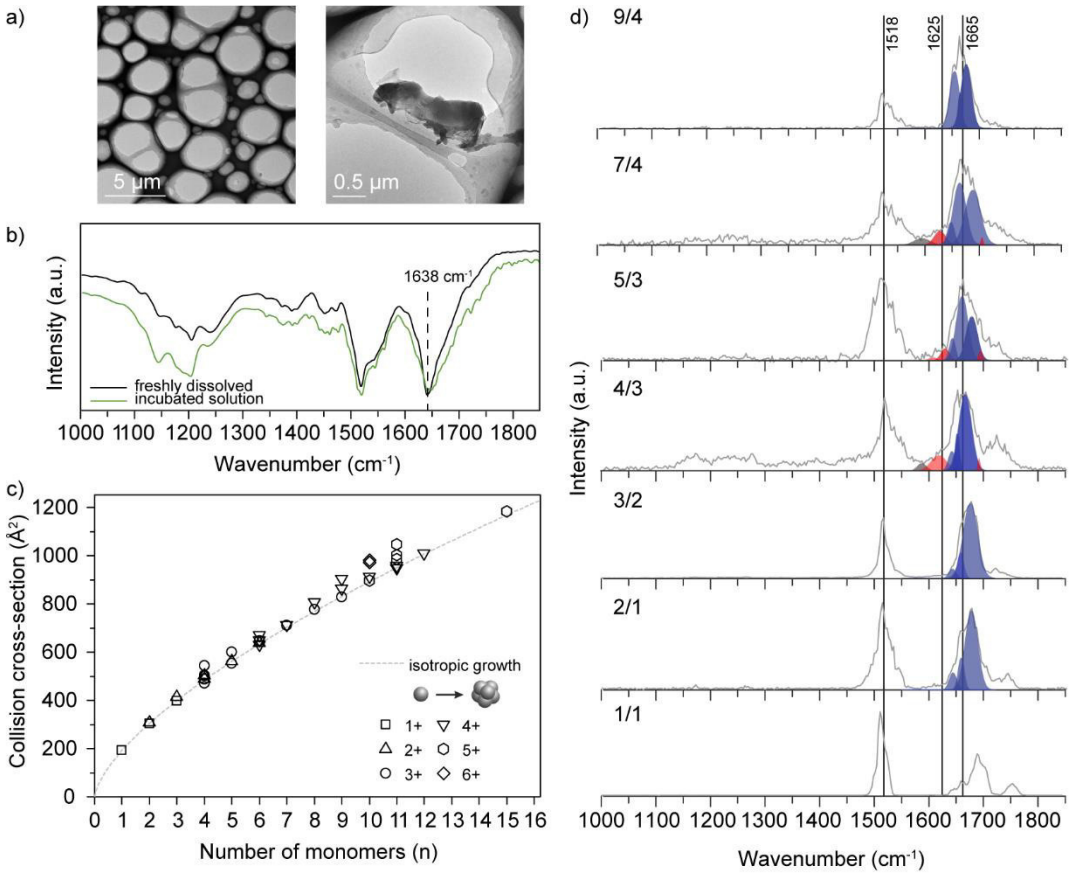


Figure 6.2: Structural analysis of VELYAL oligomers in the condensed phase and in the gas phase. a) TEM of incubated VELYAL solution. b) Solid FT-IR spectra of VELYAL for incubated (green) and freshly dissolved (black) VELYAL after lyophilization. c) $^{DT}CCS_{He}$ values of VELYAL oligomers as a function of the number of monomers. The dashed lines denotes the expected $^{DT}CCS_{He}$ values assuming isotropic growth. The statistical error of the $^{DT}CCS_{He}$ measurements is less than 1% and smaller than the actual size of the symbols. d) IRMPD spectra of m/z- and drift-time-selected VELYAL oligomer ions from the singly charged monomer ($n/z = 1/1$) up to the quadruply protonated nonamer ($n/z = 9/4$) measured in the wavenumber range 1000-1850 cm^{-1}. Fractions of the amide I band that are representative for β-sheets (1610-1640 cm^{-1} and 1690 cm^{-1}) and unordered as well as turn-like (1640-1685 cm^{-1}) structures represented in red and blue, respectively.

little deviation from the isotropic growth model (dashed line) is observed. Larger oligomers ($n > 4$), however, exhibit significantly higher $^{DT}CCS_{He}$ values than predicted by the isotropic model. **Figure 6.3c** inset shows an ATD of the triply protonated pentamers ($n/z = 5/3$), for which at least four distinct conformers can be separated (labeled I-IV). The smallest observed pentamer (5/2) fits to the isotropic model, whereas the largest pentamer (5/3 IV) is ~ 23% larger in $^{DT}CCS_{He}$ compared to isotropic oligomers. YVEALL pentamers are therefore highly

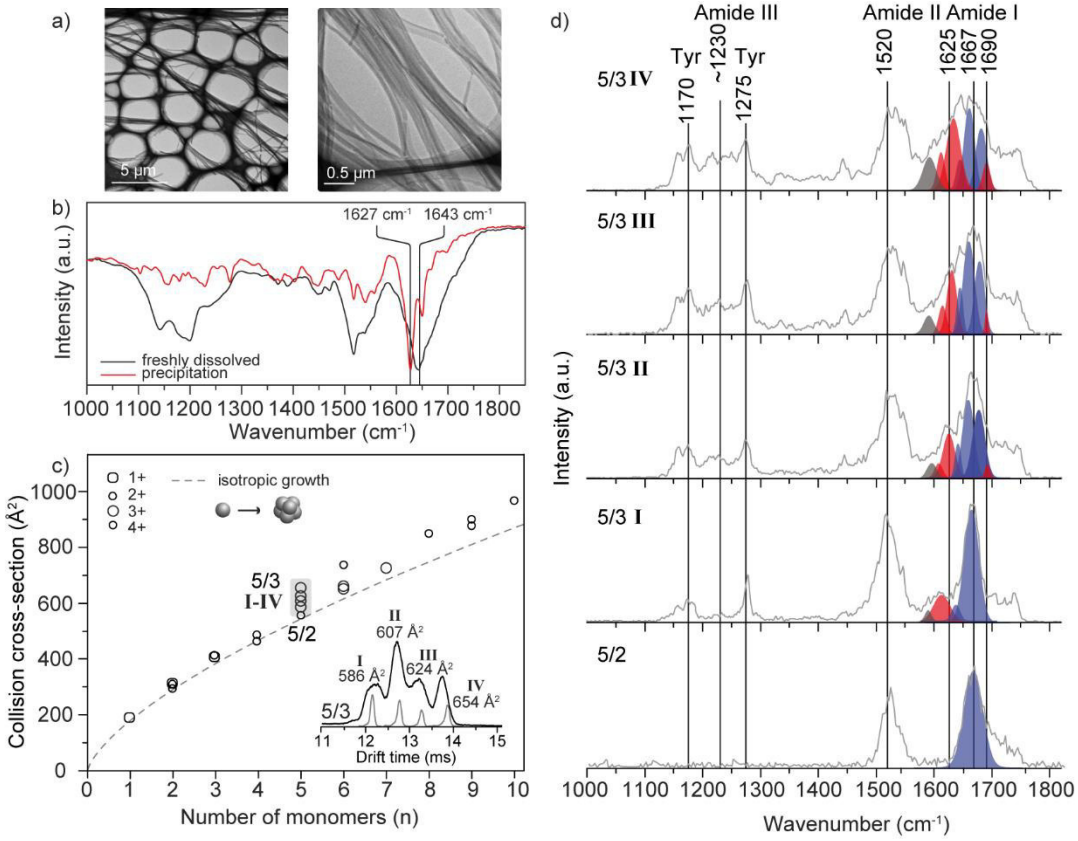


Figure 6.3: Structural analysis of YVEALL oligomers in the condensed phase and in the gas phase. a) TEM of incubated YVEALL solution. b) Solid FT-IR spectra of YVEALL for incubated precipitate (red) and freshly dissolved (black) YVEALL after lyophilization. c) Measured $^{DT}CCS_{He}$ values of YVEALL oligomers as a function of the number of monomers. The dashed line denotes isotropic growth. The statistical error of the $^{DT}CCS_{He}$ measurements is less than 1% and smaller than the actual size of the symbols. The inset shows an ATD of the triply charged pentamer ($n/z = 5/3$) for which multiple conformers with distinct $^{DT}CCS_{He}$ values are observed. The narrow peaks depicted in grey correspond to the portions of the oligomer distribution that were selected for further IRMPD-spectroscopic analysis. d) IRMPD spectra of doubly protonated pentamer ($n/z = 5/2$) and the drift-time selected species (I-IV) of the triply charged YVEALL pentamers ($n/z = 5/3$). The amide I region (1600-1700 cm^{-1}) was fitted by multiple Gaussian peaks. The blue and red Gaussians represent IR bands that correspond to β-sheets (1610-1640 cm^{-1} and 1690 cm^{-1}) and unordered as well as turn-like (1640-1685 cm^{-1}) structures represented in red and blue, respectively.

polymorph and represent a transition from globular species close to the isotropic line to much more extended states.

The IRMPD spectra of each $n/z = 5/3$ conformer are given in **Figure 6.3**d. All spectra show strong amide II and I features in the 1500-1580 cm^{-1} and 1600-1700 cm^{-1} region, respectively, and additionally exhibit pronounced signals at 1100-1300 cm^{-1}, corresponding to Tyr and amide III vibrations. It is interesting to observe how the individual IR bands change

when the pentamers become more extended. The amide II band only undergoes a small blue-shift and broadening as the $^{DT}CCS_{He}$ increases. This is not surprising because the amide II band consists of N–H and C–N bending modes, which are only marginally affected by changes in the secondary structure.[216] The diagnostic amide I band, however, shows a strong variation between different pentamers. For the doubly protonated pentamer ($n/z = 5/2$) the amide I modes are centered at 1667 cm^{-1} and do not exhibit IRMPD features in the typical β-sheet range between 1610 and 1640 cm^{-1}. However, as the $^{DT}CCS_{He}$ of the polymorph pentamers increase (from species I to IV), typical β-sheet amide I features at 1610-1640 cm^{-1} and 1690 cm^{-1} make up a considerable fraction of the overall spectrum. To allow a more quantitative assessment, the IR traces of the individual pentamers (I-IV) were fitted in the amide I region by multiple Gaussians, as shown in **Figure 6.3**d (details are given in **Table D.5**). The curves in red represent the characteristic β-sheet bands, and curves in blue correspond to turn-like features. Clearly, the β-sheet bands increase as the $^{DT}CCS_{He}$ increases. The most extended pentamer, IV, exhibits a significant amount of IRMPD features at 1610-1640 cm^{-1} and 1690 cm^{-1} that are characteristic for the existence of antiparallel β-sheet structures. For further illustration of the increased β-sheet content, difference spectra are shown in **Figure D.9**.

Additional support for a possible transition in secondary structure is provided by the increasing intensities of the IR bands around 1220-1240 cm^{-1} for conformations with increasing $^{DT}CCS_{He}$. These bands can be tentatively assigned to amide III bands typical for antiparallel β-sheets.[66, 220] Additional structural information can be drawn from amino acid side-chain vibrations. According to X-ray crystallography studies, half of all amino acid side chains in amyloid fibrils only weakly interact with other side chains or the peptide backbone. For globular oligomers, on the other hand, extensive interactions between side chains or with the backbone are essential to maintain the compact structure.[31, 208] These characteristics are also reflected in the IRMPD spectra of the corresponding structures, and lead to signatures of free side chains for fibril-like species and less-pronounced features for globular types. This behavior is revealed in the spectra of the polymorph YVEALL pentamers in **Figure 6.3**d. IR bands that can be assigned to C–O–H bending (1170 cm^{-1}) and C–OH stretching (1275 cm^{-1}) vibrations of free tyrosine residues,[66] as well as signals at 1700-1750 cm^{-1} that correspond to C=O stretching vibrations of weakly bound carboxylic acids, can all be identified with increasing intensities for species for species with increasing $^{DT}CCS_{He}$. Taken together, the

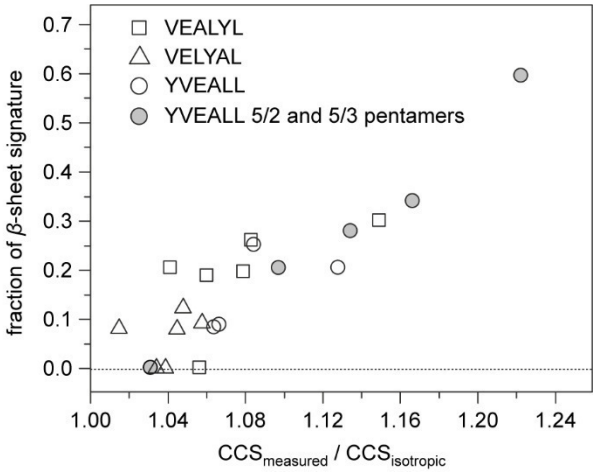


Figure 6.4: Correlation between increase in CCS and β-sheet fraction. Fraction of β-sheet in the amide I band of VEALYL, VELYAL and YVEALL oligomers as a function of relative deviation in CCS from the isotropic growth model. Shaded circles denote the polymorph pentamers (n/z = 572 and n/z = 5/3 I-IV) of YVEALL.

presented results from IM-MS and IRMPD spectroscopic experiments strongly suggest that pentamers of the fibril-forming sequence YVEALL are soluble early intermediates that undergo secondary structural transitions from a compact unordered structure to an extended assembly of repeating β-sheets.

As shown in **Figure D.10**, larger YVEALL oligomers (n/z = 7/3 and 9/4) show considerable β-sheet features, but they are less pronounced compared to the most-extended pentamer (n/z = 5/3 IV). Thus, the oligomer number n does not correlate directly with an increased β-sheet character. This suggests that the formation of larger oligomers is not necessarily initiated by the most-extended pentamer. The complex equilibrium between different oligomeric states is established during the early stages of aggregation - the lag phase - and thus a multitude of oligomers with distinct β-sheet content coexist.

Oligomers rich in β-sheet structure have larger $^{DT}CCS_{He}$ values than globular conformations (isotropic line). Hence, the deviation in $^{DT}CCS_{He}$ values from the isotropic growth model, which is similar to a shape factor[221-223] that describes the structural anisotropy in proteins, more probably correlates with the relative β-sheet character than the oligomer size n.[63] Thus, the fraction of β-sheet signature within the amide I band in all investigated species were calculated by dividing the summed area of β-sheet features in the IRMPD spectra (red curves) by the total area of the amide I band (**Table D.2**, **Table D.4**, and **Table D.6**). The result is shown in **Figure 6.4**. Oligomers of VELYAL barely deviate from the isotropic line (less than 5%) and feature β-sheet signatures that make up less than 10% of the total amide I band. The amyloid forming sequences VEALYL and YVEALL, on the other hand, show more-pronounced deviations in $^{DT}CCS_{He}$ values, with up to 59% β-sheet features in the amide I band for the largest of the four different pentameric species (n/z = 5/3) of YVEALL (filled circles). However, a clear linear correlation between deviation in $^{DT}CCS_{He}$ values from the

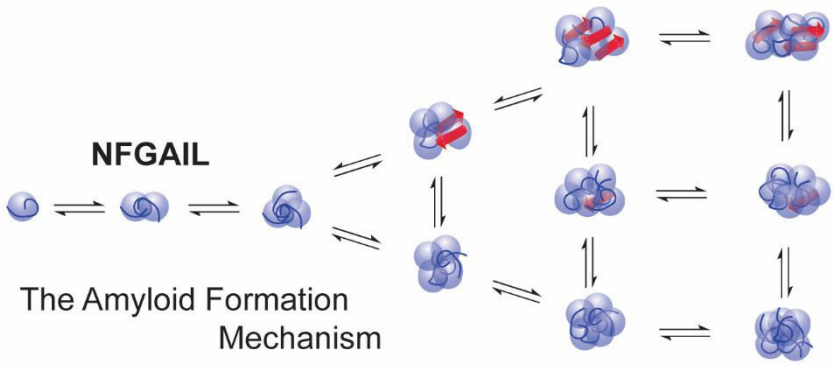
isotropic growth model and β-sheet content does not exist because the intrinsic structural heterogeneity of the system probably contributes significantly to the scatter in this plot. Nevertheless, the data indicate a positive relationship between the relative $^{DT}CCS_{He}$ and the relative amount of ordered β-sheet structure, and therefore support recent IM-MS approaches in which such a correlation was suggested.[9, 63, 214-217]

6.4 Conclusions

Combining IM-MS and IR spectroscopy has enabled us to measure directly the onset of β-sheet structure formation in isolated, early soluble oligomers of the insulin B-chain fragment VEALYL, which was previously shown to form β-sheet-rich amyloid fibrils and microcrystals.[210] In addition, the critical transition from a compact and unordered to an extended β-sheet structure was monitored for highly polymorphic pentamers of the amyloidogenic VEALYL sequence variant YVEALL. Compact YVEALL pentamers were shown to adopt predominantly turn-like or unfolded structures, and a considerably increased content of β-sheet was observed for the more-extended analogues. Further support for this structural assignment was found in amide III and side-chain vibrations, which follow the evolution of β-sheet characteristic amide I features and strongly suggest that the peptide assemblies consist of repeating β-sheets. Taken together, the data represent the first spectroscopic study that successfully identifies specific pre-fibrillar oligomer states involved in the critical transition from an unordered to a β-sheet-rich structure. For the systems investigated here, oligomers that consist of as little as four to nine peptide strands were found to exhibit a considerable fraction of β-sheet.

Furthermore, the results presented here demonstrate that there is a positive correlation between a large $^{DT}CCS_{He}$ that deviates from the isotropic growth model and the presence of β-sheet structures. This correlation greatly strengthens the interpretation given in recent IM-MS studies of amyloid systems[9, 63, 214-217] and demonstrates how IR spectroscopy combined with IM-MS can be of great use in this field of research. These combined techniques offer new possibilities to explore the secondary structures of soluble pre-fibrillar oligomers, which will be greatly beneficial for developing effective strategies to manipulate or hinder the amyloid assembly process[9, 45, 108, 214] and may ultimately help to understand the molecular basis of amyloid diseases.

7 NFGAIL Amyloid Oligomers*



7.1 Introduction

In type II diabetes (T2D) unstructured hIAPP monomers assemble into highly structured, β-sheet rich, insoluble deposits known as amyloid plaques.[224] These plaques are found in more than 95% of T2D patients, which suggests their direct involvement in β-cell dysfunction in the pancreas.[225-228] Recent studies have shown that especially early hIAPP oligomers represent the toxic species.[206, 224, 228-229] Understanding their detailed structure and the mechanism leading from monomer to fibrils is crucial to selectively modulate the assembly pathway.[224, 229-232]

Aggregation prone domains of hIAPP include the fragments 1-8, 8-20, 20-29 and 30-37, all of which form amyloid fibrils.[233-237] While each of these fragments may play a role in the assembly of the parent hIAPP, the 20-29 region has garnered the most attention since chemical modifications[238-239] or mutations[228] in this region disrupt fibril formation. The 22NFGAIL27 fragment within this region is the shortest known hIAPP sequence capable of forming amyloid fibrils at pH ~ 7 and was shown to be cytotoxic towards the pancreatic cell line RIN5fm.[240] In addition, recent evidence suggests the 20-29 domain may play a central role in fibril formation of full-length hIAPP.[241] Thus, many experimental[242-244] as well as theoretical[245-250] studies have investigated the assembly characteristics of NFGAIL. They show that the ionic strength[245] influences the NFGAIL assembly and oligomers are formed due to

* This chapter is based on the work published in W. Hoffmann, K. Folmert, J. Moschner, X. Huang, H. von Berlepsch, B. Koksch, M. T. Bowers, G. von Helden and K. Pagel, *J. Am. Chem. Soc.*, **2018**, *140*, 244-249, DOI: 10.1021/jacs.7b0951. Figures and content adapted with permission. Copyright 2018 American Chemical Society.

attractive, hydrophobic interactions between phenylalanine residues.[244, 247-248, 250] MD simulations further support that hydrophobic interactions lead to the formation of both disordered and structured aggregates.[246, 248] However, the dissociation of those disordered oligomers is the rate-limiting step for the formation of structured aggregates.[249] These can then further evolve into amyloid fibrils composed of antiparallel β-strands.[244, 250] The characterization of soluble oligomers using solution-based FT-IR, CD or NMR spectroscopy only provides ensemble averaged-information rather than information regarding individual oligomeric states. Thus, experimental evidence for the proposed assembly is still missing.

In chapter 6, IM-MS was successfully used as a preselection tool to perform gas phase IR spectroscopy on individual amyloid oligomeric states and conformations.[17] The data showed that the tetramers ($n = 4$) of the two amyloid forming peptides VEALYL and YVEALL undergo a characteristic transition from compact and unordered conformations into more extended and β-sheet rich versions.[17] All of the observed higher-order oligomers ($n \geq 4$) were shown to exhibit an elevated β-sheet content. Here, IM-MS coupled to gas-phase IR spectroscopy is used for the structural characterization of NFGAIL oligomers.

7.2 Experimental Details

Peptide Synthesis: The hexapeptide NFGAIL was synthesized manually by Kristin Folmert and Dr. Johann Moschner (Freie Universität Berlin, Berlin) according to standard Fmoc-chemistry using a preloaded TGA resin. Peptide purification was performed on a low-pressure HPLC system. Purity was validated using an analytical HPLC system and high-resolution mass spectrometry. For details see supporting information by Hoffmann et al.[16]

Thioflavin T Assay: The ThT assay was performed by Kristin Folmert (Freie Universität Berlin, Berlin). Briefly, an NFGAIL stock solution was prepared by dissolving the purified NFGAIL in HFIP (1,1,1,3,3,3-hexafluoro-2-propanol) (~18 mM) and was further sonicated for 15 minutes to dissolve all preformed aggregates. Aliquots of this stock solution were dried and then redissolved to a final concentration of 4 mM in ammonium acetate buffer (10 mM, pH ~ 7), containing 20 μM ThT. After dissolution, the sample was sonicated for 30 seconds and then incubated at 37 °C with 1300 rpm.

Fluorescence spectra were recorded using a 1 cm-path length quartz cuvette (Hellma, Müllheim, Germany) and a luminescence spectrometer LS50B (Perkin-Elmer, Boston, MA, USA). Spectra were recorded at room temperature from 470-500 nm after excitation at 420 nm (excitation slit width 5 nm; emission slit width 10 nm; scan speed = 300 nm/min;

accumulations = 3). The fluorescence intensity at 485 nm was normalized with respect to its maximum value.

Transmission Electron Microscopy Characterization: TEM images have been recorded by Dr. Xing Huang (Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin) and Dr. Hans von Berlepsch (Freie Universität Berlin, Berlin). Briefly, sample preparation was performed according to the ThT assay. The solution was further incubated at 37 °C and shaken with 1300 rpm for two days. During the incubation time, the formation of insoluble deposition was observed. Samples for staining electron microscopy were prepared by adsorbing 5 μL aliquots of peptide solution to glow-discharged carbon-coated collodium films on 400-mesh copper grids. The grids were blotted, stained with 1% phosphotungstic acid, and air-dried. TEM micrographs were taken by a FEI Talos L120 TEM operated at 120 kV for morphological characterization.

IM-MS and Drift-Time Selective IR Spectroscopy: The purified peptide was dissolved in ammonium acetate (10 mM, pH ~ 7) to yield a final peptide concentration of 1 mM. For *n*ESI, ~ 8 μL of sample were loaded into in-house prepared Pd/Pt-coated or Au-coated borosilicate capillaries, and then electrosprayed (positive ion mode, 0.6-1.0 kV) on two different IM-MS instruments: 1) HiRes[251] located in Santa Barbara (Resolution, R = 100), which allows a high IMS separation of individual conformations. 2) iMob[70] located in Berlin (described in chapter 2.6) (R = 40), which is connected to the FHI-FEL (described in chapter 2.5.2) and therefore enables the measurement of conformer- and mass-selected IR spectra.

Data Analysis: The fraction of β-sheet IR signature (F(*β*)) was visualized by fitting the amide I region with six Gaussians (1-6) using constraints in the peak center (x_c) and width (σ) as shown in chapter 6 in **Table 6.3** and calculated according to equation 6.1. The constraints were set based on known IR band positions for the individual secondary structure elements as shown in **Table 2.1**. One additional Gaussian at 1560-1600 cm^{-1} (0) is used to correct the baseline below 1600 cm^{-1}. For the calculation of β-sheet related IR frequencies a narrow IR window, as provided by reported condensed-phase experiments,[65-66] was used. For other secondary structural elements, a slightly expanded window was used to allow a better fit to the experimental data and to avoid underestimation of those elements. Thus, windows related to turns, helices or unordered structures were chosen to slightly overlap with each other, whereas β-sheet related IR windows were set adjacent to those. This further reduces an overestimation of the β-sheet IR content. The *R*-square values (R^2) for all fits are larger than 0.94 (**Table E.1**).

7.3 Results and Discussion

The hexapeptide NFGAIL is the shortest fragment of hIAPP known to form amyloid fibrils at neutral pH.[240] Those fibrils have been shown to feature antiparallel β-strands and a morphology similar to full length hIAPP fibrils.[243, 250, 252] In order to study the β-sheet formation tendency of early oligomers, the NFGAIL peptide was incubated for two days in ammonium acetate buffer. This procedure lead to the formation of fibrillar aggregates with anisometric cross-section, a mean apparent diameter of 20-30 nm and few hundreds of nanometers length (**Figure 7.1**a, b). Typical protofilaments have not been detected. The kinetics of fibril formation was monitored in real-time using ThT assay.[43] ThT intercalates into the cavity of amyloid fibrils, which leads to an increased fluorescence. Even for high concentrations (4 mM) the hexapeptide NFGAIL follows a nucleation-dependent growth mechanism (for details see supporting information by Hoffmann et al.[16]), characteristic of classical on-pathway amyloid formation. The sigmoidal growth behavior can be divided into lag-, growth-, and saturation phases. During the lag-phase (20 hours), early soluble NFGAIL intermediates are present.[229, 240] These transient oligomers are highly polydisperse, *i.e.* a multitude of oligomeric states coexist. In addition, they are most likely polymorph, meaning they occupy a wide range of different conformations for one oligomeric state. All species can interconvert or undergo an exchange of subunits, but once a so-called nucleus is formed, the auto-catalytical growth phase is initiated. The nuclei presumably act as a template to assemble monomers into mature fibrils, which are present at the saturation phase (see reference[101] for an example).

The mass spectrum (**Figure 7.1**c) shows that NFGAIL immediately forms a wide distribution of oligomers, spanning from a singly charged monomer ($n/z = 1/1$, where n = oligomer number, z = charge) up to a quadruply charged 13-mer ($n/z = 13/4$). All of these oligomers carry much less charge than their number of peptide strands and are therefore presumably derived from solution and not disturbed by Coulomb-repulsion. IM-MS further not only reveals the presence of multiple conformations but also higher oligomeric states of the same m/z value (**Figure E.1** and **Figure E.2**). The experimental $^{DT}CCS_{He}$ values of all these species (**Table E.2**) are displayed as a function of the oligomer number in **Figure 7.1**d, where the solid line represents an ideal isotropic (*i.e.* globular) growth behavior (see chapter 2.4.3).[63] Oligomers of which the $^{DT}CCS_{He}$ lie on the isotropic line adopt compact and spherical conformations that consist of turn-like and/or unordered structures. Extended oligomers that deviate from the isotropic growth model trend, on the other hand, are often partially

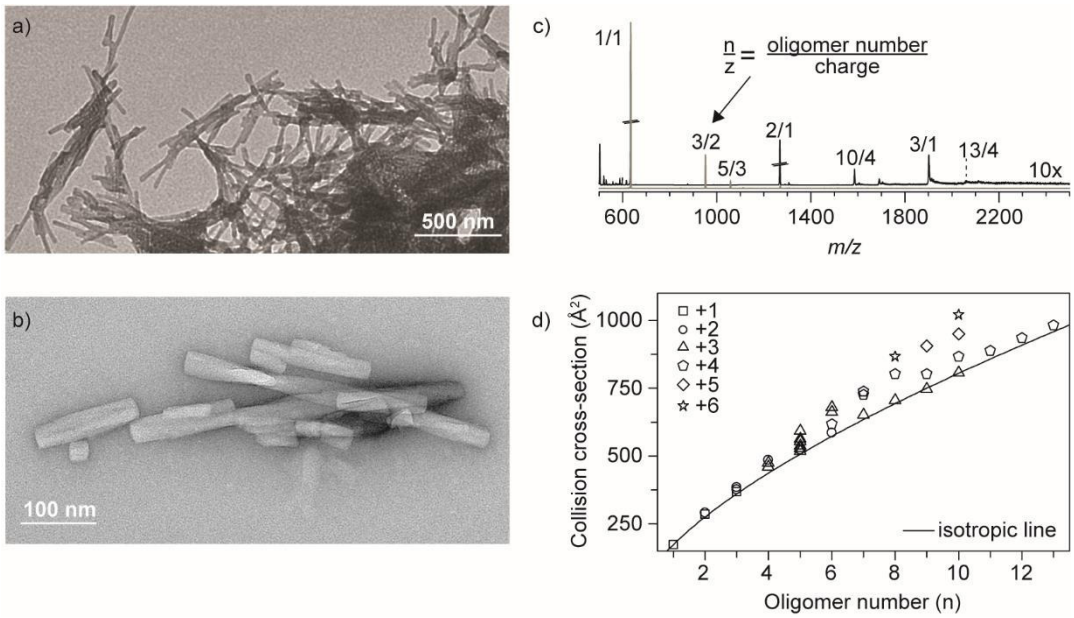


Figure 7.1: Structural investigation of NFGAIL oligomers. a, b) Negative staining TEM images of a NFGAIL solution (4 mM) in ammonium acetate buffer (10 mM, pH ~ 7) show the formation of amyloid fibrils. Scale bars: a) 500 nm and b) 100 nm. c) Mass spectra of freshly electrosprayed NFGAIL solution recorded at two different instrumental settings on the iMob instrument (gray and black line) indicate the presence of multiple oligomers, with n being the oligomer number and z the charge. d) Experimental $^{DT}CCS_{He}$ values as a function of the oligomer number n. The solid line indicates theoretical CCS_{He} expected for an idealized spherical growth. The experimental error of each $^{DT}CCS_{He}$ measurement is less than 1% and smaller than the size of the symbols.

structured and form helical or β-sheet-rich conformations.[17, 63] **Figure 7.1**d shows a variety of compact (experimental $^{DT}CCS_{He}$ values close to the isotropic line) NFGAIL oligomers, starting from the singly charged monomer ($n/z = 1/1$) up to the quadruply charged 13-mer ($n/z = 13/4$). These compact oligomers presumably adopt turn-like or unordered conformations. Interestingly, multiple conformations are present starting from the tetramer ($n = 4$). For example, the conformations of pentamer range from compact ($n/z = 5/2$ with 523 Å^2 and $n/z = 5/3$ with 520 Å^2) to highly extended structures ($n/z = 5/3$ with 540 and 566 Å^2, respectively). Similar extended conformations are also observed for higher oligomeric states (n = 6-10). Thus, the NFGAIL tetramer might represent a structural transition point from unordered/turn-like conformations to at least partially folded, presumably β-sheet rich structures. This observation is consistent with recent results on the 8-20 fragment where a significant β-sheet content is observed starting from the tetramer.[253] In full-length hIAPP, however, a β-hairpin conformer is observed for both the monomer and dimer[94, 96] and is shown to be the conformation that leads to fibrilization. Furthermore, a multidimensional

analytical approach for the full-length hIAPP shows that globular and toxic off-pathway hIAPP oligomers are formed in the presence of copper ions, but in neat hIAPP a different assembly pathway starting at tetramer is observed.[111] Another IM-MS study also observed more extended hIAPP versions for $n \geq 4$.[9]

Although IM-MS provides information regarding the stoichiometry and the overall size of amyloid oligomers, yet direct details of the fine structure cannot be deduced. Instead, IM-MS can be used to pre-select individual conformations for a subsequent analysis by orthogonal techniques such as gas phase IR spectroscopy. The combination of IM-MS and IR spectroscopy allows the individual characterization of m/z- and drift-time-selected species on the secondary- as well as the quaternary-structure level. Amide I vibrations (C=O stretching modes) are highly sensitive towards the secondary structure adopted by peptides and proteins.[65-66, 254] However, individual amide I features usually overlap due to the number of vibrational modes. A deconvolution procedure is therefore typically used in conventional condensed-phase spectroscopy to deduce the relative content of each motif.[255-256] In contrast, such relative populations cannot be directly obtained from gas phase IRMPD spectra without making assumptions about the oscillator strengths. The IRMPD process is based on the sequential absorption of multiple photons and the resulting IR intensity does not scale linearly with the number of absorbed photons. A deconvolution of IRMPD spectral features may therefore not be fully quantitative but still be a reasonably good representation of the different structural features contributing to a complex spectral band. Hence, while "quantitative" percentages of β-sheet content in the amide I band are provided, these actually represent the relative amounts between spectral bands for differing conformations or oligomeric systems but are less quantitative within a given band.

Figure 7.2a shows a set of IRMPD spectra of compact NFGAIL oligomers (n = 2-11) measured in the wavenumber range from 1400-1800 cm^{-1}. All spectra feature two bands assigned as amide II (1480-1540 cm^{-1}) and amide I (1600-1720 cm^{-1}). The amide I region was fitted with multiple Gaussian curves (see data analysis), where antiparallel β-sheets and turn-like motifs are labelled in red and blue, respectively. The center of the amide I band of the singly charged dimer (n/z = 2/1) appears at 1683 cm^{-1}, whereas for higher oligomers (n = 3-11) this feature occurs at lower wavenumbers. The extent of hydrogen bonding within the dimer is apparently lower than for larger oligomers and therefore provides less perturbation for the individual C=O oscillators. The amide I band of higher compact

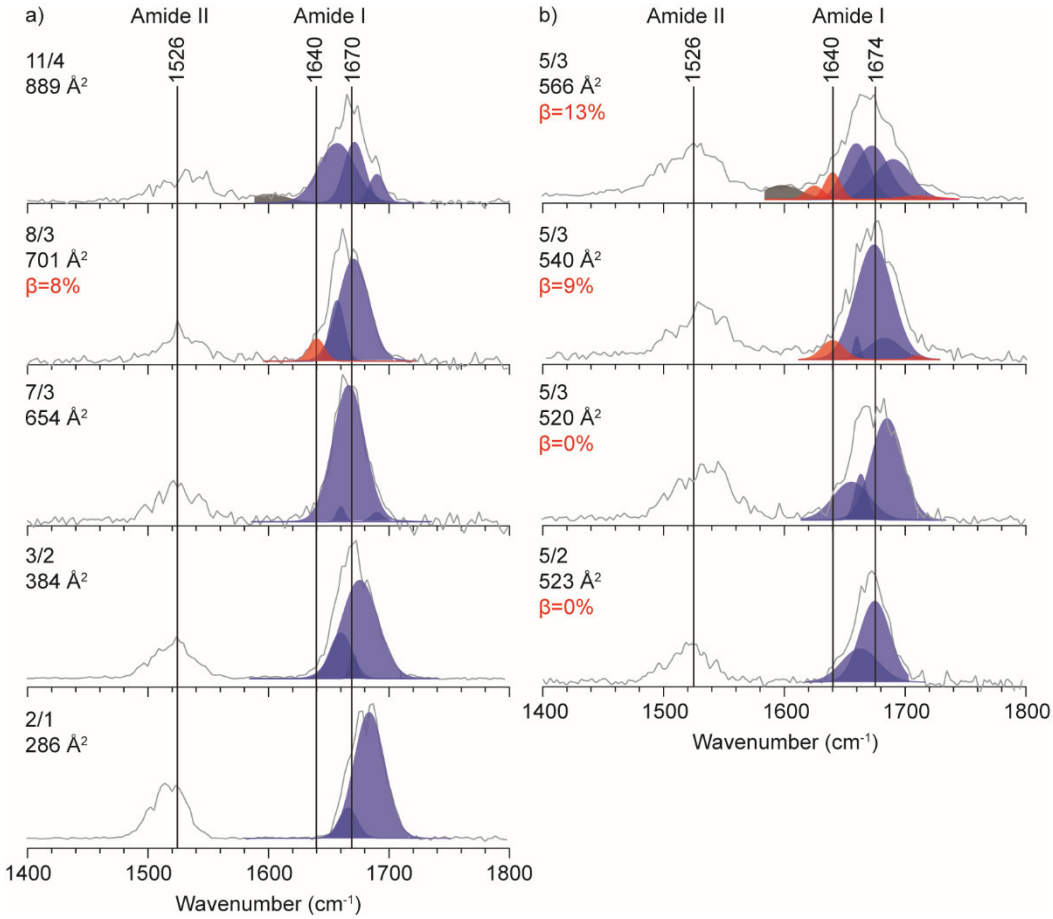


Figure 7.2: IRMPD spectra of m/z- and drift-time selected (a) compact NFGAIL oligomers, exhibiting experimental $^{DT}CCS_{He}$ values similar to theoretical values predicted by the isotropic growth model, and (b) pentamers ranging from compact ($n/z = 5/2$ with 523 Å² and $n/z = 5/3$ with 520 Å²) to extended ($n/z = 5/3$ with 540 Å² and 566 Å², respectively) structures. The amide I band (1600-1700 cm^{-1}) was deconvoluted with multiple Gaussians representing β-sheets (red) or turn-like (blue) conformations. One additional Gaussian between 1560-1600 cm^{-1} (grey) was included to compensate for an overlapping amide I and amide II band. For details about the fitting procedure see data analysis.

oligomers appears around 1670 cm^{-1}, which is indicative of turn-like structures.[17, 65-66] This observation agrees with the fact that the $^{DT}CCS_{He}$ of these oligomers represent compact, spherical structures. Interestingly, an additional shoulder around 1640 cm^{-1} emerges for the triply charged octamer (n/z = 8/3), indicating a β-sheet content of approximately 8%. The $^{DT}CCS_{He}$ of this octamer ($n/z = 8/3$ with 701 Å²), however, fits well with the theoretical value for a spherical, compact conformation (694 Å²). A small β-sheet content observed *via* IRMPD may not necessarily be correlated with an extended structural ensemble.

The first NFGAIL oligomers that considerably deviate from the isotropic line are pentamers. They are highly polymorph and range from compact ($n/z = 5/2$ with 523 Å² and

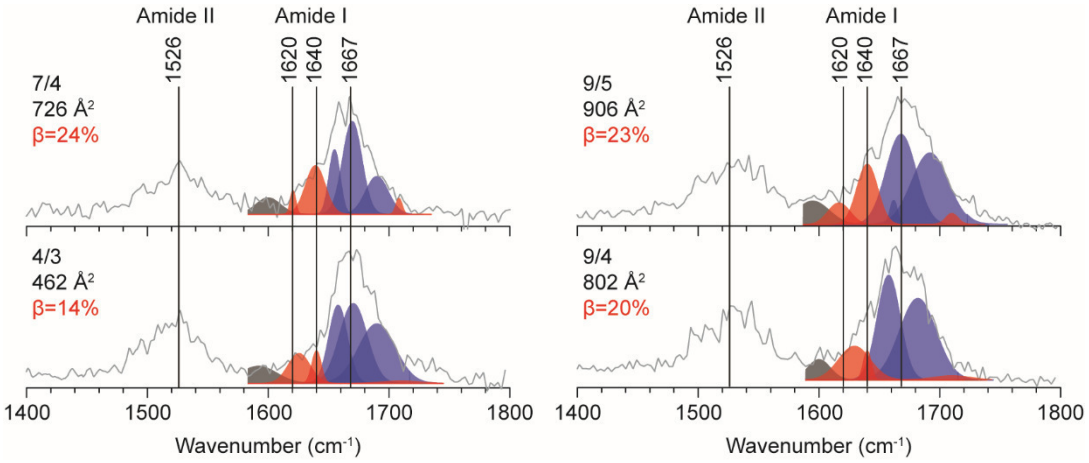


Figure 7.3: IRMPD spectra of extended NFGAIL oligomers. The amide I band (1600-1700 cm^{-1}) was deconvoluted with multiple Gaussians representing β-sheets (red) or turn-like (blue) conformations. Details of the fitting procedure are given in data analysis.

$n/z = 5/3$ with 520 $Å^2$) to highly extended ($n/z = 5/3$ with 540 $Å^2$ and 566 $Å^2$) conformations. The individual, conformer-selective IR spectra show an amide I band at 1674 cm^{-1}, which indicates a predominantly turn-like structure present for all pentamers (**Figure 7.2**b). Interestingly, the more extended pentamer isoforms exhibit a broader amide I band than its corresponding compact species and also show additional features at around 1620-1640 cm^{-1} and 1700 cm^{-1} (red Gaussians). These features are indicative of antiparallel β-strands. An amide I deconvolution reveals a β-sheet IR fraction of up to 13%. This observation is in good agreement with previous results, which suggest a higher β-sheet content for more extended conformations.[17, 63] As a result of the non-linear IR absorption process, however, it is not possible to absolutely quantify how many single peptide strands within the NFGAIL pentamer contribute to the β-sheet fraction in the IR spectrum. For a short peptide such as NFGAIL the β-sheet IR fraction originates from the non-covalent assembly of at least two individual peptide strands. In addition, a recent 2D-IR study on the full-length hIAPP further suggests that β-sheet-rich hIAPP oligomers present in the lag phase are composed of less than four β-strands.[241] Thus, the here investigated NFGAIL oligomers are likely to be composed of at least two or three β-strands.

The gas phase IR spectra of other extended NFGAIL oligomers (n = 4-9) also show a broad amide I band, with elevated intensities at 1617-1640 cm^{-1} (**Figure 7.3**). The Gaussian deconvolution analysis reveals a significant β-sheet content of up to 2%. There is a significant increase in β-sheet content from $n = 4$ (14%) to $n = 7$ (24%) but no smooth trend as n increases hence factors beyond oligomer size contribute to β-sheet formation. A complex equilibrium between oligomeric states and conformations with different β-sheet content is

established in the early stages of NFGAIL aggregation (lag phase). Secondary structural transitions might therefore occur over a multitude of oligomeric states as observed for other amyloid forming sequences such as NNQQNY.[63] Similarly to NFGAIL, also the full-length hIAPP peptide shows a diverse free energy landscape, which is more complicated than a simple transition from a random-coil structure to a perfect fiber-like nucleus.[241] Thus, the short NFGAIL peptide might serve as a good model system and provides critical insights into the full-length hIAPP assembly mechanism. In addition, the coexistence of both turn-like as well as partially β-sheet rich structured aggregates for n = 4-10 is in agreement with a previous theoretical study, which showed that unstructured NFGAIL oligomers are initially formed and their rate-limiting step for the assembly into higher, β-sheet rich oligomers is the subsequent dissociation.[249] The extended β-sheet rich NFGAIL oligomers investigated here, therefore, most likely represent on-pathway oligomers to amyloid fibrils.

7.4 Conclusions

A combination of condensed-phased methods such as ThT assay and TEM, with gas phase methods such as IM-MS coupled to gas-phase IR spectroscopy was used to study the kinetics and oligomeric structural evolution occurring in NFGAIL fibril formation. Under buffered conditions, NFGAIL follows a nucleation-dependent growth mechanism into mature fibrils. IM-MS analysis reveals that during the lag phase a variety of oligomeric states n and different conformations ranging from compact (for n = 2-13) to extended (for n = 4-10) structures are present. Interestingly, also for the full-length hIAPP peptide[9, 111] and other amyloid forming systems[17, 63, 96, 257] a similar transition from compact to more extended versions has been observed. In all cases, the conversion typically starts for oligomers as little as two to nine subunits.[9, 111] The early transition into more extended, presumably β-sheet rich structures might be therefore a general feature of amyloid forming systems.

The gas phase IR analysis of compact NFGAIL oligomers (n = 2-11) shows an amide I band centered at 1670 cm^{-1}, associated with turn-like structures. Extended NFGAIL oligomers (n = 4-9), however, exhibit additional IR features at 1617-1640 cm^{-1}, which are representative of β-sheet rich structures. Deconvolution of the amide I band indicates a β-sheet IR content of up to 24% for extended NFGAIL oligomers.

A previous theoretical study suggests that the formation of unstructured aggregates and their subsequent dissociation is the rate-limiting step to form higher-order, β-sheet rich NFGAIL oligomers.[249] The data presented here support this hypothesis by providing the first

direct secondary structure data for individual oligomers. Due to the complex assembly-disassembly-assembly cascade, the investigated oligomers range from largely unordered to significantly folded β-sheet-containing species and therefore for every extended version a more compact, unordered counterpart coexists. The conformational complexity for each NFGAIL oligomer observed here is in contrast to the in chapter 6 presented study[17] on VEALYL and YVEALL, where conformational complexity was limited to only one or two oligomer sizes. This points to coexistence of on- and off-pathway aggregates in the NFGAIL system and complex assembly/disassembly dynamics that ultimately lead to both an on-pathway β-sheet-dominated set of oligomers leading to fibrils and an off-pathway more isotropic set of oligomers up to at least the size of an 13-mer.

8 Summary and Future Perspectives

In this thesis, a combination of IM-MS with gas-phase IR Spectroscopy were used to investigate properties and the assembly of biomolecules, ranging from amino acids to peptides and proteins. First, a novel amino-acid hydrophobicity scale has been developed that now excludes entropic effects of the solvent and therefore more accurately reflects the intrinsic and true nature of the amino acid hydrophobicity. The detailed knowledge about amino acid hydrophobicity helps to identify protein domains prone to aggregation and it can be used to estimate the amyloid formation propensity of various peptides and proteins.[22, 258] In the past, more than 100 hydrophobicity scales have been introduced, with each being based on a distinct condensed-phase approach.[121-122] The comparison of hydrophobicity values and their relative ranking is, however, not easy, as the interactions between the environment and amino acid are unique to each method. This limitation is overcome here by studying the pairwise interactions of amino acids in the "clean-environment" of the gas phase, where the side-chain polarity determines the self-assembly.[136] Polar residues seek to maximize polar interactions and form more compact clusters than predicted by the idealized isotropic growth model, whereas hydrophobic amino acids assemble into more extended oligomers. Thus, a new correction factor is introduced to account for the deviation from the idealized isotropic growth model and is further used as a basis to generate a novel hydrophobicity scale. In contrast to previous scales, this new approach eliminates entropic effects of the solvent and is based only on the interaction between amino acid monomers. Thus, it more accurately reflects the true hydrophobicity of a side-chain. The novel hydrophobicity scale shows also a good qualitative and quantitative agreement with previous scales based on solvent methodologies, but now can be extended readily to non-natural derivatives, as shown for fluorinated versions.

In addition to the side-chain hydrophobicity, other intramolecular interactions are also crucial for protein folding. In particular, helices are often stabilized in the gas phase by a favorable interplay of charge-dipole interactions, aligned backbone hydrogen bonds as well as side-chain hydrogen bonding to unsaturated C=O groups at the end of the helix.[149, 151, 153] These stabilizing effects within the helical peptide $[Ac\text{-}Ala_{10}Lys + H]^+$ were selectively turned off and their impact on folding was then studied on complementary levels, ranging from the overall molecular shape *via* IM-MS, through secondary-structure sensitive gas-phase IR spectroscopy, up to atomistic resolution MD simulations.[60] Single backbone hydrogen bonds

were deleted by using alanine-to-lactic acid (amide-to-ester) substitutions and side-chain hydrogen bonding to unsaturated C=O groups was inhibited by the non-covalent attachment of a crown ether. The results indicate that a single backbone hydrogen bond only marginally alters the fine structure of the helix, whereas capping has almost no impact on the overall structure and can even slightly perturb helicity due to conformational constraints. Thus, the electrostatic dominated interaction between the charge and the helix macro-dipole is most crucial for the formation of the α-helical conformation in $[Ac\text{-}Ala_{10}Lys + H]^+$ and it is very likely that it is also one of the dominant stabilizing effects for helical structural elements in other peptides and proteins in the gas phase.[60, 151, 154]

Charged residues are not shielded by solvent molecules in the gas phase and electrostatic interactions are enhanced. Thus, Coulomb repulsion between equal charges can impede the natural folding of the protein and consequently alter and destabilize the native protein structure. In how far an *in vacuo* equivalent of the native structure of medium-sized proteins can therefore be maintained in the gas phase is still an ongoing debate. In this thesis, a combination of IM-MS with gas-phase IR spectroscopy was used to investigate the secondary and tertiary structure of myoglobin (mainly helical) and β-lactoglobulin (β-sheet rich), which were carefully transferred from solution to the gas phase. The gas-phase experiments were supported by condensed-phase secondary-structure sensitive CD and FT-IR spectroscopy measurements.[254] The data indicate that an increasing amount of charges on the protein destabilizes the native fold and yields more extended structures as indicated by higher ion mobility collision cross-sections. When the original native structure is predominantly helical, IR spectroscopy shows that the unfolded species remains helical. Helices were shown to be exceptionally stable in the gas phase[91] because they can maintain most of their hydrogen bonds while maximizing the distance between equal charges. In contrast, β-sheet-rich proteins cannot maintain their hydrogen bonds due to the Coulomb-repulsion driven unfolding and therefore undergo a conformational transition into helix-rich structures with increasing charge. Finally, at very high charge states even these helices unravel and form a string-like structure that is governed by Coulomb-repulsion and C_5-hydrogen bonds. Furthermore, the formation of string-like structures is independent from the initial native protein structure and therefore it can be seen as a new secondary structural element that highly-charged proteins can excess under vacuum conditions.[87-88, 254] In contrast, when proteins are carefully transferred from solution into the gas phase, lowly charged species with compact sizes comparable to their respective native condensed-phase structures can be observed. Furthermore, m/z- and

conformer-selected gas-phase IR spectra of these species show a remarkable agreement with their respective condensed-phase FT-IR and CD spectra. Thus, the presented data clearly demonstrate that at least structural parts of the condensed-phase secondary as well as tertiary structure can be conserved when solvent is completely removed. Thus, the here-presented study highlights the potential of MS-based methods for the structural investigation of biomolecular assemblies as these methods provide high sensitivity and selectivity which cannot be achieved using traditional condensed-phase techniques.

MS techniques can isolate and characterize a single species in the presence of many others without affecting their underlying equilibrium. This unique advantage is of exceptional use for the structural investigation of aggregating systems. In particular, transient, toxic oligomers formed by amyloidogenic peptides and proteins play a crucial role in a variety of human disorders such as Alzheimer´s disease or type II diabetes. However, these intermediates are highly dynamic, polydisperse as well as polymorph and therefore are very challenging to characterize using traditional condensed-phase methods, providing only ensemble-averaged information. Thus, in this thesis the combination of IM-MS with gas-phase IR spectroscopy was used to provide the first direct secondary-structure analysis of individual amyloid intermediates formed by the amyloidogenic peptide sequences VEALYL[17] and NFGAIL[16] derived from the insulin B-chain and the human islet amyloid polypeptide (hIAPP), respectively. Both hexapeptides were shown to be directly involved in the pathology process of type II diabetes. The data reveal that oligomers of the fibril-forming peptide segment VEALYL and its sequence-scrambled version YVEALL, which consist of as little as four to nine peptide strands, can already contain a significant amount of β-sheet. The conformational complexity for each NFGAIL oligomer is, however, in contrast to the assembly cascade of VEALYL and YVEALL, where conformational complexity was limited to only one or two oligomer sizes. This points to complex assembly/disassembly dynamics in the NFGAIL systems and the coexistence of an on-pathway, β-sheet dominated set of oligomers leading to fibrils and an off-pathway, more isotropic set of oligomers up to at least the size of a 13-mer. Taken together, the two studies have unambiguously identified for the first time specific pre-fibrillar oligomer states that are involved in the critical transition from an unordered to a β-sheet-rich structure. Such information is crucial for developing effective strategies to manipulate or hinder the amyloid formation process and will help to understand the molecular basis of amyloid-related diseases.

To further enhance the quality of experimental data and satisfy the urge of increased resolution with respect to ion mobility drift-time and m/z-separation as well as the separation of vibrational bands, the current setup of the drift-tube IM-MS instrument will undergo further modifications. The ion mobility resolving power can be easily increased by increasing the drift cell length and drift voltage. However, limitations as arcing, lab space limits and ion losses due to diffusion do not allow an unlimited extension of the drift tubes. This problem can be overcome by using new, small and cheap forms of ion optics for structure lossless ion manipulations (SLIM).[259-260] These so-called SLIM devices are based on a printed circuit board technology in which RF and DC potentials are superimposed to central electrodes flanked by guard electrodes to which a DC field is applied. The combined RF and DC fields creates a central conduit in the space between the surfaces in which ions can move or be trapped without any loss in virtually any manner desired by the user. Thus, SLIM devices do not only allow an incredibly high ion mobility resolution but also enable a variety of other ion manipulations such as ion-molecule reactions as well as ion-ion reactions and ion storages, allowing to address a broad field of research. In addition, by mounting a helium cryostat on a SLIM ion mobility device the temperature can be easily controlled and conformational transition as indicated by drift time changes can be studied by varying the temperature. This will further allow to extract defined free energy barriers for conformational transitions and/or will simplify ion mobility arrival time distributions by eliminating high temperature conformations. However, the motion of ions in a RF field within a SLIM device is so far not completely understood and therefore does not allow to calculate absolute collision cross-sections.

Another strategy to measure direct absolute collision cross-sections and to calculate free energy barriers is the implementation of an IMS-Trap-IMS setup, where a stacked hexapole RF ion trap, allowing additional optical access due to the geometry, is installed between two identical ion mobility drift tubes. Free energy barriers of conformational transitions can then be studied by IR activation and measuring the exact laser power, *i.e.* defined energy input, that causes the unfolding as indicated by the second drift tube of identical dimensions. In addition, the new setup would allow to perform action spectroscopy based on monitoring changes in drift times. A conformational transition requires less photons than the respective photofragmentation and therefore would effectively reduce current anharmonicity effects being present due to infrared multiple photon dissociation (IRMPD) and intramolecular vibrational redistribution (IVR) processes. In addition, higher

amyloidogenic oligomeric states (*e.g.* doubly-charged tetramer $n/z = 4/2$) that upon low energy input often yield fragments of the same m/z as the parent (*e.g.* singly-charged dimer $n/z = 2/1$) can then be probed by gas-phase IR action spectroscopy because drift times of individual oligomeric states of the same m/z differ.

The spectral resolution can also be further improved by cooling the ions of interest *via* collisions with cold buffer gas molecules in ion traps. The cooling of the ion of interest effectively depletes higher vibrational states and therefore the molecule is probed from low energy levels that are less disperse. In addition, the analyte ion can further be tagged with a non-covalently, weakly interaction bound species, where the new complex can then be dissociated with relatively little energy intake, thus leading to increased spectral resolution.

The current instrumental design further allows an easy modification of current time-of-flight components that have so-far very low resolution (R~500). Higher m/z-resolving power would allow to unambiguously assign charge states based on the isotopic pattern and therefore help to deconvolute overlapping species (*e.g.* higher oligomeric states) within the same m/z signal.

Overall, the outstanding potential of IM-MS as a tool for the analysis of the structure[94], assembly[52] and dynamics[46] of biomolecular clusters, in particular amyloid oligomers, was demonstrated for various examples. Moreover, its impressive potential for amyloid inhibitor screening (with up to 5000 molecules per day)[45] as well as its unique ability to quantify the mode of interaction of inhibitors[261] was shown. Few other techniques are able to access such a wide range of information on such short time scales, and it is therefore likely that the separation power of *n*ESI-IM-MS will contribute significantly to the elucidation of the assembly pathway and the development of novel amyloid inhibitors in the future. IM-MS on its own is only sensitive to the overall shape of the investigated ions and as such the obtained information content is limited. However, the real strength of the method becomes apparent when it is combined with other orthogonal techniques such as TEM or kinetic assays (*e.g.* HDX).[39, 102] More importantly, also a direct coupling to orthogonal gas-phase techniques can be performed to obtain more detailed structural information. Very recently, IM-MS was for example used as a pre-selection tool to perform FRET[262] spectroscopy and in this thesis for the first time was used to preselect individual amyloidogenic oligomeric states which then have been studied by gas-phase IR[16-17] spectroscopy. These experiments led to new insights into the secondary structure evolution of early amyloid intermediates, and it is likely that similar experiments will in the future be applied to study other biological-relevant, larger amyloid

forming systems such as full-length hIAPP and Aβ as well as the binding mode of inhibitors. In addition, the success of the here-presented methodology of isolating a specific molecular ion out of a complex mixture without changing their underlying equilibrium and then characterizing it with orthogonal techniques such as IR spectroscopy can also be readily applied for other supramolecular systems. In particular, non-biological systems which mimic artificial life are of broad interest and detailed knowledge about their assembly mechanism is crucial to answer fundamental questions about how nature evolved its complex molecular networks and, ultimately, led to the origin of life.

References

(1) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989,** *246*, 64.

(2) Jardine, I. *Nature* **1990,** *345*, 747.

(3) Henry, K. D.; Quinn, J. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1991,** *113*, 5447.

(4) Loo, J. A.; Quinn, J. P.; Ryu, S. I.; Henry, K. D.; Senko, M. W.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89*, 286.

(5) Wittig, S.; Haupt, C.; Hoffmann, W.; Kostmann, S.; Pagel, K.; Schmidt, C. *J. Am. Soc. Mass Spectrom.* **2018**.

(6) Seo, J.; Warnke, S.; Pagel, K.; Bowers, M. T.; von Helden, G. *Nat. Chem.* **2017,** *9*, 1263.

(7) Hoffmann, W.; von Helden, G.; Pagel, K. *Curr. Opin. Struct. Biol.* **2017,** *46*, 7.

(8) Zhong, Y.; Han, L.; Ruotolo, B. T. *Angew. Chem. Int. Ed.* **2014,** *53*, 9209.

(9) Young, L. M.; Cao, P.; Raleigh, D. P.; Ashcroft, A. E.; Radford, S. E. *J. Am. Chem. Soc.* **2014,** *136*, 660.

(10) Wyttenbach, T.; Pierson, N. A.; Clemmer, D. E.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2014,** *65*, 175.

(11) Qiu, B.; Liu, J.; Qin, Z.; Wang, G.; Luo, H. *Chem. Commun. (Cambridge, U. K.)* **2009**, 2863.

(12) Koch, K. J.; Aggerholm, T.; Nanita, S. C.; Cooks, R. G. *J. Mass Spectrom.* **2002,** *37*, 676.

(13) J. Hofmann; H. S. Hahm; P. H. Seeberger; Pagel, K. *Nature* **2015,** *526*, 241.

(14) Khanal, N.; Masellis, C.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. *Analyst* **2018,** *143*, 1846.

(15) Seo, J.; Hoffmann, W.; Malerz, S.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G. *Int. J. Mass Spectrom.* **2017**, doi: 10.1016/j.ijms.2017.06.011.

(16) Hoffmann, W.; Folmert, K.; Moschner, J.; Huang, X.; von Berlepsch, H.; Koksch, B.; Bowers, M. T.; von Helden, G.; Pagel, K. *J. Am. Chem. Soc.* **2018,** *140*, 244.

(17) Seo, J.; Hoffmann, W.; Warnke, S.; Huang, X.; Gewinner, S.; Schöllkopf, W.; Bowers, M. T.; von Helden, G.; Pagel, K. *Nat. Chem.* **2017,** *9*, 39.

(18) Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010,** *39*, 1633.

(19) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2017,** *86*, 27.

(20) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006,** *75*, 333.

(21) Murray, J. C.; Lopez, D. A., *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Harvard University Press: Cambridge, 1996.

(22) Caflisch, A. *Curr. Opin. Chem. Biol.* **2006,** *10*, 437.

(23) M., D. C.; Andrej, Š.; Martin, K. *Angew. Chem. Int. Ed.* **1998,** *37*, 868.

(24) Dobson, C. M. *Nature* **2003,** *426*, 884.

(25) Dill, K. A.; Chan, H. S. *Nat. Struct. Mol. Biol.* **1997,** *4*, 10.

(26) Richardson, J. S. *Adv. Protein Chem.* **1981,** *34*, 167.

(27) Eisenberg, D.; Jucker, M. *Cell* **2012,** *148*, 1188.

(28) Stefani, M.; Dobson, C. M. *J. Mol. Med.* **2003,** *81*, 678.

(29) Paravastu, A. K.; Leapman, R. D.; Yau, W. M.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105*, 18349.

(30) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005,** *435*, 773.

(31) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riekel, C.; Eisenberg, D. *Nature* **2007,** *447*, 453.

(32) Xiao, Y.; Ma, B.; McElheny, D.; Parthasarathy, S.; Long, F.; Hoshi, M.; Nussinov, R.; Ishii, Y. *Nat. Struct. Mol. Biol.* **2015,** *22*, 499.

(33) Wälti, M. A.; Ravotti, F.; Arai, H.; Glabe, C. G.; Wall, J. S.; Böckmann, A.; Güntert, P.; Meier, B. H.; Riek, R. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113*, E4976.

(34) Astbury, W. T.; Dickinson, S.; Bailey, K. *Biochem. J.* **1935,** *29*, 2351.

(35) Eanes, E. D.; Glenner, G. G. *J. Histochem. Cytochem.* **1968,** *16*, 673.

(36) Sunde, M.; Blake, C., The Structure of Amyloid Fibrils by Electron Microscopy and X-Ray Diffraction. In *Adv. Protein Chem.*, Richards, F. M.; Eisenberg, D. S.; Kim, P. S., Eds. Academic Press: 1997; Vol. 50, pp 123.

(37) Yamada, N.; Ariga, K.; Naito, M.; Matsubara, K.; Koyama, E. *J. Am. Chem. Soc.* **1998,** *120*, 12192.

(38) Lee, S.-W.; Mou, Y.; Lin, S.-Y.; Chou, F.-C.; Tseng, W.-H.; Chen, C.-h.; Lu, C.-Y. D.; Yu, S. S. F.; Chan, J. C. C. *J. Mol. Biol.* **2008,** *378*, 1142.

(39) Carulla, N.; Caddy, G. L.; Hall, D. R.; Zurdo, J.; Gairi, M.; Feliz, M.; Giralt, E.; Robinson, C. V.; Dobson, C. M. *Nature* **2005,** *436*, 554.

(40) Sánchez, L.; Madurga, S.; Pukala, T.; Vilaseca, M.; López-Iglesias, C.; Robinson, C. V.; Giralt, E.; Carulla, N. *J. Am. Chem. Soc.* **2011,** *133*, 6505.

(41) Koffie, R. M.; Meyer-Luehmann, M.; Hashimoto, T.; Adams, K. W.; Mielke, M. L.; Garcia-Alloza, M.; Micheva, K. D.; Smith, S. J.; Kim, M. L.; Lee, V. M.; Hyman, B. T.; Spires-Jones, T. L. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106*, 4012.

(42) Woods, L. A.; Radford, S. E.; Ashcroft, A. E. *Biochim. Biophys. Acta* **2013,** *1834*, 1257.

(43) Wolfe, L. S.; Calabrese, M. F.; Nath, A.; Blaho, D. V.; Miranker, A. D.; Xiong, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107*, 16863.

(44) Kalapothakis, Jason M. D.; Morris, Ryan J.; Szavits-Nossan, J.; Eden, K.; Covill, S.; Tabor, S.; Gillam, J.; Barran, Perdita E.; Allen, Rosalind J.; MacPhee, Cait E. *Biophys. J.* **2015,** *108*, 2300.

(45) Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E. *Nat. Chem.* **2015,** *7*, 73.

(46) Leney, A. C.; Pashley, C. L.; Scarff, C. A.; Radford, S. E.; Ashcroft, A. E. *Mol. BioSyst.* **2014,** *10*, 412.

(47) Young, L. M.; Mahood, R. A.; Saunders, J. C.; Tu, L.-H.; Raleigh, D. P.; Radford, S. E.; Ashcroft, A. E. *Analyst* **2015,** *140*, 6990.

(48) Larson, M. E.; Lesne, S. E. *J. Neurochem.* **2012,** *120 Suppl 1*, 125.

(49) Laganowsky, A.; Liu, C.; Sawaya, M. R.; Whitelegge, J. P.; Park, J.; Zhao, M.; Pensalfini, A.; Soriaga, A. B.; Landau, M.; Teng, P. K.; Cascio, D.; Glabe, C.; Eisenberg, D. *Science* **2012,** *335*, 1228.

(50) Apostol, M. I.; Perry, K.; Surewicz, W. K. *J. Am. Chem. Soc.* **2013,** *135*, 10202.

(51) B., K. A.; Prabha, D.; Maggie, T.; Laura, M.; H., H. H. *J. Mass Spectrom.* **2008,** *43*, 1.

(52) Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J.-E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *Nat. Chem.* **2009,** *1*, 326.

(53) Wang, J. S.; Whitehead, S. N.; Yeung, K. K. *J. Am. Soc. Mass Spectrom.* **2018,** *29*, 786.

(54) Cernescu, M.; Stark, T.; Kalden, E.; Kurz, C.; Leuner, K.; Deller, T.; Göbel, M.; Eckert, G. P.; Brutschy, B. *Anal. Chem.* **2012,** *84*, 5276.

(55) Stark, T.; Lieblein, T.; Pohland, M.; Kalden, E.; Freund, P.; Zangl, R.; Grewal, R.; Heilemann, M.; Eckert, G. P.; Morgner, N.; Göbel, M. W. *Biochemistry* **2017,** *56*, 4840.

(56) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996,** *100*, 16082.

(57) Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996,** *261*, 86.

(58) Shvartsburg, A. A.; Hudgins, R. R.; Dugourd, P.; Jarrold, M. F. *J. Phys. Chem. A* **1997,** *101*, 1684.

(59) Warnke, S.; Seo, J.; Boschmans, J.; Sobott, F.; Scrivens, J. H.; Bleiholder, C.; Bowers, M. T.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2015,** *137*, 4236.

(60) Hoffmann, W.; Marianski, M.; Warnke, S.; Seo, J.; Baldauf, C.; von Helden, G.; Pagel, K. *Phys. Chem. Chem. Phys.* **2016,** *18*, 19950.

(61) Revercomb, H. E.; Mason, E. A. *Anal. Chem.* **1975,** *47*, 970.

(62) Mason, E. A.; McDaniel, E. W., *Transport Properties of Ions in Gases*. John Wiley & Sons: New York, United States, 1988.

(63) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. *Nat. Chem.* **2011,** *3*, 172.

(64) Smith, D.; Griffin, J. *Science* **1978,** *199*, 1214.

(65) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995,** *30*, 95.

(66) Barth, A. *Biochim. Biophys. Acta* **2007,** *1767*, 1073.

(67) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. *Int. J. Mass Spectrom.* **2006,** *254*, 1.

(68) K K Lehmann; G Scoles, a.; Pate, B. H. *Annu. Rev. Phys. Chem.* **1994,** *45*, 241.

(69) Schöllkopf, W.; Gewinner, S.; Junkes, H.; Paarmann, A.; von Helden, G.; Bluem, H.; Todd, A. M. M. In *The new IR and THz FEL facility at the Fritz Haber Institute in Berlin*, Proc. of SPIE: 2015; pp 95121L.

(70) Warnke, S.; von Helden, G.; Pagel, K. *Proteomics* **2015,** *15*, 2804.

(71) Chen, F.; Gulbakan, B.; Weidmann, S.; Fagerer, S. R.; Ibanez, A. J.; Zenobi, R. *Mass Spectrom. Rev.* **2016,** *35*, 48.

(72) McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. *J. Am. Chem. Soc.* **2015,** *137*, 12667.

(73) Papadopoulos, G.; Svendsen, A.; Boyarkin, O. V.; Rizzo, T. R. *Faraday Discuss.* **2011,** *150*, 243.

(74) Tim, M.; Valérie, G.; Helmut, G.; Modesto, O. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013,** *3*, 408.

(75) Siuzdak, G.; Bothner, B.; Yeager, M.; Brugidou, C.; Fauquet, C. M.; Hoey, K.; Chang, C. M. *Chem. Biol.* **1996,** *3*, 45.

(76) Joost, S.; J., R. R.; David, V.; E., J. J.; R., H. A. J. *Angew. Chem. Int. Ed.* **2013,** *52*, 4020.

(77) Loo, J. A.; Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y.; Kaufman, S. L.; Chernushevich, I. V. *J. Am. Soc. Mass Spectrom.* **2005,** *16*, 998.

(78) Sharon, M.; Witt, S.; Glasmacher, E.; Baumeister, W.; Robinson, C. V. *J. Biol. Chem.* **2007,** *282*, 18448.

(79) Barrera, N. P.; Di Bartolo, N.; Booth, P. J.; Robinson, C. V. *Science* **2008,** *321*, 243.

(80) Barrera, N. P.; Robinson, C. V. *Annu. Rev. Biochem.* **2011,** *80*, 247.

(81) Zhou, M.; Morgner, N.; Barrera, N. P.; Politis, A.; Isaacson, S. C.; Matak-Vinković, D.; Murata, T.; Bernal, R. A.; Stock, D.; Robinson, C. V. *Science* **2011,** *334*, 380.

(82) Pierson, N. A.; Chen, L.; Valentine, S. J.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2011,** *133*, 13810.

(83) Shi, L.; Holliday, A. E.; Shi, H.; Zhu, F.; Ewing, M. A.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2014,** *136*, 12702.

(84) Warnke, S.; von Helden, G.; Pagel, K. *J. Am. Chem. Soc.* **2013,** *135*, 1177.

(85) Goth, M.; Lermyte, F.; Schmitt, X. J.; Warnke, S.; von Helden, G.; Sobott, F.; Pagel, K. *Analyst* **2016,** *141*, 5502.

(86) Shelimov, K. B.; Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997,** *119*, 2240.

(87) González Flórez, A. I.; Mucha, E.; Ahn, D.-S.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; von Helden, G. *Angew. Chem. Int. Ed.* **2016,** *55*, 3295.

(88) Warnke, S.; Hoffmann, W.; Seo, J.; De Genst, E.; von Helden, G.; Pagel, K. *J. Am. Soc. Mass Spectrom.* **2016**.

(89) Oomens, J.; Polfer, N.; Moore, D. T.; van der Meer, L.; Marshall, A. G.; Eyler, J. R.; Meijer, G.; von Helden, G. *Phys. Chem. Chem. Phys.* **2005,** *7*, 1345.

(90) Jarrold, M. F. *Phys. Chem. Chem. Phys.* **2007,** *9*, 1659.

(91) Kohtani, M.; Jones, T. C.; Schneider, J. E.; Jarrold, M. F. *J. Am. Chem. Soc.* **2004,** *126*, 7420.

(92) Pithadia, A.; Brender, J. R.; Fierke, C. A.; Ramamoorthy, A. *J. Diabetes Res.* **2016,** *2016*, 2046327.

(93) Butler, A. E.; Jang, J.; Gurlo, T.; Carty, M. D.; Soeller, W. C.; Butler, P. C. *Diabetes* **2004,** *53*, 1509.

(94) Dupuis, N. F.; Wu, C.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2009,** *131*, 18283.

(95) Abedini, A.; Raleigh, D. P. *Biochemistry* **2005,** *44*, 16284.

(96) Dupuis, N. F.; Wu, C.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2011,** *133*, 7240.

(97) Haass, C.; Selkoe, D. J. *Nat. Rev. Mol. Cell Biol.* **2007,** *8*, 101.

(98) Selkoe, D. J.; Hardy, J. *EMBO Mol. Med.* **2016,** *8*, 595.

(99) Lesné, S.; Koh, M. T.; Kotilinek, L.; Kayed, R.; Glabe, C. G.; Yang, A.; Gallagher, M.; Ashe, K. H. *Nature* **2006,** *440*, 352.

(100) Bitan, G.; Kirkitadze, M. D.; Lomakin, A.; Vollers, S. S.; Benedek, G. B.; Teplow, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100*, 330.

(101) Economou, N. J.; Giammona, M. J.; Do, T. D.; Zheng, X.; Teplow, D. B.; Buratto, S. K.; Bowers, M. T. *J. Am. Chem. Soc.* **2016,** *138*, 1772.

(102) Serra-Vidal, B.; Pujadas, L.; Rossi, D.; Soriano, E.; Madurga, S.; Carulla, N. *ACS Chem. Biol.* **2014,** *9*, 2678.

(103) Li, K. S.; Rempel, D. L.; Gross, M. L. *J. Am. Chem. Soc.* **2016,** *138*, 12090.

(104) Kłoniecki, M.; Jabłonowska, A.; Poznański, J.; Langridge, J.; Hughes, C.; Campuzano, I.; Giles, K.; Dadlez, M. *J. Mol. Biol.* **2011,** *407*, 110.

(105) Pujol-Pina, R.; Vilaprinyó-Pascual, S.; Mazzucato, R.; Arcella, A.; Vilaseca, M.; Orozco, M.; Carulla, N. *Sci. Rep.* **2015,** *5*, 14809.

(106) Hochberg, G. K. A.; Ecroyd, H.; Liu, C.; Cox, D.; Cascio, D.; Sawaya, M. R.; Collier, M. P.; Stroud, J.; Carver, J. A.; Baldwin, A. J.; Robinson, C. V.; Eisenberg, D. S.; Benesch, J. L. P.; Laganowsky, A. *Proc. Natl. Acad. Sci. U.S.A.* **2014,** *111*, E1562.

(107) Sinha, S.; Lopes, D. H. J.; Du, Z.; Pang, E. S.; Shanmugam, A.; Lomakin, A.; Talbiersky, P.; Tennstaedt, A.; McDaniel, K.; Bakshi, R.; Kuo, P.-Y.; Ehrmann, M.; Benedek, G. B.; Loo, J. A.; Klärner, F.-G.; Schrader, T.; Wang, C.; Bitan, G. *J. Am. Chem. Soc.* **2011,** *133*, 16958.

(108) Zheng, X.; Liu, D.; Klärner, F.-G.; Schrader, T.; Bitan, G.; Bowers, M. T. *J. Phys. Chem. B* **2015,** *119*, 4831.

(109) Riba, I.; Barran, P. E.; Cooper, G. J. S.; Unwin, R. D. *Int. J. Mass Spectrom.* **2015,** *391*, 47.

(110) Li, H.; Ha, E.; Donaldson, R. P.; Jeremic, A. M.; Vertes, A. *Anal. Chem.* **2015,** *87*, 9829.

(111) Lee, S. J. C.; Choi, T. S.; Lee, J. W.; Lee, H. J.; Mun, D.-G.; Akashi, S.; Lee, S.-W.; Lim, M. H.; Kim, H. I. *Chem. Sci.* **2016,** *7*, 5398.

(112) Susa, A. C.; Wu, C.; Bernstein, S. L.; Dupuis, N. F.; Wang, H.; Raleigh, D. P.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2014,** *136*, 12912.

(113) G. D. Rose; Wolfenden, R. *Annu. Rev. Biophys. Biomol. Struct.* **1993,** *22*, 381.

(114) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins: Struct., Funct., Bioinf.* **2000,** *41*, 415.

(115) Yutani, K.; Ogasahara, K.; Tsujita, T.; Sugino, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1987,** *84*, 4441.

(116) Tripet, B.; Wagschal, K.; Lavigne, P.; Mant, C. T.; Hodges, R. S. *J. Mol. Biol.* **2000,** *300*, 377.

(117) Chiti, F.; Taddei, N.; Baroni, F.; Capanni, C.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Nat. Struct. Mol. Biol.* **2002,** *9*, 137.

(118) Gerling, U. I. M.; Salwiczek, M.; Cadicamo, C. D.; Erdbrink, H.; Czekelius, C.; Grage, S. L.; Wadhwani, P.; Ulrich, A. S.; Behrends, M.; Haufe, G.; Koksch, B. *Chem. Sci.* **2014,** *5*, 819.

(119) Young, L.; Jernigan, R. L.; Covell, D. G. *Protein Sci.* **1994,** *3*, 717.

(120) Efremov, R. G.; Chugunov, A. O.; Pyrkov, T. V.; Priestle, J. P.; Arseniev, A. S.; Jacoby, E. *Curr. Med. Chem.* **2007,** *14*, 393.

(121) Palliser, C. C.; Parry, D. A. *Proteins* **2001,** *42*, 243.

(122) Biswas, K. M.; DeVido, D. R.; Dorsey, J. G. *J. Chrom. A* **2003,** *1000*, 637.

(123) J. -L. Fauchère; Pliska, V. *Eur. J. Med. Chem.* **1983,** *18*, 369.

(124) Nozaki, Y.; Tanford, C. *J. Biol. Chem.* **1971,** *246*, 2211.

(125) Rose, G.; Geselowitz, A.; Lesser, G.; Lee, R.; Zehfus, M. *Science* **1985,** *229*, 834.

(126) Bull, H. B.; Breese, K. *Arch. Biochem. Biophys.* **1974,** *161*, 665.

(127) Kovacs, J. M.; Mant, C. T.; Hodges, R. S. *Peptide Science* **2006,** *84*, 283.

(128) Mant, C. T.; Kovacs, J. M.; Kim, H. M.; Pollock, D. D.; Hodges, R. S. *Biopolymers* **2009,** *92*, 573.

(129) Meek, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1980,** *77*, 1632.

(130) Tripet, B.; Cepeniene, D.; Kovacs, J. M.; Mant, C. T.; Krokhin, O. V.; Hodges, R. S. *J. Chrom. A* **2007,** *1141*, 212.

(131) Wilce, M. C. J.; Aguilar, M.-I.; Hearn, M. T. W. *Anal. Chem.* **1995,** *67*, 1210.

(132) Wilce, M. C. J.; Aguilar, M. I.; Hearn, M. T. W. *J. Chrom. A* **1991,** *536*, 165.

(133) Li, L.; Li, C.; Zhang, Z.; Alexov, E. *J. Chem. Theory Comput.* **2013,** *9*, 2126.

(134) Meng, C. K.; Fenn, J. B. *Org. Mass Spectrom.* **1991,** *26*, 542.

(135) Nemes, P.; Schlosser, G.; Vékey, K. *J. Mass Spectrom.* **2005,** *40*, 43.

(136) Do, T. D.; de Almeida, N. E. C.; LaPointe, N. E.; Chamas, A.; Feinstein, S. C.; Bowers, M. T. *Anal. Chem.* **2016,** *88*, 868.

(137) Vulpetti, A.; Dalvit, C. *Drug Discov. Today* **2012,** *17*, 890.

(138) Gottler, L. M.; de la Salud Bea, R.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *Biochemistry* **2008,** *47*, 9243.

(139) Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I.; Ye, S.; Koksch, B. *Chem. Soc. Rev.* **2012,** *41*, 2135.

(140) Berger, A. A.; Voller, J. S.; Budisa, N.; Koksch, B. *Acc. Chem. Res.* **2017,** *50*, 2093.

(141) Robalo, J. R.; Huhmann, S.; Koksch, B.; Vila Verde, A. *Chem 3*, 881.

(142) Shoemaker, K. R.; Kim, P. S.; York, E. J.; Stewart, J. M.; Baldwin, R. L. *Nature* **1987,** *326*, 563.

(143) Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009,** *78*, 929.

(144) Chakrabartty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci.* **1994,** *3*, 843.

(145) Marqusee, S.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987,** *84*, 8898.

(146) Counterman, A. E.; Clemmer, D. E. *J. Phys. Chem. B* **2003,** *107*, 2111.

(147) Counterman, A. E.; Clemmer, D. E. *J. Phys. Chem. B* **2002,** *106*, 12045.

(148) Counterman, A. E.; Clemmer, D. E. *J. Am. Chem. Soc.* **2001,** *123*, 1490.

(149) Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998,** *120*, 12974.

(150) Kohtani, M.; Jarrold, M. F. *J. Am. Chem. Soc.* **2004,** *126*, 8454.

(151) Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **1999,** *121*, 3494.

(152) Schubert, F.; Rossi, M.; Baldauf, C.; Pagel, K.; Warnke, S.; von Helden, G.; Filsinger, F.; Kupser, P.; Meijer, G.; Salwiczek, M.; Koksch, B.; Scheffler, M.; Blum, V. *Phys. Chem. Chem. Phys.* **2015,** *17*, 7373.

(153) Rossi, M.; Blum, V.; Kupser, P.; von Helden, G.; Bierau, F.; Pagel, K.; Meijer, G.; Scheffler, M. *J. Phys. Chem. Lett.* **2010,** *1*, 3465.

(154) Ko, J. Y.; Heo, S. W.; Lee, J. H.; Oh, H. B.; Kim, H.; Kim, H. I. *J. Phys. Chem. A* **2011,** *115*, 14215.

(155) Wieczorek, R.; Dannenberg, J. J. *J. Phys. Chem. B* **2008,** *112*, 1320.

(156) Scheike, J. A.; Baldauf, C.; Spengler, J.; Albericio, F.; Pisabarro, M. T.; Koksch, B. *Angew. Chem. Int. Ed.* **2007,** *46*, 7766.

(157) Tkatchenko, A.; Rossi, M.; Blum, V.; Ireta, J.; Scheffler, M. *Phys. Rev. Lett.* **2011,** *106*, 118102.

(158) Fu, Y.; Gao, J.; Bieschke, J.; Dendle, M. A.; Kelly, J. W. *J. Am. Chem. Soc.* **2006,** *128*, 15948.

(159) Deechongkit, S.; Dawson, P. E.; Kelly, J. W. *J. Am. Chem. Soc.* **2004,** *126*, 16762.

(160) Wang, M.; Wales, T. E.; Fitzgerald, M. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103*, 2600.

(161) Lide, D. R., *Handbook of Chemistry and Physics*. 71st ed.; CRC Press: Boca Raton, 1990.

(162) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008,** *4*, 435.

(163) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001,** *105*, 6474.

(164) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996,** *118*, 11225.

(165) McAliley, J. H.; Bruce, D. A. *J. Chem. Theory Comput.* **2011,** *7*, 3756.

(166) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. *J. Comput. Chem.* **2001,** *22*, 1340.

(167) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007,** *126*, 014101.

(168) Becke, A. D. *J. Chem. Phys.* **1993,** *98*, 5648.

(169) Frisch, M. J. *Gaussian 09*, Rev D.01; Gaussian, Inc.: Wallingford, CT, USA, 2009.

(170) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2005,** *127*, 14534.

(171) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990,** *11*, 431.

(172) Lin, Y.-S.; Bowman, Gregory R.; Beauchamp, Kyle A.; Pande, Vijay S. *Biophys. J.* **2012,** *102*, 315.

(173) Karle, I. L.; Das, C.; Balaram, P. *Biopolymers* **2001,** *59*, 276.

(174) Ohyama, T.; Oku, H.; Hiroki, A.; Maekawa, Y.; Yoshida, M.; Katakai, R. *Biopolymers* **2000,** *54*, 375.

(175) Chapman, E.; Thorson, J. S.; Schultz, P. G. *J. Am. Chem. Soc.* **1997,** *119*, 7151.

(176) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997,** *36*, 11314.

(177) Pandey, A.; Mann, M. *Nature* **2000,** *405*, 837.

(178) Ruotolo, B. T.; Robinson, C. V. *Curr. Opin. Chem. Biol.* **2006,** *10*, 402.

(179) Heck, A. J. R. *Nat. Methods* **2008,** *5*, 927.

(180) Engen, J. R. *Anal. Chem.* **2009,** *81*, 7870.

(181) Oh, H. B.; Moon, B. *Mass Spectrom. Rev.* **2015,** *34*, 116.

(182) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. *Nat. Chem.* **2014,** *6*, 281.

(183) Bohrer, B. C.; Merenbloom, S. I.; Koeniger, S. L.; Hilderbrand, A. E.; Clemmer, D. E. *Annu. Rev. Anal. Chem.* **2008,** *1*, 293.

(184) Wyttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2007,** *58*, 511.

(185) Polfer, N. C. *Chem. Soc. Rev.* **2011,** *40*, 2211.

(186) Polfer, N. C.; Oomens, J. *Mass Spectrom. Rev.* **2009,** *28*, 468.

(187) Simons, J. P. *Mol. Phys.* **2009,** *107*, 2435.

(188) Wyttenbach, T.; Bowers, M. T. *J. Phys. Chem. B* **2011,** *115*, 12266.

(189) Segev, E.; Wyttenbach, T.; Bowers, M. T.; Gerber, R. B. *Phys. Chem. Chem. Phys.* **2008,** *10*, 3077.

(190) Pagel, K.; Kupser, P.; Bierau, F.; Polfer, N. C.; Steill, J. D.; Oomens, J.; Meijer, G.; Koksch, B.; von Helden, G. *Int. J. Mass Spectrom.* **2009,** *283*, 161.

(191) Fung, Y. M.; Besson, T.; Lemaire, J.; Maitre, P.; Zubarev, R. A. *Angew. Chem. Int. Ed.* **2009,** *48*, 8340.

(192) Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99*, 15863.

(193) Liu, L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2011,** *22*, 310.

(194) Liu, L.; Bagal, D.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2009,** *131*, 15980.

(195) Gross, D. S.; Zhao, Y.; William, E. R. *J. Am. Soc. Mass Spectrom.* **1997,** *8*, 519.

(196) Schenk, E. R.; Almeida, R.; Miksovska, J.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. *J. Am. Soc. Mass Spectrom.* **2015,** *26*, 555.

(197) Shelimov, K. B.; Jarrold, M. F. *J. Am. Chem. Soc.* **1997,** *119*, 2987.

(198) Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta* **2005,** *1751*, 119.

(199) Casal, H. L.; Köhler, U.; Mantsch, H. H. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **1988,** *957*, 11.

(200) Meersman, F.; Smeller, L.; Heremans, K. *Biophys. J.* **2002,** *82*, 2635.

(201) Caughey, B.; Peter T. Lansbury, J. *Annu. Rev. Neurosci.* **2003,** *26*, 267.

(202) Selkoe, D. J. *Nature* **2003,** *426*, 900.

(203) Walsh, D. M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W. K.; Anwyl, R.; Wolfe, M. S.; Rowan, M. J.; Selkoe, D. J. *Nature* **2002,** *416*, 535.

(204) Goedert, M.; Spillantini, M. G. *Science* **2006,** *314*, 777.

(205) Winner, B.; Jappelli, R.; Maji, S. K.; Desplats, P. A.; Boyer, L.; Aigner, S.; Hetzer, C.; Loher, T.; Vilar, M.; Campioni, S.; Tzitzilonis, C.; Soragni, A.; Jessberger, S.; Mira, H.; Consiglio, A.; Pham, E.; Masliah, E.; Gage, F. H.; Riek, R. *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108*, 4194.

(206) Lin, C.-Y.; Gurlo, T.; Kayed, R.; Butler, A. E.; Haataja, L.; Glabe, C. G.; Butler, P. C. *Diabetes* **2007,** *56*, 1324.

(207) Gurlo, T.; Ryazantsev, S.; Huang, C.-j.; Yeh, M. W.; Reber, H. A.; Hines, O. J.; O'Brien, T. D.; Glabe, C. G.; Butler, P. C. *Am. J. Pathol.* **2010,** *176*, 861.

(208) Ivanova, M. I.; Sievers, S. A.; Sawaya, M. R.; Wall, J. S.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106*, 18990.

(209) Stromer, T.; Serpell, L. C. *Microsc. Res. Tech.* **2005,** *67*, 210.

(210) Matthes, D.; Daebel, V.; Meyenberg, K.; Riedel, D.; Heim, G.; Diederichsen, U.; Lange, A.; de Groot, B. L. *J. Mol. Biol.* **2014,** *426*, 362.

(211) Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclaye, S.; Dufrene, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2009,** *421*, 415.

(212) Celej, M. S.; Sarroukh, R.; Goormaghtigh, E.; Fidelio, G. D.; Ruysschaert, J. M.; Raussens, V. *Biochem. J.* **2012,** *443*, 719.

(213) Buchanan, L. E.; Dunkelberger, E. B.; Tran, H. Q.; Cheng, P.-N.; Chiu, C.-C.; Cao, P.; Raleigh, D. P.; de Pablo, J. J.; Nowick, J. S.; Zanni, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110*, 19285.

(214) Bleiholder, C.; Do, T. D.; Wu, C.; Economou, N. J.; Bernstein, S. S.; Buratto, S. K.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2013,** *135*, 16926.

(215) Do, T. D.; LaPointe, N. E.; Economou, N. J.; Buratto, S. K.; Feinstein, S. C.; Shea, J.-E.; Bowers, M. T. *J. Phys. Chem. B* **2013,** *117*, 10759.

(216) Do, T. D.; Economou, N. J.; Chamas, A.; Buratto, S. K.; Shea, J.-E.; Bowers, M. T. *J. Phys. Chem. B* **2014,** *118*, 11220.

(217) Do, T. D.; LaPointe, N. E.; Sangwan, S.; Teplow, D. B.; Feinstein, S. C.; Sawaya, M. R.; Eisenberg, D. S.; Bowers, M. T. *J. Phys. Chem. B* **2014,** *118*, 7247.

(218) Atherton, E.; C Sheppard, R., *Solid Phase Peptide Synthesis : A Practical Approach*. Oxford Univ. Press: 1989.

(219) Knowles, T. P.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. *Science* **2009,** *326*, 1533.

(220) Cai, S.; Singh, B. R. *Biochemistry* **2004,** *43*, 2541.

(221) Bleiholder, C.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass Spectrom.* **2011,** *308*, 1.

(222) Wyttenbach, T.; Bleiholder, C.; Bowers, M. T. *Anal. Chem.* **2013,** *85*, 2191.

(223) Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. *Structure* **2015,** *23*, 791.

(224) Cao, P.; Abedini, A.; Raleigh, D. P. *Curr. Opin. Struct. Biol.* **2013,** *23*, 82.

(225) Clark, A.; Cooper, G. J.; Lewis, C. E.; Morris, J. F.; Willis, A. C.; Reid, K. B.; Turner, R. C. *Lancet* **1987,** *2*, 231.

(226) Luskey, K. L. *Diabetes care* **1992,** *15*, 297.

(227) Kahn, S. E.; Andrikopoulos, S.; Verchere, C. B. *Diabetes* **1999,** *48*, 241.

(228) Cao, P.; Marek, P.; Noor, H.; Patsalo, V.; Tu, L. H.; Wang, H.; Abedini, A.; Raleigh, D. P. *FEBS Lett.* **2013,** *587*, 1106.

(229) Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C. *Endocr. Rev.* **2008,** *29*, 303.

(230) Lorenzo, A.; Razzaboni, B.; Weir, G. C.; Yankner, B. A. *Nature* **1994,** *368*, 756.

(231) Zraika, S.; Hull, R. L.; Verchere, C. B.; Clark, A.; Potter, K. J.; Fraser, P. E.; Raleigh, D. P.; Kahn, S. E. *Diabetologia* **2010,** *53*, 1046.

(232) Cooper, G. J.; Willis, A. C.; Clark, A.; Turner, R. C.; Sim, R. B.; Reid, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **1987,** *84*, 8628.

(233) Gilead, S.; Gazit, E. *Exp. Diabetes Res.* **2008,** *2008*, 256954.

(234) Scrocchi, L. A.; Ha, K.; Chen, Y.; Wu, L.; Wang, F.; Fraser, P. E. *J. Struct. Biol.* **2003,** *141*, 218.

(235) Ehud, G. *Curr. Med. Chem.* **2002,** *9*, 1725.

(236) Nilsson, M. R.; Raleigh, D. P. *J. Mol. Biol.* **1999,** *294*, 1375.

(237) Ilitchev, A. I.; Giammona, M. J.; Do, T. D.; Wong, A. G.; Buratto, S. K.; Shea, J.-E.; Raleigh, D. P.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **2016,** *27*, 1010.

(238) Kapurniotu, A.; Schmauder, A.; Tenidis, K. *J. Mol. Biol.* **2002,** *315*, 339.

(239) Tatarek-Nossol, M.; Yan, L. M.; Schmauder, A.; Tenidis, K.; Westermark, G.; Kapurniotu, A. *Chem. Biol.* **2005,** *12*, 797.

(240) Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M. L.; Voelter, W.; Brunner, H.; Kapurniotu, A. *J. Mol. Biol.* **2000,** *295*, 1055.

(241) Shim, S. H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106*, 6614.

(242) Soriaga, A. B.; Sangwan, S.; Macdonald, R.; Sawaya, M. R.; Eisenberg, D. *J. Phys. Chem. B* **2016,** *120*, 5810.

(243) Deng, W.; Cao, A.; Lai, L. *Protein Sci.* **2008,** *17*, 1102.

(244) Azriel, R.; Gazit, E. *J. Biol. Chem.* **2001,** *276*, 34156.

(245) Cai, Z.; Li, J.; Yin, C.; Yang, Z.; Wu, J.; Zhou, R. *J. Phys. Chem. B* **2014,** *118*, 48.

(246) Melquiond, A.; Gelly, J.-C.; Mousseau, N.; Derreumaux, P. *J. Chem. Phys.* **2007,** *126*, 065101.

(247) Wu, C.; Lei, H.; Duan, Y. *Biophys. J.* **2005,** *88*, 2897.

(248) Wu, C.; Lei, H.; Duan, Y. *J. Am. Chem. Soc.* **2005,** *127*, 13530.

(249) Wu, C.; Lei, H.; Duan, Y. *Biophys. J.* **2004,** *87*, 3000.

(250) Zanuy, D.; Ma, B.; Nussinov, R. *Biophys. J.* **2003,** *84*, 1884.

(251) Kemper, P. R.; Dupuis, N. F.; Bowers, M. T. *Int. J. Mass Spectrom.* **2009,** *287*, 46.

(252) Kayed, R.; Bernhagen, J.; Greenfield, N.; Sweimeh, K.; Brunner, H.; Voelter, W.; Kapurniotu, A. *J. Mol. Biol.* **1999,** *287*, 781.

(253) Wang, L.; Ilitchev, A. I.; Giammona, M. J.; Li, F.; Buratto, S. K.; Bowers, M. T. *J. Phys. Chem. B* **2016,** *120*, 11905.

(254) Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G. *Angew. Chem. Int. Ed.* **2016,** *55*, 14173.

(255) Arrondo, J. L. R.; Goñi, F. M. *Prog. Biophys. Mol. Biol.* **1999,** *72*, 367.

(256) Bramanti, E.; Benedetti, E. *Biopolymers* **1996,** *38*, 639.

(257) Do, T. D.; LaPointe, N. E.; Nelson, R.; Krotee, P.; Hayden, E. Y.; Ulrich, B.; Quan, S.; Feinstein, S. C.; Teplow, D. B.; Eisenberg, D.; Shea, J.-E.; Bowers, M. T. *J. Am. Chem. Soc.* **2016,** *138*, 549.

(258) Pawar, A. P.; Dubay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. *J. Mol. Biol.* **2005,** *350*, 379.

(259) Tolmachev, A. V.; Webb, I. K.; Ibrahim, Y. M.; Garimella, S. V. B.; Zhang, X.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2014,** *86*, 9162.

(260) Webb, I. K.; Garimella, S. V. B.; Tolmachev, A. V.; Chen, T.-C.; Zhang, X.; Norheim, R. V.; Prost, S. A.; LaMarche, B.; Anderson, G. A.; Ibrahim, Y. M.; Smith, R. D. *Anal. Chem.* **2014,** *86*, 9169.

(261) Young, L. M.; Saunders, J. C.; Mahood, R. A.; Revill, C. H.; Foster, R. J.; Ashcroft, A. E.; Radford, S. E. *Methods* **2016,** *95*, 62.

(262) Daly, S.; Kulesza, A.; Poussigue, F.; Simon, A.-L.; Choi, C. M.; Knight, G.; Chirot, F.; MacAleese, L.; Antoine, R.; Dugourd, P. *Chem. Sci.* **2016,** *7*, 1609.

Appendix A

Novel Hydrophobicity Scale for Amino Acids

Table A.1: $^{DT}CCS_{He}$ of protonated valine meta clusters

n	$\tilde{z}$	$^{DT}CCS_{He}$ / Å^2
1	1	59
2	1	100
4	1	136
6	1	183
8	1	225
8	2	244
9	2	273
10	2	283
11	2	296
12	2	326
13	2	343
14	2	367
16	2	412
21	3	495
22	3	524
24	3	567
26	4	615
32	4	744
34	4	777
36	4	810

Table A.2: $^{DT}CCS_{He}$ of isoleucine meta clusters in positive and negative ion mode

n	*z*	$^{DT}CCS_{He}$ (+) / Å^2	$^{DT}CCS_{He}$ (-) / Å^2
1	1	66	63
2	1	108	103
2	1		105
8	2	273	
9	2	302	
10	2	315	
11	2	329	330
12	2	351	348
13	2	377	371
14	2	401	388
15	2	418	411
16	2	445	431
17	2		451
18	2		473
19	2		496
20	3	528	
21	3	543	
22	3	572	
24	3	614	
26	4	661	
27	3		645
30	4	746	

Table A.3: $^{DT}CCS_{He}$ of phenylalanine meta clusters in positive and negative ion mode

n	*z*	$^{DT}CCS_{He}$ (+) / Å²	$^{DT}CCS_{He}$ (-) / Å²
1	1	73	72
2	1	111	116
2	1		120
8	2	304	
9	2	329	
10	2	348	
11	2	379	353
12	2	399	373
13	2	414	395
14	2	436	416
15	2	456	429
16	2		453
17	2		465
17	3	500	
18	2		480
18	3	519	
19	2		501
19	3	538	
20	2		519
20	3	558	
21	2		539
21	3	580	
22	2		559
22	3	590	
23	3	607	

Table A.4: $^{DT}CCS_{He}$ of protonated tryptophan meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	82
2	1	121
2	1	123
3	1	166
4	1	197
5	1	227
7	1	280
7	2	284
8	2	317
9	2	343
10	2	373
11	2	393
12	2	417
13	2	438
14	2	444
15	3	488
18	3	554
20	4	632
21	3	583
26	4	705
28	4	694

Table A.5: $^{DT}CCS_{He}$ of protonated leucine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	65
2	1	109
2	1	111
8	2	279
10	2	322
11	2	340
12	2	361
13	2	379
14	2	402
15	2	422
16	2	447
17	2	452
18	2	472
19	3	495
20	2	502
21	3	532
24	3	578
27	3	627
30	3	678
36	4	785
40	4	849

Table A.6: $^{DT}CCS_{He}$ of protonated alanine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	52
2	1	80
4	1	127
11	2	253
12	2	259
14	2	302
15	2	338
17	3	354
22	3	422
24	3	455
26	3	487
28	3	518

Table A.7: $^{DT}CCS_{He}$ of protonated proline meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	59
2	1	92
2	1	94
3	1	130
4	1	142
5	1	169
6	1	186
9	2	248
10	2	270
11	2	270
11	2	282
12	2	291
13	2	266
14	2	317
15	2	324
16	2	339
18	2	365
18	2	377
20	2	385
20	2	397
21	3	411
22	3	420
24	3	447
27	3	467
27	3	474
29	3	501
30	3	511
40	4	650

Table A.8: $^{DT}CCS_{He}$ of protonated methionine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	68
2	1	111
7	1	240
8	2	288
9	2	309
10	2	330
11	2	350
12	2	374
13	2	389
14	2	405
15	2	430
16	2	445
17	2	465
18	2	480
18	3	494
19	3	511
20	3	520
21	3	530
22	3	557
22	4	588
23	3	572
24	3	589
26	3	632
26	4	668
27	3	646
30	4	721
32	4	749
34	4	776
36	4	784

Table A.9: $^{DT}CCS_{He}$ of protonated lysine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	65
2	1	104
3	1	134
4	1	160
5	1	191
6	1	211
7	2	241
8	2	265
9	2	285
10	2	309
11	2	328
12	2	344
13	2	358
14	2	378
14	3	393
14	3	400
15	3	410
16	3	419
17	3	432
18	3	445
19	3	458
20	3	477
21	3	491
28	4	598

Table A.10: $^{DT}CCS_{He}$ of arginine meta clusters in positive and negative ion mode

n	*z*	$^{DT}CCS_{He}$ (+) / Å^2	$^{DT}CCS_{He}$ (-) / Å^2
1	1	72	72
2	1	115	111
3	1	145	143
4	1	174	175
5	1	198	200
6	1	222	
7	1	247	
7	2	255	
8	2	276	
9	2	291	
10	2	310	
11	2	330	
12	2	352	
12	3	365	
13	3	374	
13	3	381	
14	2	388	
14	3	392	
15	3	408	
16	3	424	
17	3	449	
18	3	461	
21	3	507	
21	4	520	
22	4	533	
23	4	545	
24	4	559	
28	4	619	

Table A.11: $^{DT}CCS_{He}$ of protonated histidine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	67
2	1	109
3	1	142
4	1	163
5	1	198
6	1	222
7	2	264
8	2	294
9	2	290
9	2	298
10	2	314
11	2	327
12	2	350
14	3	429
15	3	446
17	3	454
18	3	468
20	4	525

Table A.12: $^{DT}CCS_{He}$ of protonated glutamic acid meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	63
2	1	105
3	1	135
4	1	159
5	1	176
6	1	195
7	1	211
8	1	233
8	2	250
9	1	251
9	2	266
10	2	284
11	2	300
12	2	311
13	2	308
14	2	328
15	2	342
16	2	356
18	2	380
18	3	413
19	3	416
20	3	423
21	3	437
22	3	447
23	3	456
24	3	467
27	3	497

Table A.13: $^{DT}CCS_{He}$ of protonated aspartic acid meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	58
2	1	93
6	1	179
7	1	196
10	2	254
11	2	268
12	2	285
13	2	288
14	2	297
15	2	315
16	2	330
17	2	342
18	2	357
18	3	369
19	3	383
20	3	392
21	3	396
22	3	409
23	3	420
24	3	433
25	3	436
26	3	450
27	3	458
28	3	463
36	4	557

Table A.14: $^{DT}CCS_{He}$ of protonated tyrosine meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	75
2	1	116
5	1	221
6	1	245
8	2	300
9	2	325
10	2	348
11	2	372
12	2	391
13	2	413
13	2	421
18	3	515

Table A.15: $^{DT}CCS_{He}$ of serine meta clusters in positive and negative ion mode

n	*z*	$^{DT}CCS_{He}$ (+) / Å^2	$^{DT}CCS_{He}$ (-) / Å^2
1	1	51	52
1	1	52	
2	1	82	81
2	1	83	
6	1		155
8	1	190	
9	2	219	
10	2		234
11	2	238	236
12	2	252	239
13	2	260	255
14	2	272	270
14	2	275	
15	2	277	282
16	2	290	292
16	2		297
17	2	300	303
18	2	311	313
19	2	323	321
20	2	330	
21	3	355	
22	3	365	
23	3	369	
24	3	383	
25	3	389	
26	3	398	
27	3	408	410
28	3	416	
29	3	424	
31	3	442	

Table A.16: $^{DT}CCS_{He}$ of protonated threonine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	58
2	1	91
2	1	92
8	2	230
9	2	252
10	2	262
11	2	275
12	2	291
13	2	303
14	2	321
15	2	334
16	2	349
17	2	358
18	2	374
19	2	382
20	3	408
21	3	420
22	3	428
23	3	445
24	3	448
25	3	470
26	3	476
27	3	492

Table A.17: $^{DT}CCS_{He}$ of glutamine meta clusters in positive and negative ion mode

n	*z*	$^{DT}CCS_{He}$ (+) / Å²	$^{DT}CCS_{He}$ (-) / Å²
1	1	64	62
2	1	106	102
3	1	140	
4	1	170	138
5	1		170
6	1		190
7	1		210
7	2	251	
8	1		226
8	2	276	242
9	2	291	256
10	2	307	274
11	2	323	288
12	2	337	301
13	2	350	321
13	3	382	
14	2	363	330
14	3	392	
15	2	410	344
16	2		352
16	3	417	
17	2		370
17	3	427	
18	3	442	399
19	3	453	
20	3	462	
21	3	477	436
22	4	552	
24	3		468
26	4	606	
28	4	582	

Table A.18: $^{DT}CCS_{He}$ of protonated asparagine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	59
2	1	94
3	1	123
4	1	150
5	1	174
6	1	190
7	1	208
8	2	242
9	2	258
10	2	272
11	2	291
12	2	300
13	2	314
14	2	326
15	3	357
16	2	350
16	3	369
17	3	379
18	3	388
19	3	400
20	3	414
21	3	424
22	3	429
23	3	442
24	3	458
25	3	469
28	4	511
29	4	523
32	4	551

Table A.19: $^{DT}CCS_{He}$ of protonated cysteine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	58
2	1	91
10	2	252
11	2	273
12	2	300

Table A.20: $^{DT}CCS_{He}$ of protonated glycine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	44
2	1	70

Table A.21: $^{DT}CCS_{He}$ of protonated 1oF-phenylalanine meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å²
1	1	74
2	1	116
8	2	293
9	2	333
10	2	352
11	2	380
12	2	396
13	2	414
14	2	435
15	3	473
16	3	494
17	3	506
18	3	525
19	3	538

Table A.22: $^{DT}CCS_{He}$ of protonated 1mF-phenylalanine meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å²
1	1	74
2	1	114
3	1	154
5	1	228
8	2	301
9	2	327
10	2	347
11	2	383
12	2	402
13	2	414
14	2	435
15	2	451
16	2	471
17	2	496
18	3	522
19	3	537
20	3	557
21	3	574
22	3	598
24	3	632

Table A.23: $^{DT}CCS_{He}$ of protonated 1pF-phenylalanine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	74
2	1	114
6	1	248
8	2	301
9	2	330
10	2	352
11	2	385
12	2	401
13	2	420
14	2	439
15	2	459
16	2	479
17	2	493
17	2	478
18	2	510
19	2	533
20	2	548
22	3	604
23	3	626
24	3	647
25	3	660
26	3	680
27	3	701
32	4	827
34	4	847
36	4	880

Table A.24: $^{DT}CCS_{He}$ of protonated 5F-phenylalanine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	81
2	1	129
3	1	168
9	2	335
10	2	357
11	2	385
12	2	407
13	2	431
14	2	455

Table A.25: $^{DT}CCS_{He}$ of protonated P-phenylalanine meta clusters

n	z	$^{DT}CCS_{He}$ / Å^2
1	1	95
2	1	145
3	1	186
4	1	220

Table A.26: $^{DT}CCS_{He}$ of protonated PF5-phenylalanine meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	95
2	1	141
3	1	182
4	2	238
5	2	265

Table A.27: $^{DT}CCS_{He}$ of protonated F3-isoleucine meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	68
2	1	112
3	1	151
9	2	311
11	2	351
12	2	377
13	2	404
14	2	427
15	2	441
16	2	472
18	2	518
24	3	641
27	3	685

Table A.28: $^{DT}CCS_{He}$ of protonated F3-valine(S,S) meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	64
2	1	106
4	1	166
8	2	266
11	2	322
12	2	347
13	2	371
15	2	417

Table A.29: $^{DT}CCS_{He}$ of protonated F3-valine(S,R) meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	64
2	1	104
10	2	310
11	2	332
12	2	353
13	2	373
14	2	395

Table A.30: $^{DT}CCS_{He}$ of protonated aminoisobutyric acid (Aib) meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	54
2	1	90
3	1	114
4	1	120
5	1	143
8	2	205
9	2	209
9	2	213
10	2	232
11	2	245
12	2	276
16	2	349

Table A.31: $^{DT}CCS_{He}$ of protonated F3-aminoisobutyric acid (F3-Aib) meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	61
2	1	94
2	1	95
3	1	132
7	2	265
9	2	313
11	2	357
13	2	393
15	2	432

Table A.32: $^{DT}CCS_{He}$ of protonated aminobutyric acid (Abu) meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	55
2	1	90
3	1	118
3	1	124
8	2	219
9	2	249
10	2	259
11	2	270
12	2	295
13	2	310
14	2	330
15	2	343
16	2	368
17	2	382
18	2	400
21	3	454
22	3	475
24	3	509
25	3	524
27	3	544
28	3	564

Table A.33: $^{DT}CCS_{He}$ of protonated F2-aminobutyric acid (F2-Abu) meta clusters

n	*z*	$^{DT}CCS_{He}$ / Å^2
1	1	58
2	1	91
8	2	234
9	2	257
10	2	268
11	2	285
15	3	346

Table A.34: Amino acids (three letter code) and their protonated (+) or deprotonated (-) monomer $^{DT}CCS_{He}$ (σ_1) as well as hydrophobicity score (α) with its goodness of fit (R^2) according to equation 3.1

Amino acid	σ_1 / Å^2 (+)	*a*	R^2	σ_1 / Å^2 (-)	*a*	R^2
Ala	52	1.003	0.991			
Arg	72	0.909	0.998	72	0.938	0.998
Asn	59	0.964	0.994			
Asp	58	0.856	0.994			
Cys	58	0.956	0.997			
Glu	63	0.886	0.982	62	0.899	0.994
Gln	64	1.043	0.992			
His	67	1.029	0.990			
Ile	66	1.048	0.998	63	1.107	0.998
Leu	65	1.101	0.999			
Lys	65	1.003	0.999			
Met	68	1.053	0.998			
Phe	73	1.042	0.999	72	0.969	0.999
Pro	59	0.897	0.992			
Ser	51	0.862	0.993	52	0.886	0.993
Thr	58	0.932	0.997			
Trp	82	0.933	0.998			
Tyr	75	0.998	1.000			
Val	59	1.053	0.986			
1oF-Phe	74	1.026	0.999			
1mF-Phe	74	1.021	0.998			
1pF-Phe	74	1.014	0.998			
5F-Phe	81	0.951	0.999			
P-Phe	95	0.906	0.993			
PF5-Phe	95	0.980	0.992			
F3-Ile	68	1.118	0.998			
F3-Val(S,S)	64	1.061	0.998			
F3-Val(S,R)	64	1.080	1.000			
Aib	54	0.924	0.984			
F3-Aib	61	1.245	0.993			
Abu	55	1.057	0.997			
F2-Abu	58	1.000	0.999			

Appendix B

Assessing the Stability of Alanine-based Helices

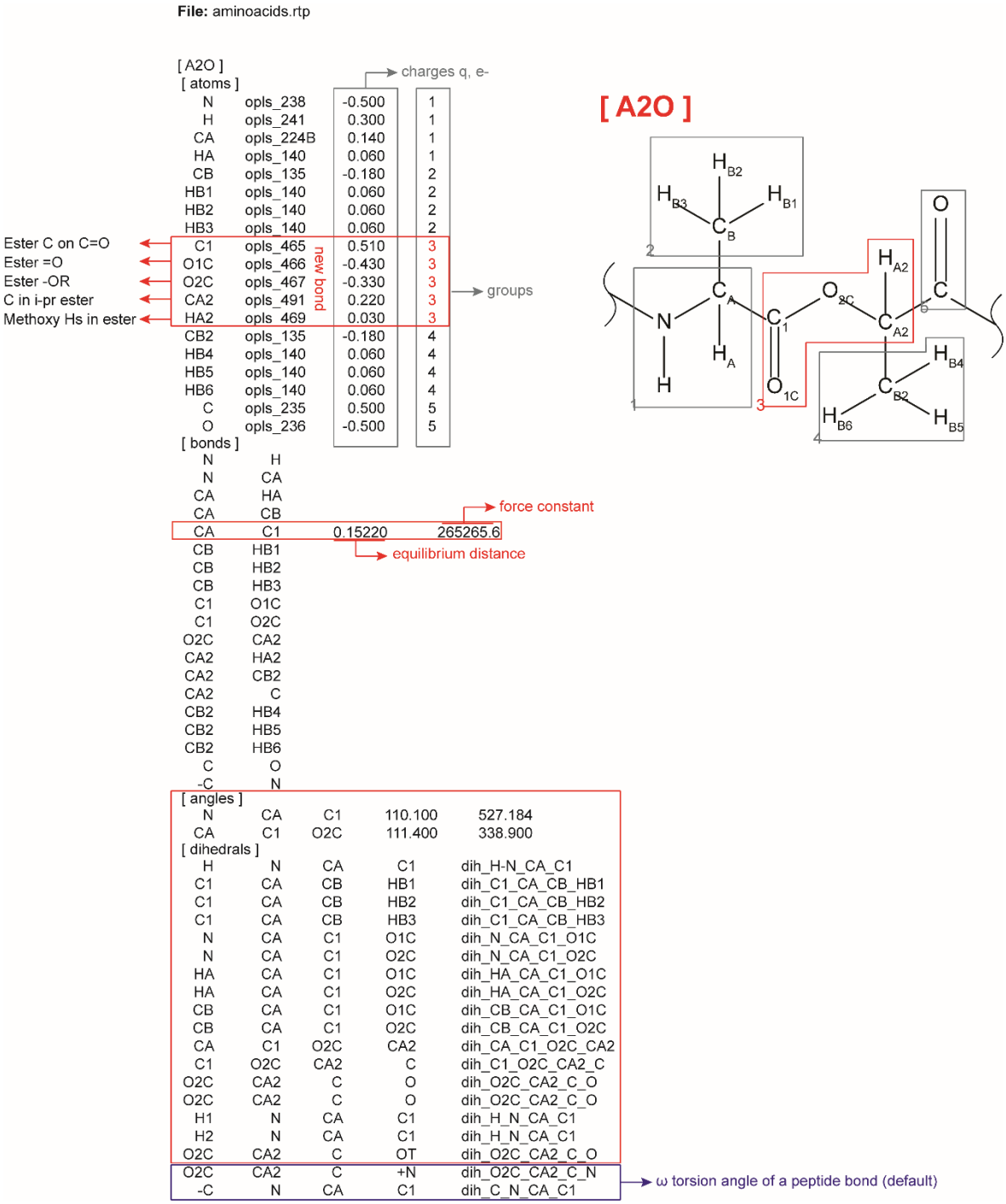


Figure B.1: Modified OPLS-AA force field. A detailed description for the lactic acid building block [A2O] is given.

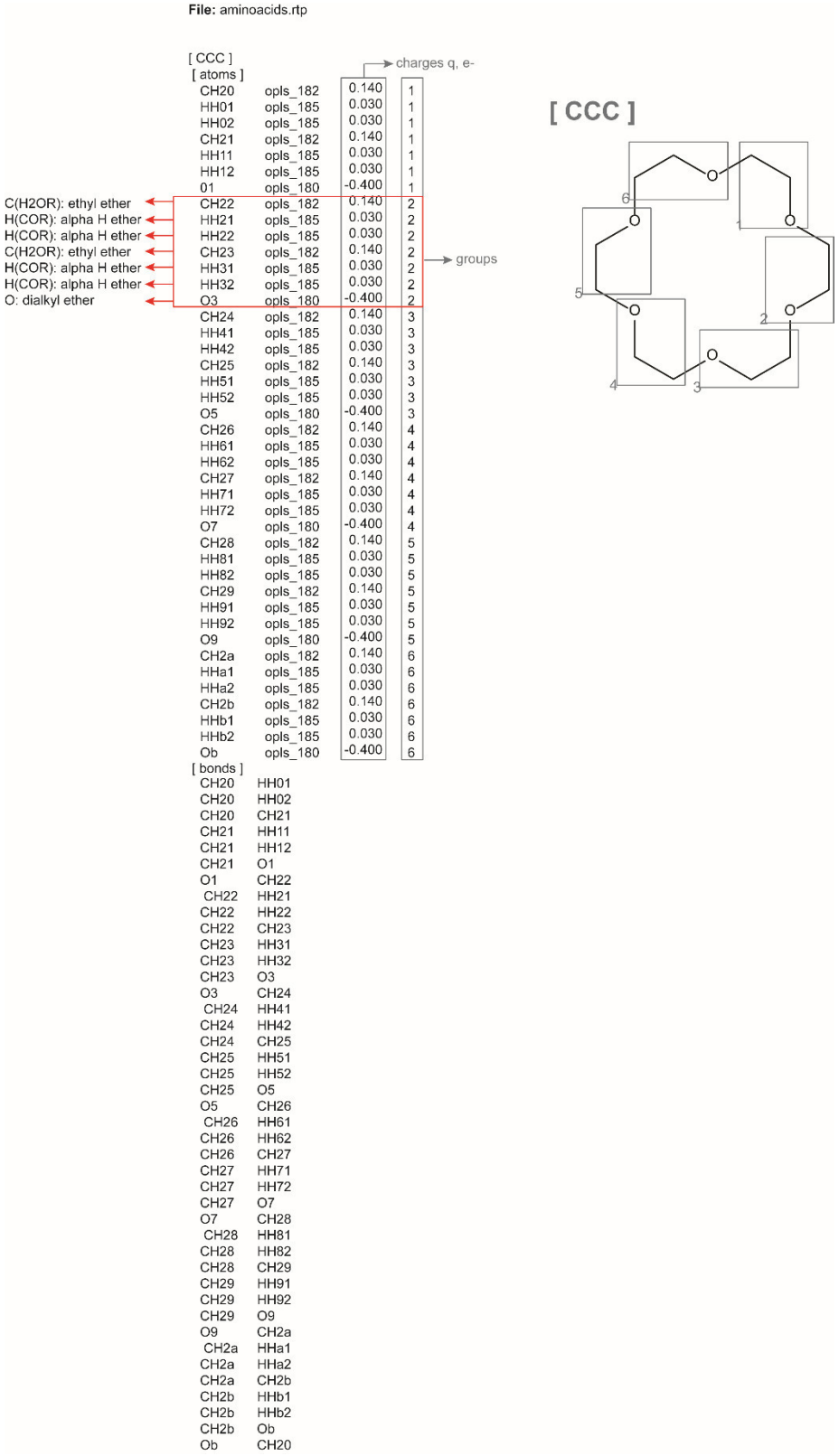


Figure B.2: Modified OPLS-AA force field. A detailed description for the 18-crown-6 ether [CCC] is given.

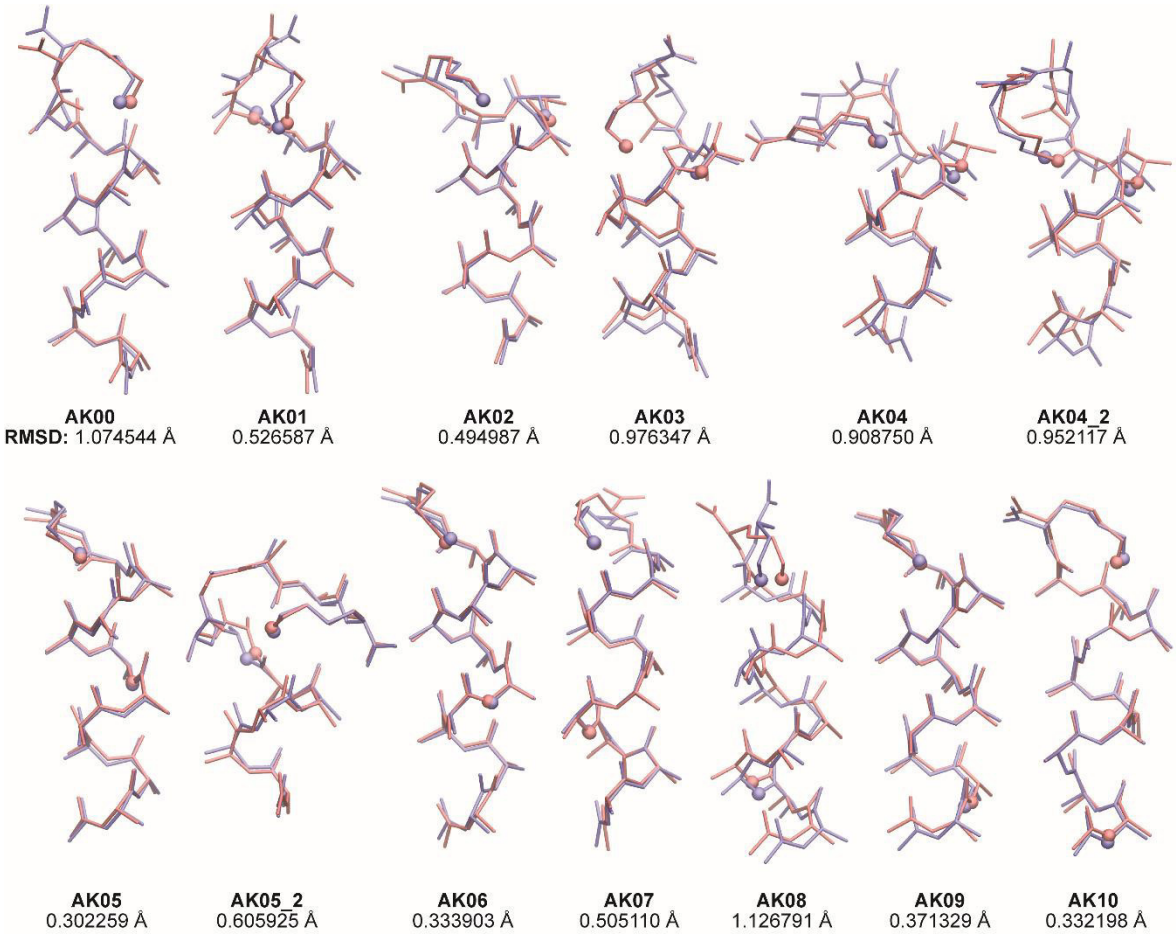


Figure B.3: Overlays of the central structures obtained by MD simulations (blue ribbon) and their respective DFT optimized structures (red ribbon). The nitrogen atom of the lysine side-chain and the oxygen atom of the depsi peptide bond are highlighted. Calculated RMSD values of each pair are given in below each overlay.

Table B.1: Comparison of experimental $^{DT}CCS_{He}$ of protonated AK00 and its depsi peptide analogues to theoretical $^{TM}CCS_{He}$ of optimized structures.

	$^{DT}CCS_{He}$ (Å^2)	$^{TM}CCS_{He}$ (Å^2)	Deviation (%)
AK00	237	240 ± 1	1.0
AK01	240	238 ± 1	-0.7
AK02	237	235 ± 2	-2.9
AK03	242	234 ± 1	-3.2
AK04	237	239 ± 2	0.8
AK04_2		241 ± 2	1.6
AK05	241	238 ± 2	-1.0
AK05_2		230 ± 2	-4.8
AK06	237	243 ± 2	2.5
AK07	236	238 ± 2	1.0
AK08	234	238 ± 1	1.7
AK09	236	239 ± 1	1.0
AK10	237	239 ± 2	1.2

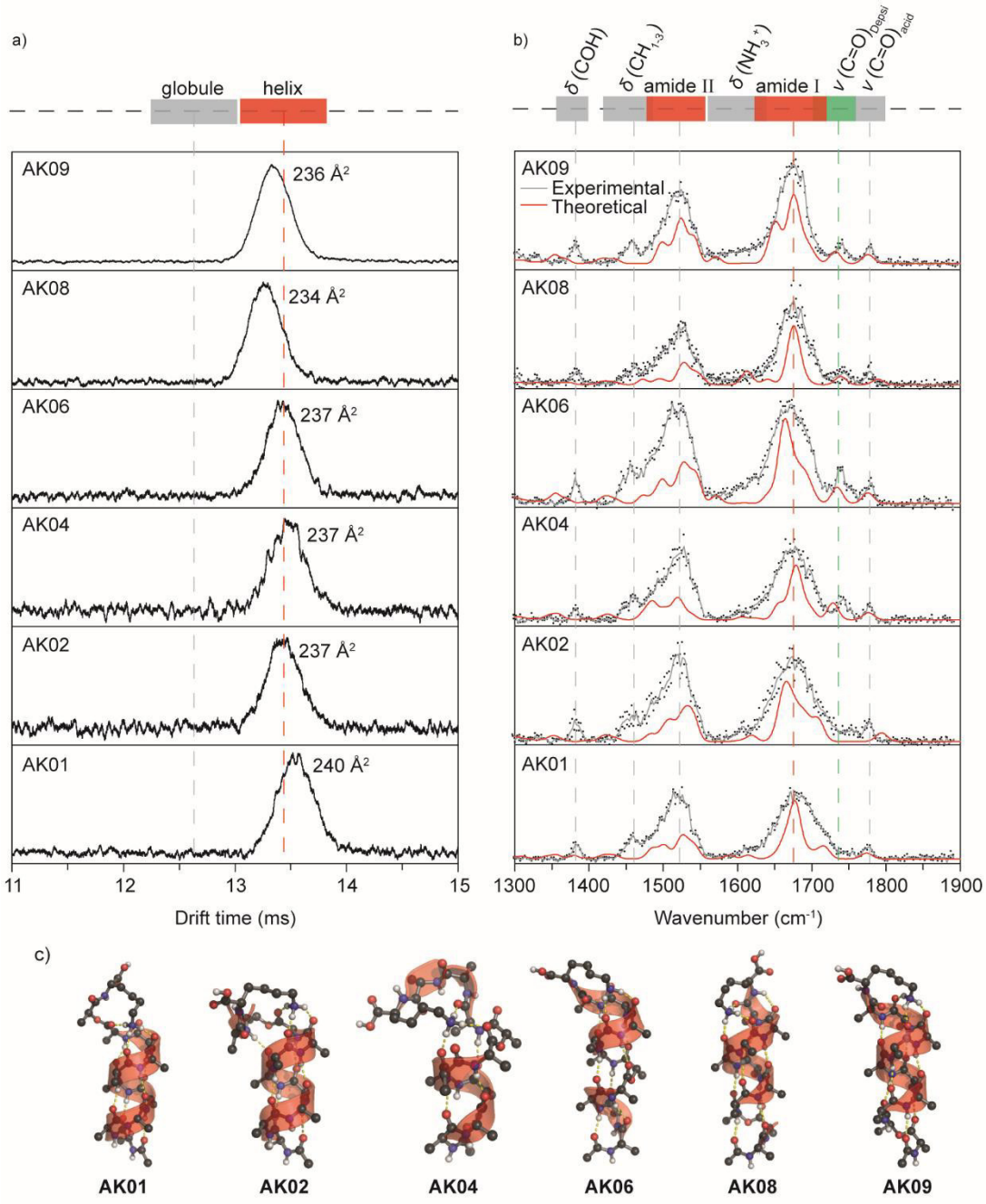


Figure B.4: a) ATDs of selected depsi peptides as well as their corresponding $^{DT}CCS_{He}$ values are shown in black. Dashed red and gray lines represent drift-time positions of calculated $^{DT}CCS_{He}$ reported for helical and globular conformations of $[Ac\text{-}Ala10Lys + H]^+$, respectively. b) Gas-phase IR spectra of ion-mobility- and m/z-selected conformations. Data from two individual scans per species are depicted as black dots. The solid gray lines represent an average of these two individual scans. Theoretical vibrational spectra of geometry-optimized structures c) are displayed as solid red lines. Band assignment is schematically shown in the upper panel.

Appendix C

Retention of Native Protein Structures in the Gas Phase

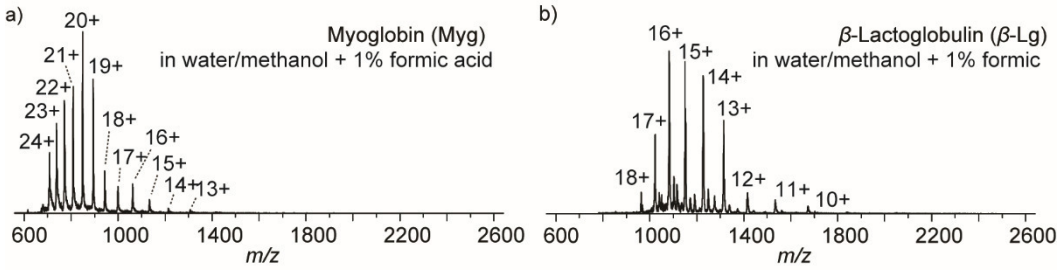


Figure C.2: Charge state distributions of proteins in acidic solution. a) Mass spectrum of myoglobin in water/methanol with 1% formic acid. The heme is readily detached from myoglobin in the acidic water/methanol, and thus only apo-form of myoglobin is observed. b) Mass spectrum of 1/1 mixture of β-lactoglobulin and palmitic acid in acidic water/methanol. Palmitic acid is not bound to β-lactoglobulin under these solvent conditions and only bare proteins are detected at high charge states.

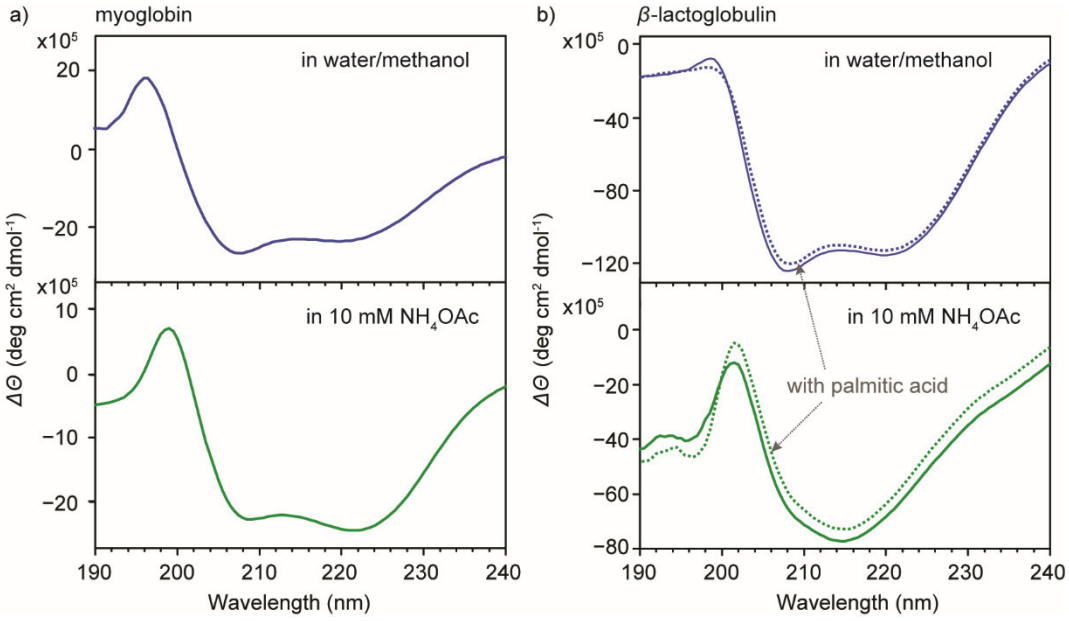


Figure C.1: CD spectra. a) Myoglobin and b) β-lactoglobulin either in water/methanol (blue) or buffered water (green). Dotted lines denote CD spectra of 1/1 mixture of β-lactoglobulin and palmitic acid.

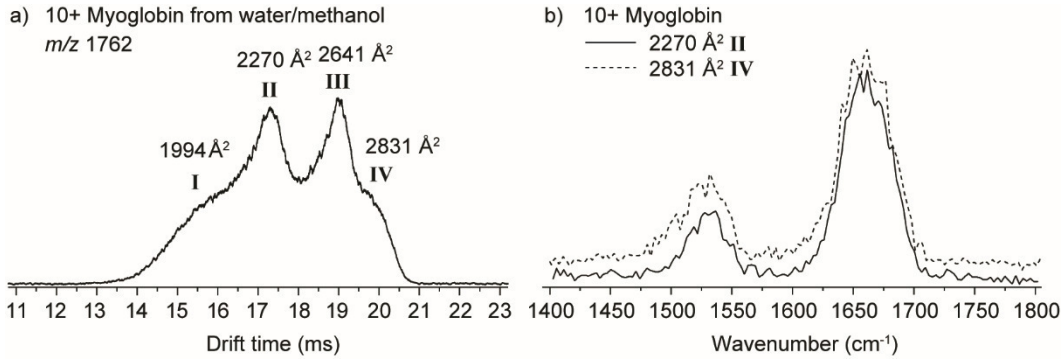


Figure C.3: ATDs and IR spectra of two conformers of myoglobin 10+. a) ATD of 10+ myoglobin sprayed from water/methanol. At least four different conformers (I-IV) with vastly different $^{DT}CCS_{He}$ values (1994, 2270, 2641, and 2831 Å²) are observable. b) Gas-phase IR spectra of two different conformers (solid line for II and dotted line for IV) of myoglobin 10+.

Appendix D

β-Sheet Formation in Peptide Amyloid Assemblies

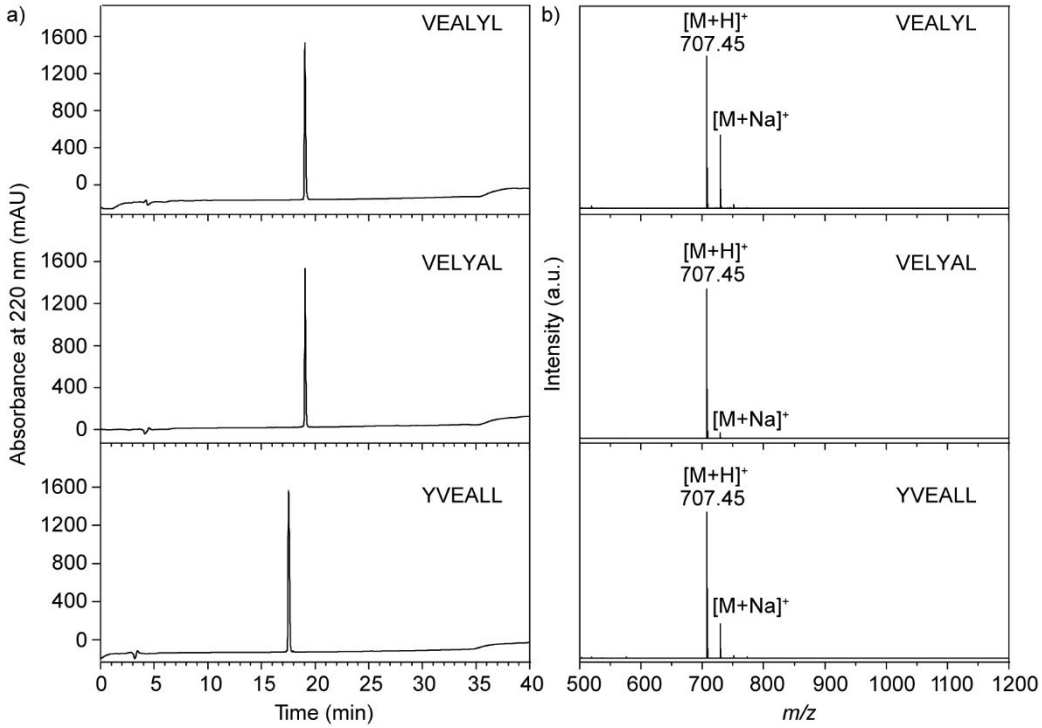


Figure D.1: a) RP-HPLC chromatograms and b) *n*ESI mass spectra of purified VEALYL, VELYAL and YVEALL. According to the chromatograms and mass spectra, purities of the synthesized peptides are more than 99%.

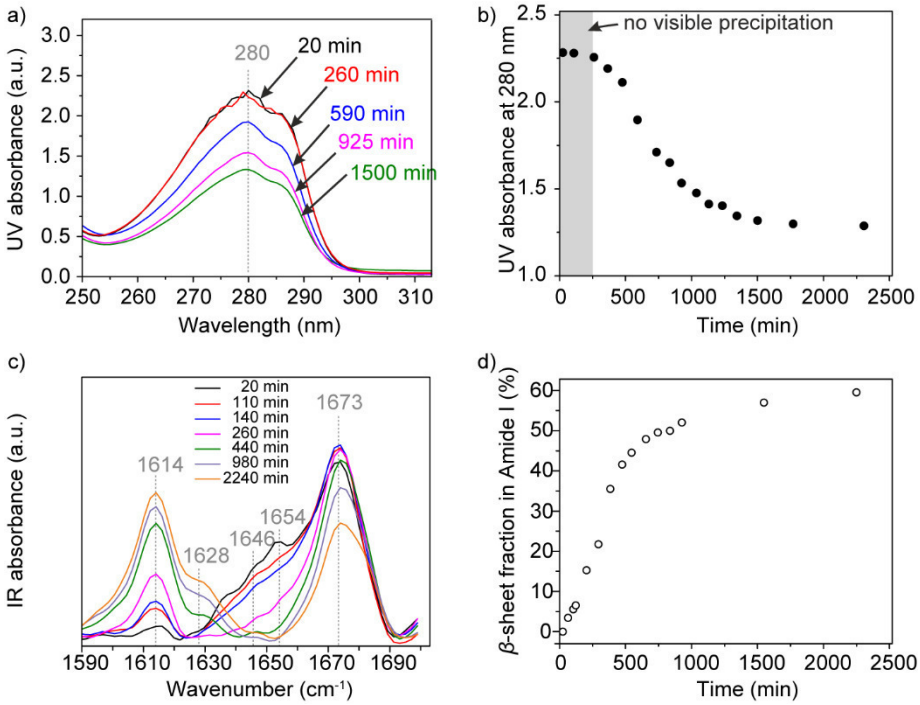


Figure D.2: a) UV absorption spectra of VEALYL solution in D2O/CH_3OD after 20, 260, 590, 925, and 1500 min of incubation at room temperature and b) measured UV absorption at 280 nm as a function of incubation time. c) FT-IR spectra of VEALYL solution after different incubation times and d) determined β-sheet fractions as a function of incubation time. The amide I feature was fit with multiple Gaussians, and the area of Gaussians in 1600-1630 cm^{-1} are used to determine the β-sheet fraction.

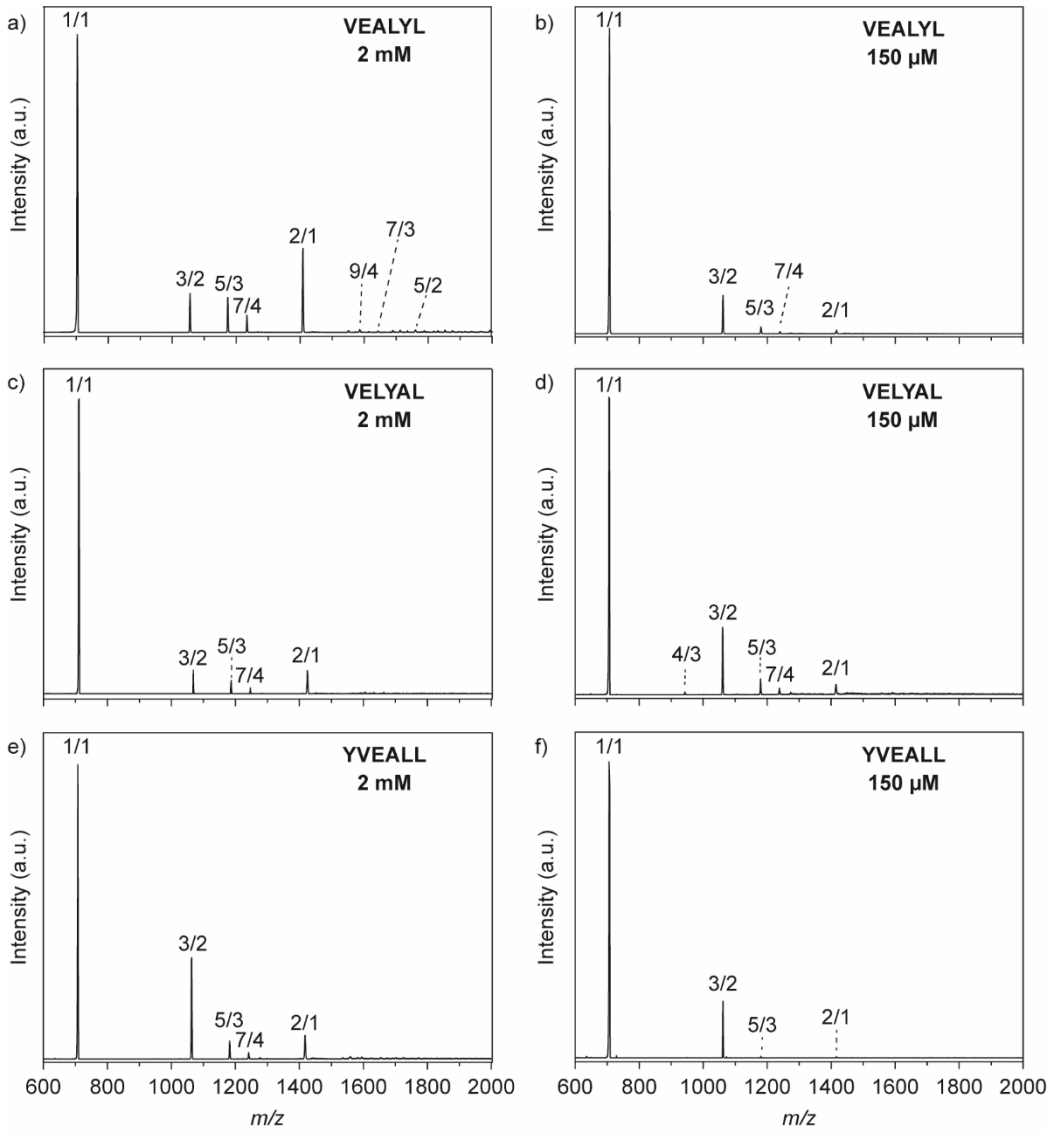


Figure D.3: nESI time-of-flight mass spectra of 2 mM and 0.15 mM VEALYL (a, b), VELYAL (c, d) and YVEALL (d, e) solutions in water/methanol. The abundant peaks are annotated with their corresponding n/z, where n stands for the oligomer number n and z for the charage.

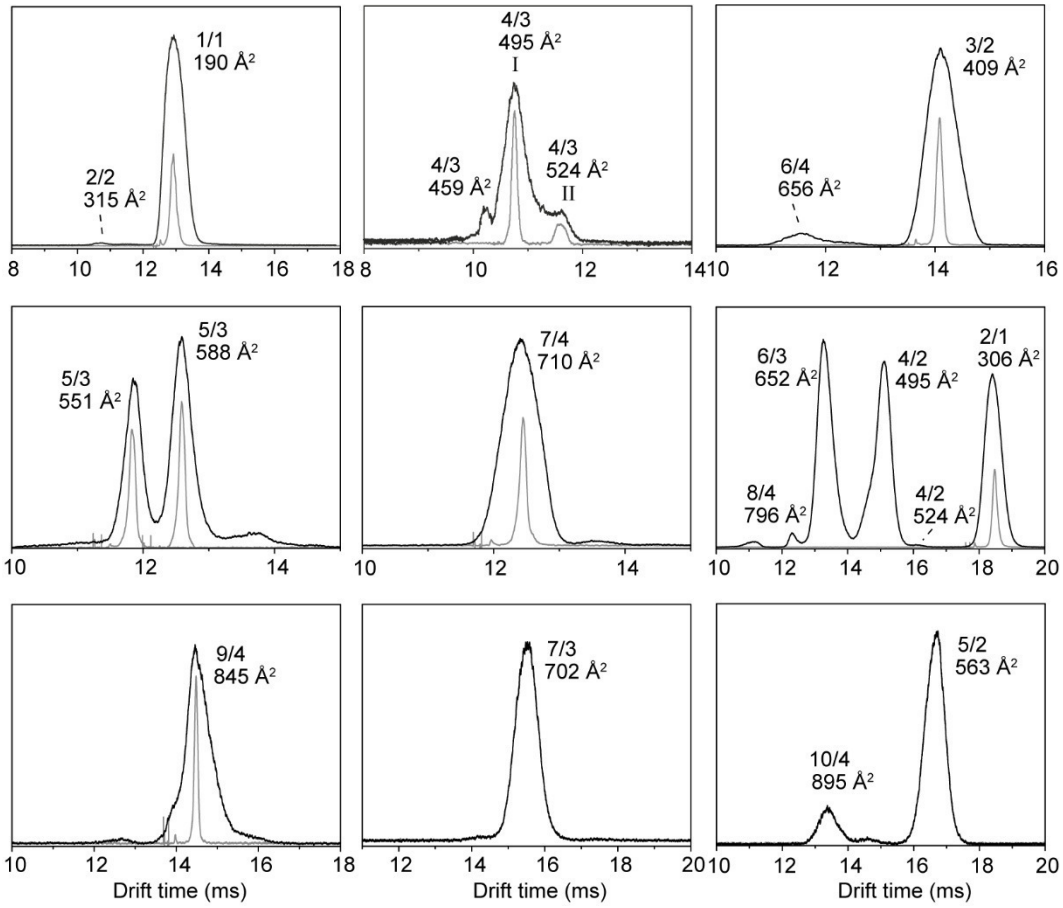


Figure D.4: ATDs of VEALYL recorded for 2 mM sample solution (water/methanol = 1/1, v/v). The recorded data is shown in black., while the 100 μs-wide sliced ATD used for IRMPD spectroscopy is shown in gray. Each ATD feature is labelled by n/z with n being the oligomer number, and z the overall charge. The $^{DT}CCS_{He}$ values are given in Å².

Table D.1: Fitted parameters for amide I features of VEALYL oligomers

	Fitted parameters of *n*-th Gaussian: Peak center (FWHM) Relative area							R^2
	n = 0	*n* = 1	*n* = 2	*n* = 3	*n* = 4	*n* = 5	*n* = 6	
3/2	-	-	-	1646 (18) 32	1658 (8) 22	1671 (25) 100	-	0.99
4/3 I	1582 (25) 24	1609 (16) 17	1627 (31) 61	1645 (15) 52	1661 (17) 100	1680 (29) 98	1692 (10) 15	0.99
4/3 II	1581 (25) 26	1608 (17) 16	1626 (33) 53	1645 (14) 22	1657 (29) 100	1677 (29) 81	1690 (11) 18	0.99
5/3	1579 (24) 12	1609 (18) 9	1626 (30) 35	1641 (10) 14	1657 (24) 100	1676 (32) 77	1690 (10) 9	0.99
7/4	1575 (25) 10	1612 (18) 14	1638 (17) 17	1646 (15) 33	1659 (11) 36	1672 (25) 100	1688 (8) 8	0.99
9/4	1577 (14) 6	1609 (23) 18	1638 (19) 26	1648 (13) 38	1662 (14) 100	1677 (15) 52	1690 (7) 8	0.99

Table D.2: $^{DT}CCS_{He}$ values and β-sheet fractions in the amide I feature of VEALYL oligomers

n	*z*	$^{DT}CCS_{He}$ (Å^2)	$^{DT}CCS_{He}$ /$CCS_{isotropic}$	*β*-sheet ratio in amide I
1	1	190	1.05	0
2	1	306	1.06	0
2	2	315	1.09	
3	2	409	1.08	0
4	2	495	1.09	
		524	1.14	
4	3	459	1.00	
		595	1.08	0.26
		524	1.15	0.30
5	2	563	1.06	
5	3	551	1.04	0.22
		588	1.11	0.22
6	6	652	1.09	
6	4	656	1.09	
7	3	702	1.06	
7	4	710	1.07	0.19
8	4	796	1.10	
9	4	845	1.08	0.20
10	4	895	1.07	

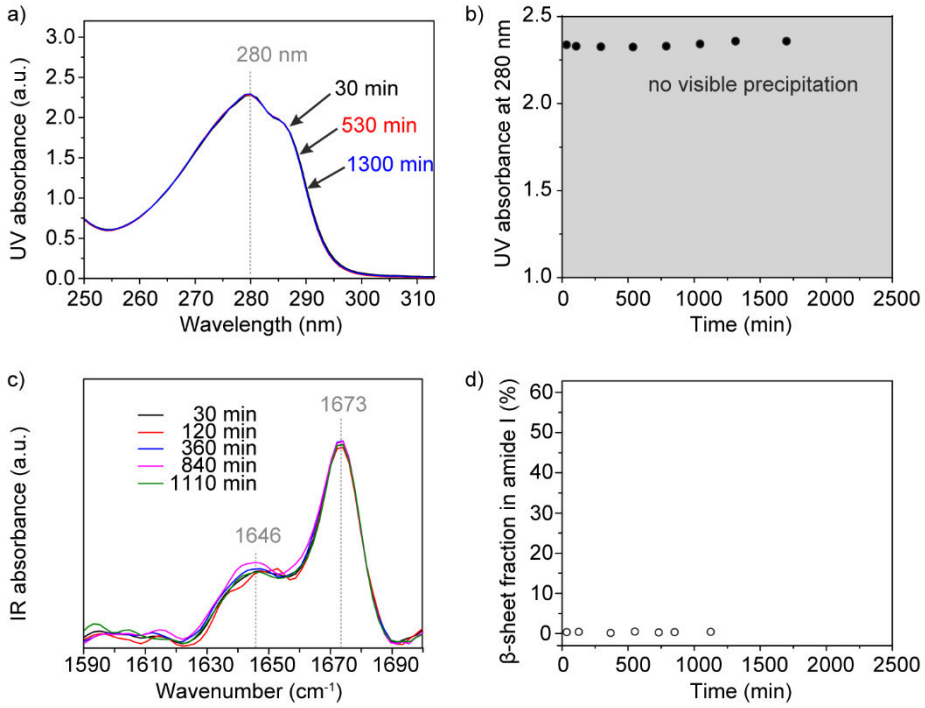


Figure D.5: a) UV absorption spectra of VELYAL solution in D_2O/CH_3OD after 30, 530, and 1300 min of incubation at room temperature and b) measured UV absorption at 280 nm as a function of incubation time. c) FT-IR spectra of VELYAL solution after different incubation times and d) determined β-sheet fractions as a function of incubation time. The amide I feature was fitted with multiple Gaussians, and the area of Gaussians in 1600-1630 cm^{-1} are used to determine the β-sheet fraction.

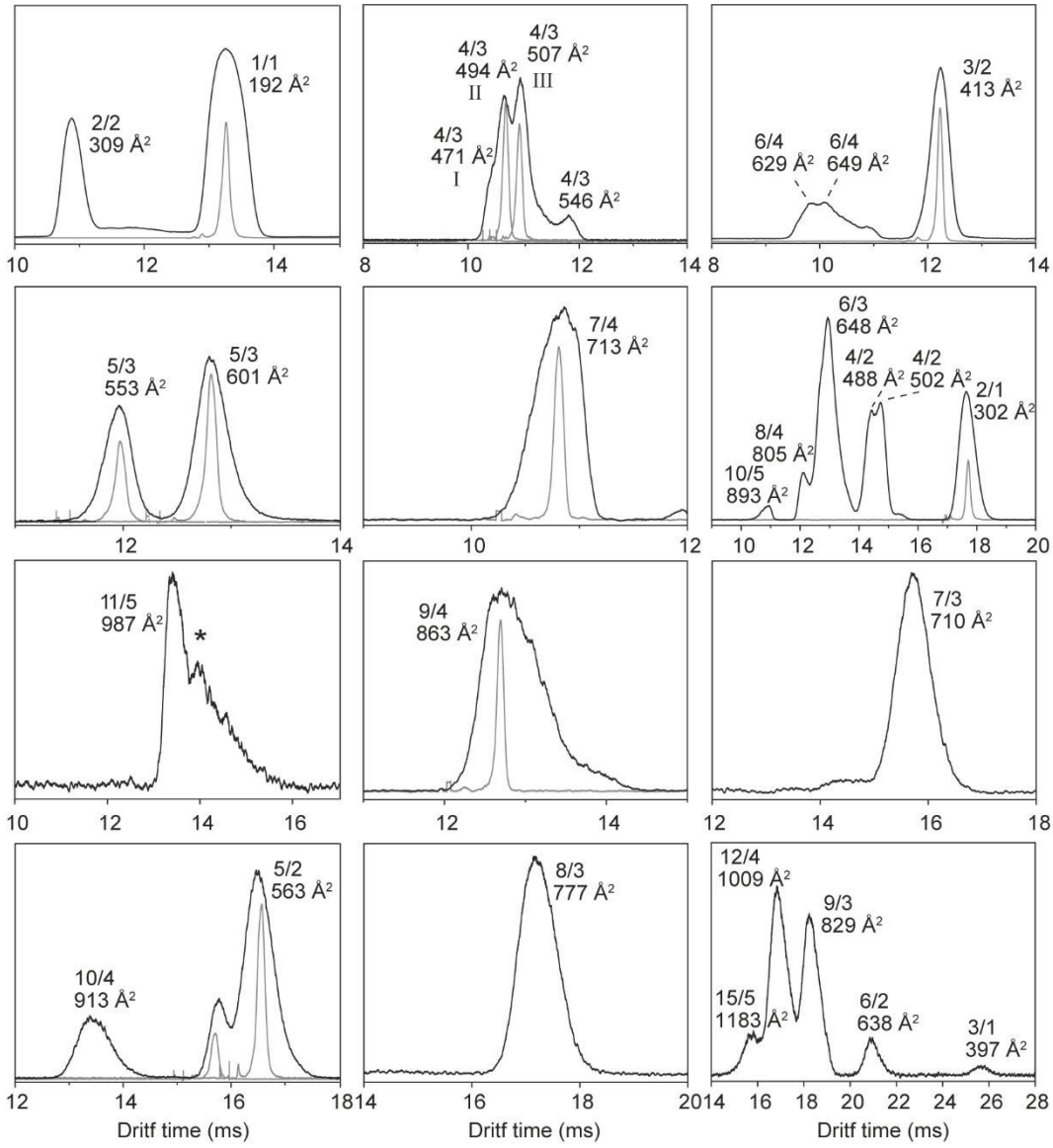


Figure D.6: ATDs of VELYAL recorded for 2 mM sample solution (water/methanol = 1/1, v/v). The recorded data is shown in black, while the 100 μs-wide sliced ATD used for IRMPD spectroscopy is shown in gray. Each ATD feature is labelled by n/z notation with n being the oligomer number and z the charge.

Table D.3: Fitted parameters for amide I features of VELYAL oligomers

	Fitted parameters of *n*-th Gaussian: Peak center (FWHM) Relative area							R^2
	n = 0	*n* = 1	*n* = 2	*n* = 3	*n* = 4	*n* = 5	*n* = 6	
2/1	-	-	-	1643 (15) 12	1658 (12) 18	1676 (26) 100	-	0.99
3/2	-	-	-	1643 (14) 8	1657 (14) 16	1676 (31) 100	-	1.00
4/3	1588 (20) 8	1619 (34) 21	-	1645 (14) 16	1656 (8) 16	1667 (25) 100	1692 (5) 3	0.99
5/3	-	1609 (19) 4	1630 (19) 11	1644 (12) 16	1661 (24) 100	1676 (22) 52	1695 (9) 5	0.99
7/4	1589 (34) 13	-	1625 (25) 20	1645 (14) 18	1662 (27) 93	1685 (34) 100	1702 (4) 2	0.99
9/4	-	-	-	1648 (22) 93	1667 (8) 21	1680 (21) 100	-	0.99

Table D.4: $^{DT}CCS_{He}$ values and β-sheet fractions in the amide I feature of VELYAL oligomers

n	*z*	$^{DT}CCS_{He}$ (Å^2)	$^{DT}CCS_{He}$ /$CCS_{isotropic}$	*β*-sheet ratio in amide I
1	1	192	1.00	
2	1	302	1.00	
2	2	309	1.01	
3	1	397	1.00	
3	2	413	1.00	0
4	2	488	1.01	
		502	1.04	
4	3	471	1.00	
		494	1.02	
		507	1.05	0.12
5	2	563	1.00	
5	3	553	1.00	0.092
		601	1.07	0.091
6	2	638	1.00	
6	3	648	1.02	
6	4	629	1.00	
		649	1.02	
7	3	710	1.01	
7	4	713	1.01	0.095
8	3	777	1.01	
8	4	805	1.05	
9	3	829	1.00	
9	4	863	1.04	0
10	4	913	1.00	
10	5	893	1.02	
11	4	948	1.00	
		959	1.01	
11	5	987	1.04	
12	4	1009	1.00	
15	5	1183	1.01	

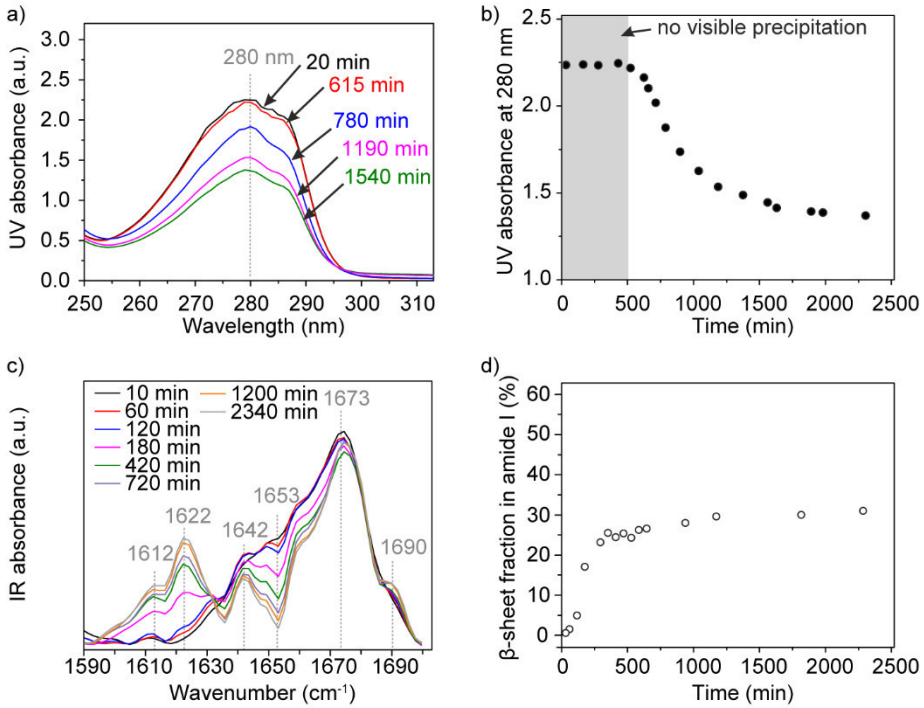


Figure D.7: a) UV absorption spectra of YVEALL solution in D_2O/CH_3OD after 20, 615, 780, 1190, and 1540 min of incubation at room temperature and b) measured UV absorption at 280 nm as a function of incubation time. c) FT-IR spectra of YVEALL solution after different incubation times and d) determined β-sheet fraction as a function of incubation time. The amide I feature was fit with multiple Gaussians, and the area of Gaussians in 1600-1630 cm^{-1} are used to determine the β-sheet fraction.

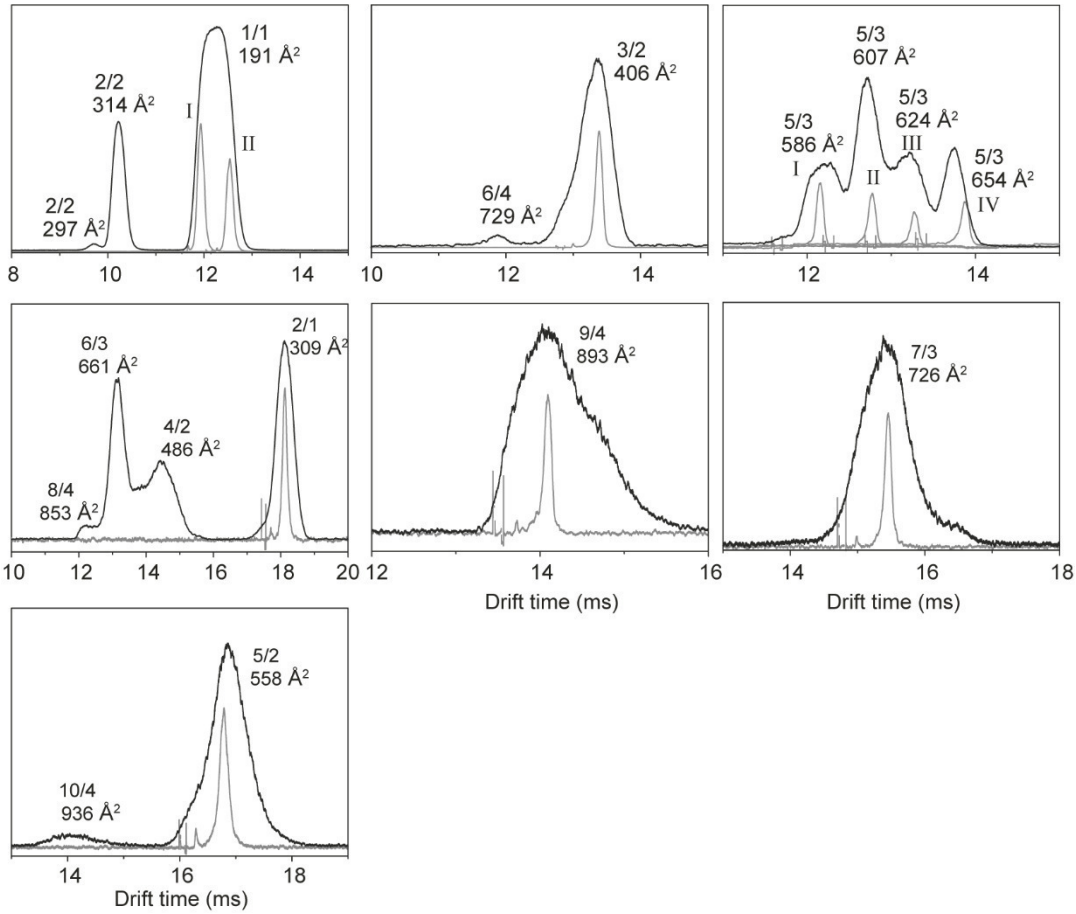


Figure D.8: ATDs of YVEALL for a 2 mM sample solution (water/methanol = 1/1, v/v). The recorded data is shown in black, while the 100 μs-wide sliced ATD used for IRMPD spectroscopy is shown in gray. Each ATD feature is labelled by n/z with n being the oligomer number and z the charge.

Table D.5: Fitted parameters for amide I features of YVEALL oligomers

	Fitted parameters of *n*-th Gaussian: Peak center (FWHM) Relative area							R^2
	n = 0	*n* = 1	*n* = 2	*n* = 3	*n* = 4	*n* = 5	*n* = 6	
2/1	1590 (25) 12	1619 (16) 14	-	1644 (20) 37	1668 (18) 100	1683 (19) 93	1698 (9) 8	0.99
3/2	1591 (25) 9	1618 (25) 15	-	1645 (24) 51	1666 (25) 100	1685 (25) 84	1699 (10) 6	0.99
5/2	-	-	-	-	1666 (35) 100	-	-	1.00
5/3 I	1588 (13) 4	1613 (29) 29	-	1642 (12) 8	1666 (30) 100	-	-	1.00
5/3 II	1594 (22) 33	1614 (10) 11	1626 (24) 62	1644 (9) 21	1660 (21) 100	1676 (22) 94	1692 (9) 11	0.99
5/3 III	1593 (23) 33	1613 (17) 25	1630 (20) 58	1645 (12) 25	1662 (23) 100	1678 (21) 69	1694 (8) 12	0.99
5/3 IV	1580 (24) 37	1613 (15) 29	1631 (33) 100	1646 (14) 16	1660 (27) 48	1672 (28) 42	1691 (12) 22	0.99
7/3	1584 (32) 13	1615 (25) 14	1635 (33) 28	1648 (33) 52	1668 (12) 10	1673 (35) 100	1692 (17) 8	0.99
9/4	1577 (23) 7	1613 (21) 21	1636 (22) 35	1647 (17) 46	1663 (16) 97	1674 (18) 100	1688 (12) 13	0.99

Table D.6: $^{DT}CCS_{He}$ values and β-sheet fractions in the amide I feature of YVEALL oligomers

n	z	$^{DT}CCS_{He}$ (Å^2)	$^{DT}CCS_{He}$ /$CCS_{isotropic}$	*β*-sheet ratio in amide I
1	1	191	1.04	
2	1	309	1.06	0.082
2	2	297	1.02	
		314	1.08	
3	2	406	1.06	0.084
4	2	461	1.00	
		486	1.05	
5	2	558	1.04	0
5	3	586	1.10	0.21
		607	1.13	0.22
		624	1.16	0.34
		654	1.22	0.59
6	3	661	1.10	
6	4	729	1.21	
7	3	726	1.08	0.23
8	4	853	1.17	
9	4	893	1.12	0.21
10	4	936	1.10	

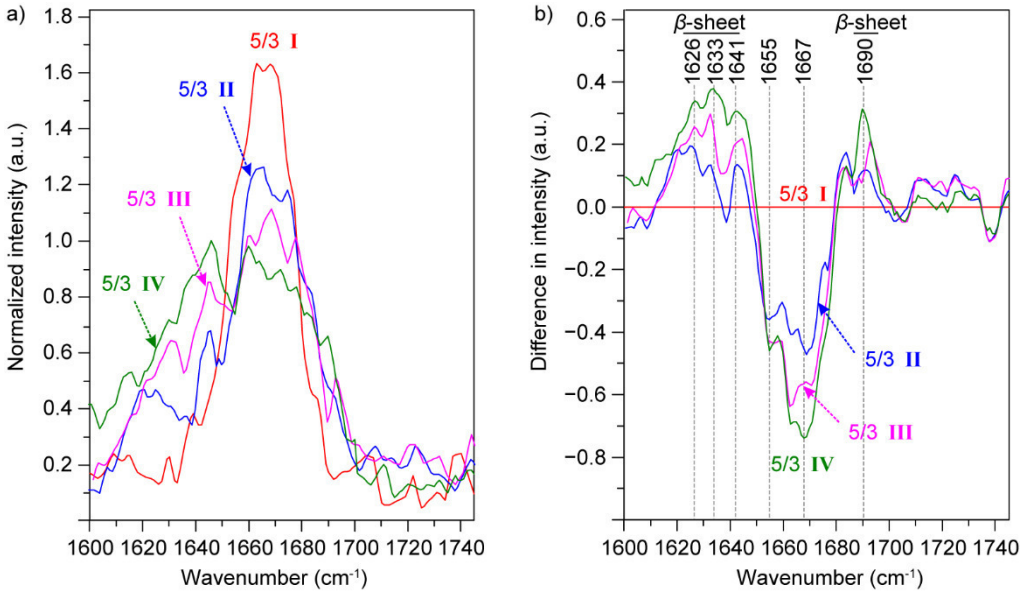


Figure D.9: a) The area-normalized IRMPD spectra and b) the IR difference spectra of YVEALL pentamers (n/z = 5/3 I-IV) in amide I region. The area of each IR spectrum in 1580-1750 cm^{-1} was determined by integration and used to normalize the IR intensities in amide I region. Difference IR spectrum was built from those area-normalized IRMPD spectra.

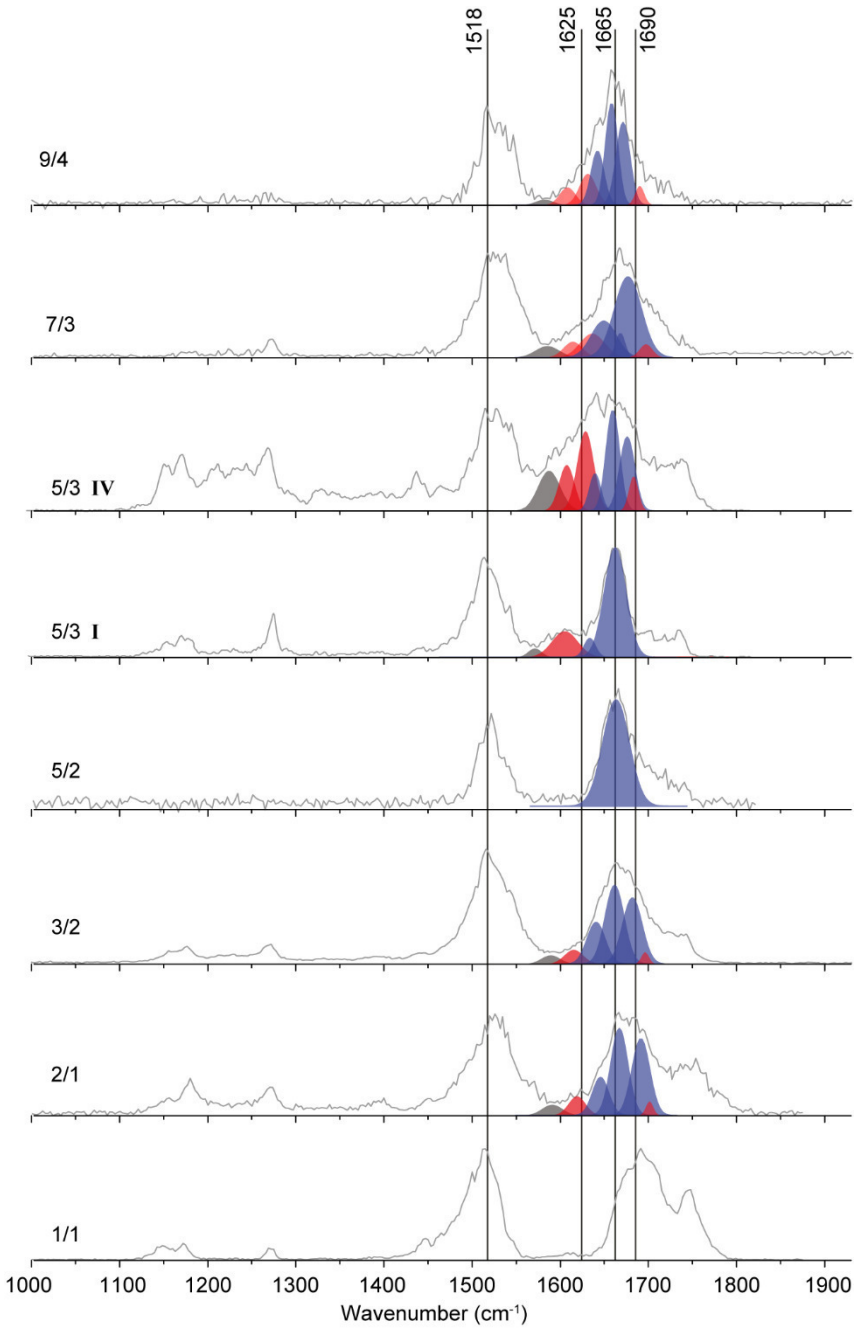


Figure D.10: IR spectra of YVEALL monomer and oligomers up to the quadruply protonated nonamer (n/z = 9/4). The amide I region (1600-1700 cm^{-1}) was fit with multiple Gaussians to deduce the secondary structural components. Red and blue envelopes denote the amide I features for β-sheet (1610-1640 cm^{-1} and ~ 1690 cm^{-1}) and turn-like structures (1660-1685 cm^{-1}).

Appendix E

NFGAIL Amyloid Oligomers

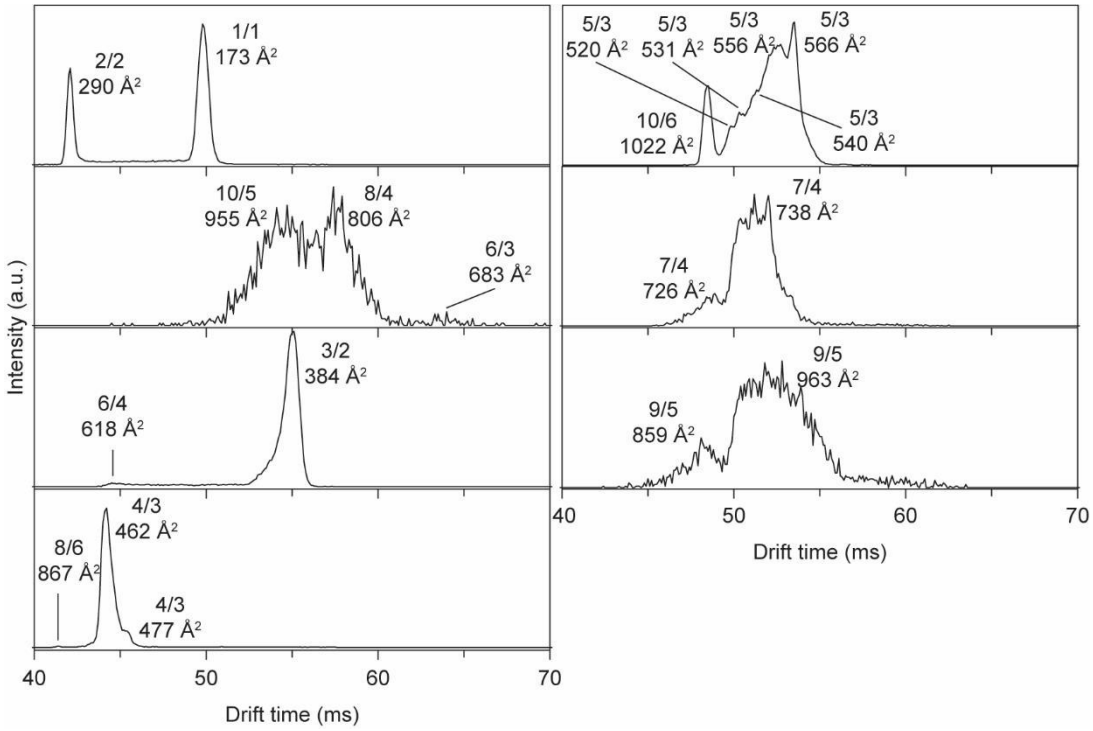


Figure E.1: ATDs of NFGAIL (1 mM) in ammonium acetate buffer (10 mM, pH~7) recorded at a home-built high-resolution IM-MS instrument (HiRes, Resolution, R=100) located in Santa Barbara. Each ATD feature is labelled by its n/z notation with n being the oligomer number and z the overall charge. The respective $^{DT}CCS_{He}$ values are given in Å².

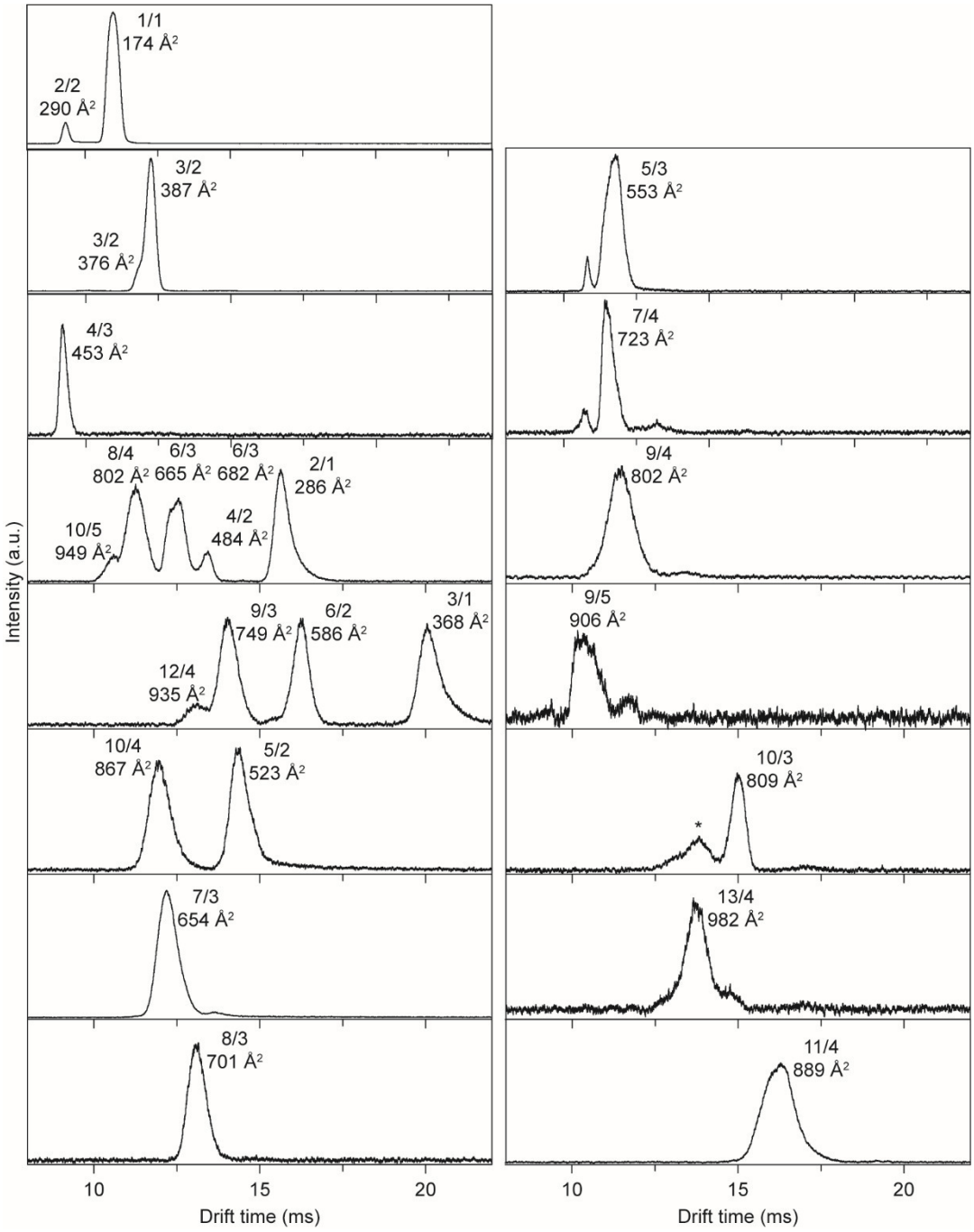


Figure E.2: ATDs of NFGAIL (1 mM) in ammonium acetate buffer (10 mM, pH~7) recorded at a home-built high-resolution IM-MS instrument (iMob, R=40) located in Berlin. Each ATD feature is labelled by its n/z notation with n being the oligomer number and z the overall charge. The respective $^{DT}CCS_{He}$ values are given. Overlapping noise features are shown with an asterisk.

Table E.1: Gaussian fitting parameters for amide I bands of NFGAIL oligomers

	Fitted parameters of *n*-th Gaussian: Peak center (FWHM) Relative area							R^2
	n = 0	*n* = 1	*n* = 2	*n* = 3	*n* = 4	*n* = 5	*n* = 6	
2/1	-	-	-	-	1666 (15) 15	1684 (25) 100	-	0.99
3/2	-	-	-	1660 (18) 27	1671 (6) 5	1676 (30) 100	-	0.99
4/3	1592 (30) 28	1626 (18) 30	1640 (8) 14	1658 (15) 66	1670 (23) 100	1690 (30) 100	1710 (30) 4	0.95
5/2	-	-	-	-	1663 (30) 50	1675 (25) 100	-	0.99
5/3 520 Å^2	-	-	-	1656 (30) 42	1663 (10) 17	1685 (26) 100	-	0.95
5/3 540 Å^2	-	-	1640 (18) 10	1660 (5) 3	1674 (30) 100	1682 (25) 15	1710 (10) 1	0.96
5/3 556 Å^2	1598 (30) 28	1625 (14) 14	1640 (13) 23	1660 (22) 82	1672 (27) 100	1690 (30) 82	1710 (30) 5	0.99
7/3	-	-	-	1660 (6) 3	1667 (26) 100	1690 (9) 3	-	0.99
7/4	1600 (24) 21	1621 (5) 8	1639 (19) 54	1655 (11) 41	1670 (19) 100	1690 (22) 49	1708 (6) 5	0.94
8/3	-	-	1640 (13) 11	1657 (12) 28	1671 (25) 100	-	-	0.99
9/4	1600 (18) 15	1629 (26) 35	1640 (6) 8	-	1658 (19) 80	1682 (30) 100	1710 (30) 5	0.94
9/5	1595 (27) 29	1617 (18) 18	1640 (17) 47	1660 (6) 7	1666 (25) 100	1690 (30) 97	1710 (10) 5	0.94
11/4	1600 (30) 14	-	-	1657 (30) 100	1671 (17) 57	1690 (12) 20	-	0.99

Table E.2: $^{DT}CCS_{He}$ values (measured on the iMob[70] and HiRes[251] instruments) and β-sheet IR intensity ratio in the amide I feature of NFGAIL oligomers. Bold $^{DT}CCS_{He}$ values have been used to constructed **Figure 7.1**d.

n	*z*	$^{DT}CCS_{He}$ (Å^2) iMob[70]	$^{DT}CCS_{He}$ (Å^2) HiRes	*β*-sheet ratio in amide I
1	1	**174**	**173**	0
2	1	**286**		0
2	2	**290**	**290**	
3	1	**368**		
3	2	376		
		387	**384**	0
4	2	**484**		
4	3	**453**	**462**	0.14
			477	
5	2	**523**		0
5	3		**520**	0
			531	
			540	0.09
		553	**556**	
			566	0.13
6	2	**586**		
6	3	**665**		
		682	**683**	
6	4		**618**	
7	3	**654**		0
7	4	**723**	**726**	0.24
			738	
8	3	**701**		0
8	4	**802**	**806**	
8	6		**867**	
9	3	**749**		
9	4	**802**		0.20
9	5	**906**	859	0.23
			963	
10	3	**809**		
10	4	**867**		
10	5	**949**	**955**	
10	6		**1022**	
11	4	**889**		0
12	4	**935**		
13	4	**982**		

List of Publications

1. **W. Hoffmann**, J. Hofmann, K. Pagel, Energy Resolved Ion Mobility Mass Spectrometry – A Concept to Improve the Separation of Isomeric Carbohydrates, *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 471-479. → Ron Hites Award
2. **W. Hoffmann**, M. Marianski, S. Warnke, J. Seo, C. Baldauf, G. von Helden, K. Pagel, Assessing the Stability of Alanine-Based Helices by Conformer Selective IR Spectroscopy, *Phys. Chem. Chem. Phys.* **2016**, *18*, 19950-19954.
3. J. Seo, **W. Hoffmann**, S. Warnke, M. T. Bowers, K. Pagel, G. von Helden, Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study, *Angew. Chem. Int. Ed.* **2016**, *55*, 14173-14176. → Back cover: Angew. Chem. Int. Ed. **2016**, *45*
4. S. Warnke, **W. Hoffmann**, J. Seo, E. De Genst, G. von Helden, K. Pagel, From Compact to String – The Role of Secondary and Tertiary Structure in Charge-Induced Unzipping of Gas-Phase Proteins, *J. Am. Soc. Mass Spectrom.* **2016**, *28*, 638-646.
5. J. Seo‡, **W. Hoffmann**‡, S. Warnke, X. Huang, S. Gewinner, W. Schöllkopf, M. T. Bowers, G. von Helden, K. Pagel, An Infrared Spectroscopy Approach to Follow β-Sheet Formation in Peptide Amyloid Assemblies, *Nat. Chem.* **2017**, *9*, 39-44.
6. **W. Hoffmann**, G. von Helden, K. Pagel, Ion Mobility-Mass Spectrometry and Orthogonal Gas-Phase Techniques to Study Amyloid Formation and Inhibition, *Curr. Opin. Struct. Biol.* **2017**, *46*, 7-15.
7. E. Mucha, A. I. González Flórez, M. Marianski, D. A. Thomas, **W. Hoffmann**, W. B. Struwe, H. S. Hahm, S. Gewinner, W. Schöllkopf, P. H. Seeberger, G. von Helden, K. Pagel, Glycan Fingerprinting using Cold-Ion Infrared Spectroscopy, *Angew. Chem. Int. Ed.* **2017**, 56, 11248-11251.
8. J. Seo, **W. Hoffmann**, S. Malerz, S. Warnke, M. T. Bowers, K. Pagel, G. von Helden, Side Chain Effect on the Structures of Protonated Amino Acid Dimers: A Gas-Phase Infrared Spectroscopy Study, *Int. J. Mass Spectrom.* **2018**, 429, 115-120.
9. **W. Hoffmann**, K. Pagel, Von Normalen Proteinen zu Unlöslichen Ablagerungen, *Nachr. Chem.* **2017**, *65*, 874-878.

‡ The authors contributed equally to this work.

10. **W. Hoffmann**, K. Folmert, J. Moschner, H. Xing, H. von Berlepsch, B. Koksch, M. T. Bowers, G. von Helden, K. Pagel, The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation, *J. Am. Chem. Soc.* **2018**, *140*, 244-249.

11. U. Warzok, M. Marianski, **W. Hoffmann**, L. Turunen, K. Rissanen, K. Pagel, C. Schalley, Large [N···I^+···N] Halogen-Bonded Supramolecular Capsules: Structure and Reactivity Elucidated by Mass Spectrometry and Ion Mobility, **submitted for review**.

12. S. Wittig, C. Haupt, **W. Hoffmann**, S. Kostmann, K. Pagel, C. Schmidt, Chemical Cross-Linking and Native Mass Spectrometry reveal Oligomerisation of Synaptobrevin-2, *J. Am. Soc. Mass Spectrom.* **2018**, **in press**.

13. C. Huber, K. Pagel, **W. Hoffmann**, S. Heiles, K. C. Schäfer, M. Kompauer, S. Gerbig, J. Soltwisch, Trendbericht: Analytische Chemie 2016/2017, *Nachr. Chem.* **2018**, 66, 389-399.

14. M. Göth, F. Witte, M. Quennet, P. Jungk, G. Podolan, D. Lentz, **W. Hoffmann**, K. Pagel, H. U. Reissig, B. Paulus, C. Schalley, To Anion-π or not to Anion-π: The Case of Anion-Binding to Divalent Fluorinated Pyridines in the Gas Phase, *Chem. Eur. J.* **2018**, **in press**.

15. **W. Hoffmann**, R. Chang, M. Marianski, J. Seo, M. T. Bowers, G. von Helden, K. Pagel, Supramolecular Assembly of Proline Mediated by Alkali Metal Cations, **submitted for review**.

16. **W. Hoffmann**, J. Langenhan, S. Huhmann, J. Moschner, M. Accorsi, J. Seo, J. Rademann, B. Koksch, M. T. Bowers, G. von Helden, K. Pagel, Novel Ion Mobility-Mass Spectrometry based Hydrophobicity Scale for Amino Acids, **submitted for review**.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel genutzt habe. Alle wörtlich oder inhaltlich übernommenen Stellen habe ich als solche gekennzeichnet.
Ich versichere außerdem, dass ich die vorliegende Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und, dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

Berlin, 19. Juli 2018 _____