Investigation of Protein-Ligand Complexes by Native Mass Spectrometry and Ion Mobility-Mass Spectrometry

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

Melanie Göth

from Amberg, Germany

2017

The work reported in this thesis was performed from September 2013 to October 2017, 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:	7.12.2017

Danksagung

Viele Leute haben mich in den letzten vier Jahren auf meinem Weg begleitet und in unterschiedlichster Form dazu beigetragen, dass ich diese Arbeit erstellen konnte. Zuallererst möchte ich mich bei meinem Erstgutachter und Betreuer Prof. Dr. Kevin Pagel bedanken. Danke, für die Aufnahme in deine Gruppe und für dieses sehr vielseitige und spannende Thema. Ich bin sehr dankbar für die Freiheiten, die du mir für meine Forschung gegeben hast, aber auch gleichzeitig für die Betreuung und Ratschläge, die immer genau zur richtigen Zeit kamen. Danke auch für die Möglichkeit, dass ich zahlreiche Konferenzen, sowohl im In- als auch im Ausland, besuchen durfte und somit meinen Erfahrungsschatz erweitern konnte.

Weiterhin möchte ich Prof. Dr. Christoph Schalley danken für die Übernahme des Zweitgutachters meiner Dissertation.

Natürlich wäre die heutige Arbeit nicht ohne die nötige Finanzierung möglich gewesen. Aus diesem Grund möchte ich mich bei der Bayer AG für die Finanzierung meiner Doktorandenstelle bedanken und bei allen Leuten, die daran beteiligt waren: Dr. Anke Müller-Fahrnow und Dr. Andreas Becker, die dieses Kooperationsprojekt ermöglicht haben. Dr. Volker Badock, mein direkter Ansprechpartner und Betreuer bei Bayer, der alle Hebel in Bewegung gesetzt hat, damit ich ein paar Monate im Rahmen eines Projekts direkt in den Laboren von Bayer forschen durfte. Danke für die Möglichkeit, einmal in die tägliche Routine der pharmazeutischen Industrie zu schnuppern. Ich fand die Zeit bei euch sehr spannend und habe viel gelernt! Danke auch für die schnellen Antworten und Ratschläge, das mehrmalige Korrekturlesen und natürlich auch für das ein oder andere Treffen im Biergarten, bei denen wir wissenschaftliche und nicht-wissenschaftliche Themen über einem Bier besprochen haben. Ein großes Dankeschön geht auch an Dr. Benno Kuropka, den ich zunächst also Postdoc bei Bayer kennengelernt habe. Es hat mir sehr gut gefallen mit dir zu arbeiten, danke für deine Geduld und ebenso danke für nette Mittagspausen, sei es bei Bayer oder jetzt an der FU. Danke auch, dass du so engagiert mit unserem Paper warst und auch für das Korrekturlesen meiner Arbeit. An dieser Stelle möchte ich auch Dr. Jörg Weiske, Dr. Hanna Meyer und der kompletten Massenspektrometrieabteilung danken für die freundliche Aufnahme, die netten Gespräche und für eure Zeit, Rat und Tat.

Als nächstes möchte ich dem Fritz-Haber-Institut und allen Mitgliedern der Molekülphysik, insbesondere Prof. Gerard Meijer und Prof. Dr. Gert von Helden danken für die Finanzierung meines ersten Promotionsjahres und dass ich auch danach noch als Gast bleiben durfte und meinen wunderbaren Schreibtisch behalten durfte. Ein großes Dankeschön geht an alle Mitglieder der Pagel und von Helden Gruppe. Ihr habt diese Zeit zu einer unvergesslichen

Schönen gemacht! Die freundschaftliche Atmosphäre, unser guter Zusammenhalt und der wissenschaftliche Input, aber auch abendliche Sitzungen auf der MP-Terrasse und unsere gemeinsamen Skikurztrips oder sonstige Unternehmungen haben unter anderem dazu beigetragen, dass ich mich von Anfang an sofort wohl gefühlt habe. Besonders möchte ich auch meinen „Office Buddies“ Jakob, Leo und Frank danken für die lustigen, die Aha- und auch die stillen Momente. Danke Johanna, beste Konferenzzimmerteilpartnerin überhaupt! Ich könnte jetzt weiter seitenweise zu Jedem von euch etwas schreiben, warum ich euch dankbar bin und froh bin euch kennen gelernt zu haben. Leider sprengt das den Rahmen und deswegen sag ich es euch einfach beim nächsten Bier persönlich. Tausend Dank auf jeden Fall für eure Unterstützung während der letzten Monate! Weiterhin möchte ich allen meinen Bachelorstudenten und Forschungspraktikanten für euren Fleiß und eure gute Arbeit danken. Danke auch an die Core-Facility der FU, insbesondere an Dr. Andreas Springer und Fabian Klautzsch fürs Ultima-sitten, fürs Troubleshooting und auch sonst für die netten Pläuschchen zwischendurch.

Während meiner Zeit als Doktorandin durfte ich mit vielen hervorragenden Wissenschaftlern zusammenarbeiten: Thank you, Dr. Frederik Lermyte and Prof. Dr. Frank Sobott for interesting insights in ETD-MS and a (culinary) guided tour through Antwerp. Dankeschön, Dr. Oren Moscovitz for infinite hours of measuring and learning that some proteins are just as hard as stones and won't break, not even when you put them *under pressure* ;). Measuring with you was so much fun (although a lot did not work) and who knows, someday we have our natce-paper together (*don't stop believing*). Thank you for motivating me and being a friend. Danke Dr. Olaf Jahn, dass du Kevin und mich in die Calmodulin-Familie aufgenommen hast und uns damit die Tür zu einem sehr spannenden Projekt und einer tollen Zusammenarbeit geöffnet hast.

Zuletzt möchte ich mich bei meiner Familie und meinen Freunden für eure bedingungslose Unterstützung bedanken. Ich bin so froh, dass ich euch hab und sehr sehr dankbar, sei es für Finanzspritzen während der Studienzeit, für offene Ohren und gutes Zureden oder einfach ab und zu eine Portion Ablenkung. Ich hätte das alles definitiv nicht ohne euch geschafft! Danke Peter, für deine Liebe, dein Verständnis und, dass du für mich da bist.

Kurzzusammenfassung

Intermolekulare Wechselwirkungen von Proteinen untereinander oder mit anderen Molekülen, nehmen eine Schlüsselrolle in allen Prozessen lebender Organismen ein. Aus diesem Grund stellen Proteine wichtige therapeutische Angriffspunkte (Targets) dar und ihre strukturelle Aufklärung ist essenziell für die Entwicklung neuer Wirkstoffe. Native Massenspektrometrie (MS) ist ein attraktives Werkzeug, um Proteine und deren Komplexe zu untersuchen. Moleküle werden sanft ionisiert und aus der Lösung in die Gasphase überführt mit dem Ziel, inter- und intramolekulare Wechselwirkungen und die dreidimensionale Struktur aufrechtzuerhalten. Die ionischen Spezies werden durch ihr Masse-zu-Ladungsverhältnis detektiert, was zu Informationen über Masse, Ladung und Stöchiometrie führt. Native MS ist mit anderen Gasphasentechniken, wie Ionenmobilitätsspektrometrie (IMS), gut kompatibel und die Kombination beider Methoden zu Ionenmobilitätsmassenspektrometrie (IM-MS) liefert weitere strukturelle Informationen über die Größe und Form eines Moleküls.

In dieser Arbeit wurden Protein-Ligandkomplexe mittels nativer MS und nativer IM-MS mit dem Ziel untersucht, sie hinsichtlich ihres Potentials für die Hochdurchsatzanalyse der Wirkstoffentwicklung zu bewerten. Zuerst wurde native MS eingesetzt, um den Einfluss des Standardlösungsmittels Dimethylsulfoxid (DMSO) auf die Proteingasphasenstruktur, sowie die Protein-Ligandaffinität zu analysieren. Es wurde gezeigt, dass sich je nach DMSO Konzentration in der Probe die Ladungszustandsverteilung und somit vermutlich auch die Gasphasenstruktur eines Proteins ändert. Weiterhin verringerte sich die Affinität des Liganden mit steigender DMSO Konzentration. In einer zweiten Studie wurde das Potential von nativer MS für die fragmentbasierte Wirkstoffentwicklung getestet. In diesem Ansatz werden kleine Molekülfragmente gegen ein Target gescreent, wodurch im Vergleich zum Standard-hochdurchsatzverfahren höhere Erfolgsraten bei gleichzeitig geringerer Probenanzahl möglich sind. Ergebnisse von vier getesteten Targetproteinen zeigten, dass native MS im Moment noch keinen großen Durchsatz erlaubt, jedoch wichtige Einblicke in Protein-Ligandkomplexe liefern kann, welche durch andere Methoden nicht ohne Weiteres zugänglich sind. In einer weiteren Studie wurde in Kombination mit IMS der Einfluss der unmittelbaren Proteinumgebung auf dessen Gasphasenstruktur untersucht. Dafür wurden Kronenether-moleküle nichtkovalent gebunden, um positiv geladene Seitenketten zu mikrosolvatisieren und sie dadurch vor einem Zurückfalten auf das Proteinrückgrat zu bewahren. Mit Hilfe von Tandem MS und IM-MS wurden Bindungsstellen der Kronenether identifiziert und darüber hinaus gezeigt, dass bestimmte Proteinseitenketten die Gasphasenstruktur auch ohne Bindung von Kronenether stabilisieren können. In der letzten Studie wurde IM-MS erfolgreich als Methode zum Konformationsscreening von Protein-Peptidkomplexen eingesetzt. Die

Komplexe zeigten nur geringe strukturelle Unterschiede, weshalb sie mittels kollisionsinduzierter Entfaltung und Dissoziation weiter untersucht und ihre Unterschiede in Stabilität und Entfaltungsverhalten analysiert wurden.

Zusammenfassend stellen native MS und IM-MS wertvolle Werkzeuge zur Charakterisierung von Protein-Ligandwechselwirkungen dar. Gegenwärtig sind die Methoden zwar noch nicht für eine große Probenanzahl geeignet, aber durch ständige Entwicklungen werden sie in naher Zukunft sicherlich eine bedeutendere Rolle im Prozess des Wirkstoffdesigns einnehmen können.

Abstract

Intermolecular interactions of proteins with each other or with other molecules play a key role in all processes in living organisms. Therefore, proteins represent important therapeutic targets and their structural elucidation is essential for the development of new drugs. Native mass spectrometry (MS) is an attractive tool for the investigation of proteins and their complexes. Molecules are ionized gently and transferred from solution into the gas phase with the aim to maintain inter- and intramolecular non-covalent interactions and the three-dimensional structure. By measuring the mass-to-charge ratio of the ionic species, information on mass, charge, and stoichiometry can be obtained. Native MS is readily compatible with other gas-phase techniques, such as ion mobility spectrometry (IMS) and the combination of both methods, so-called ion mobility-mass spectrometry (IM-MS), provides further structural information on the overall size and shape of a molecule.

In this thesis, protein-ligand complexes were investigated using native MS and IM-MS with the aim of evaluating their potential for application in high-throughput analysis for drug discovery. First, native MS was used to elucidate the influence of the standard solvent dimethyl sulfoxide (DMSO) on the protein gas-phase structure and protein-ligand affinity. It was shown that the protein charge-state distribution and likely the gas-phase structure is altered depending on the DMSO concentration. In addition, the protein-ligand affinity decreased with increasing DMSO levels. In a second study, the potential of native MS for fragment-based drug discovery was evaluated. This approach is based on the screening of small molecular fragments against a target and promises higher hit rates and smaller library sizes compared to standard high-throughput screening. Data on four protein systems showed that native MS currently presents a medium- to low-throughput method but can provide valuable insights into protein-ligand interactions that are inaccessible by other techniques. In combination with IMS, the influence of a protein's microenvironment on its gas-phase structure was investigated. To do so, crown-ether molecules were attached non-covalently to microsolvate positively charged protein side chains, preventing them from collapsing onto the protein backbone. Using tandem MS and IM-MS, the crown-ether binding sites were identified and it was shown that specific side chains stabilize the gas-phase structure even without crown-ether binding. In the last study, IM-MS was successfully tested as a tool for conformational screening of protein-peptide complexes. Only subtle structural differences between the complexes were observed, and a further investigation by collision-induced unfolding and collision-induced dissociation displayed differences in complex stability and unfolding behavior. In summary, native MS and IM-MS are valuable tools for the characterization of protein-ligand interactions. Currently, the methods are limited to a small number of samples, but ongoing developments promise a decisive role in drug discovery in the near future.

Contents

1 Introduction

Proteins represent the most versatile macromolecules in living systems. Constructed from the same set of only 20 amino acid building blocks, thousands of different proteins are formed, each folded into a distinct three-dimensional structure with a specific task. Proteins are involved in cell growth and degradation or in the cell-cell communication, but they can also catalyze reactions or provide immune protection. All these functions are usually based on molecular recognition, which implies the interaction of proteins with each other or with other molecules (carbohydrates, lipids, DNA).[1] However, when a protein does not properly fulfill its natural function any longer or when a protein's structure changes irreversibly, also known as protein misfolding, this can also contribute to the cause of severe diseases. The structural elucidation of proteins and their assemblies, as well as the characterization of protein-ligand interactions is the fundamental basis for a better understanding of biochemical pathways in our organism. Furthermore, this knowledge is essential for medicinal chemistry to treat diseases and develop novel therapeutics. In this context, one of the key steps during the early drug discovery process is the identification of small molecules that bind specifically to the target protein. This binding can induce conformational changes in the protein or prevent the natural ligand from entering the active site.

Today, a variety of biological methods is available for the structural investigation of proteins and their complexes. Each method has its strengths and weaknesses for which reason they are typically used in combination to increase confidence and the informational content. Techniques such as nuclear magnetic resonance (NMR) spectroscopy[2] and X-ray crystallography[3] provide high-resolution information and represent the gold standards in the field of drug discovery, while isothermal titration calorimetry (ITC)[4] and surface plasmon resonance (SPR)[5] are especially appreciated for the competence in determining binding affinities. Another method that is ideally suited for the analysis of non-covalent interactions is mass spectrometry (MS). Using this technique, the analytes of interest are ionized, transferred into the gas phase and subsequently analyzed based on their mass-to-charge (*m/z*) ratio. Thus, information on mass, charge, stoichiometry, and relative stability can be obtained with the advantage of using very low amounts of sample in a short time frame. Mass spectrometry-based methods have contributed to a great extent to the identification of the human proteome,[6] in pharmaceutical industry, mass spectrometry has been established for quality control and to follow pharmacodynamics or pharmacokinetics.[7] In recent years, native

MS,[8] a particular concept of mass spectrometry, attracted interest for application in drug discovery. By using neutral aqueous solutions and gentle conditions, non-covalent protein-ligand interactions are maintained upon transfer from solution into the gas phase, which allows a direct analysis of binding and relative binding affinities. Native MS can be readily combined with other gas-phase techniques, such as ion mobility spectrometry (IMS), to obtain further information. In IMS, ionic species are guided through a gas-filled cell and are separated according to charge, size, and shape. The overall size of the molecule can be estimated in form of a rotationally-averaged collision cross section. Thus, ion mobility-mass spectrometry (IM-MS) allows the simultaneous analysis of mass, charge, size, and shape. Furthermore, it enables to follow conformational changes and therefore presents a promising tool for the investigation of protein-ligand interactions.

Motivation and Outline of the Thesis

In this thesis, protein-ligand interactions are investigated using native MS and IM-MS with the aim of evaluating their potential for routine application in drug discovery. **Chapter 2** introduces the basic principles of the applied methods and important applications in the field. In addition, characteristics of the two utilized instruments are briefly described.

Besides the capability for high-throughput and automation, simple and fast sample handling is a key factor for a successful application in screening assays. Dimethyl sulfoxide (DMSO) has established as the standard solvent for compound storage and handling. Depending on the solubility of the molecules and the biological method that is used, the level of DMSO in the final samples can be comparatively high. In **Chapter 3**, the effect of DMSO on the protein gas-phase structure and on protein-ligand affinity is studied with native MS based on two protein domains and a high-affinity binding ligand. The results show a significantly altered charge-state distribution in the mass spectra with differing DMSO concentrations. Furthermore, a lower binding affinity is observed for increasing DMSO level. Possible mechanisms for this effect and a potential solution to replace DMSO are discussed.

To investigate protein-ligand interactions in detail and select the ligand with the highest binding affinity, it can be of great benefit to focus at first on the characterization of small systems. In recent years, fragment-based screening (FBS) has evoked as a promising alternative to standard high-throughput screening (HTS). In FBS, small fragments of a molecule up to a molecular size of ~300 Da are investigated on their interactions with proteins. The question whether native MS has the potential to be applied for FBS is evaluated in **Chapter 4**. To do so, small sets of systematically chosen fragments are screened against four proteins of different classes. MS results are compared to data from thermal shift assay (TSA), which presents a well-established method in screening campaigns. The results strongly indicate that native MS is not yet ready to be implemented in a high-throughput scaffold, but provides valuable information on protein-ligand interactions as a complementary method.

IM-MS likely presents an attractive tool for applications in pharmaceutical industry, as it cannot only provide information on the overall shape of a target protein, but also monitor conformational changes. In general, the gas-phase structure of a protein can be manipulated simply by attaching crown-ether molecules non-covalently to positively charged side chains of a protein. It was shown previously that this microsolvation prevents a collapse of the charged side chains onto the protein backbone. In **Chapter 5**, IM-MS is applied in combination with tandem MS to identify such crown-ether binding sites on the protein ubiquitin. The results show that specific side chains play a decisive role for a structural preservation in the gas phase. Residues that are involved in salt bridges stabilize the protein structure and do not bind favorably to crown ethers.

Upon ligand binding, a conformational rearrangement can be induced in the protein, which can be readily detected by IM-MS. In **Chapter 6**, IM-MS is employed successfully as a tool for conformational screening of protein-peptide complexes. The investigated complexes show only subtle structural differences, which is assigned to a collapse due to a flexible linker in the protein. In addition, unfolding and dissociation of the complexes are investigated to provide information on unfolded structures and on relative binding affinities.

Chapter 7 concludes the thesis and gives future prospects for native MS and IM-MS.

2 Fundamentals

2.1 Proteins and their Structural Elements

Proteins represent one of the most important classes of biomolecules in every organism. They are built from a repertoire of 20 different α-*L*-amino acid building blocks (Figure 2.1), which are connected to each other *via* covalent peptide bonds to a linear sequence. This sequence, referred to as the primary structure (Figure 2.1), determines predominantly how the protein backbone folds into a specific three-dimensional shape. In addition, protein folding is influenced by interactions between the peptide backbone and amino acid side chains as well as between the molecule and its environment. These interactions are based for example on ionic attractions, hydrogen bonds, van der Waals or hydrophobic forces and lead to the generation of further structural elements. A local three-dimensional organization of the peptide chain results in the formation of secondary structural elements, such as α-helices, β-sheets, and turns (Figure 2.1). These elements can further arrange to form the tertiary structure and multiple proteins or protein subunits assemble to the quaternary structure (Figure 2.1).[9] The three-dimensional protein structure is closely related to its function (structure-activity-relationship). Thus, to be able to interact specifically with natural binding partners, the protein has to adopt and maintain its *native* structure, which represents commonly the most thermodynamically stable structure under physiological conditions.[10]

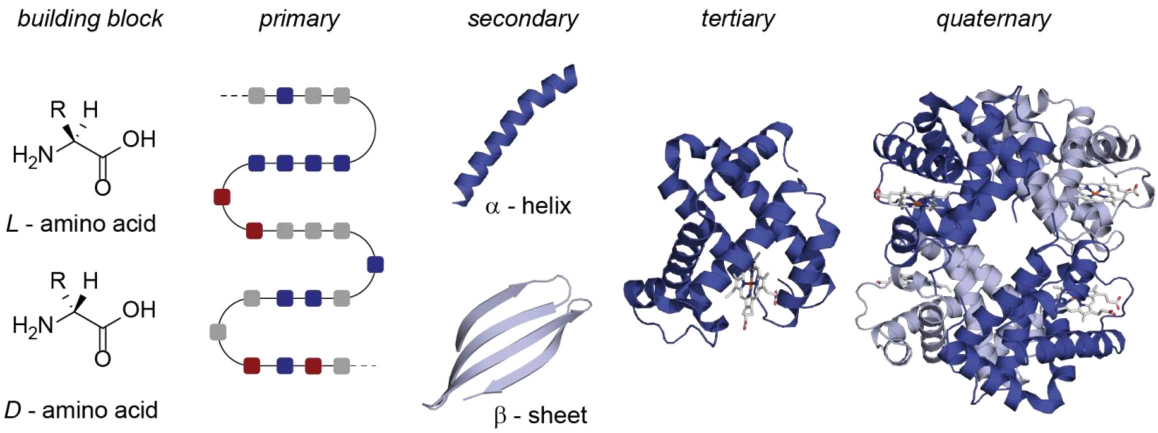


Figure 2.1: Classification of protein structure elements from left to right: structures of α-*L*-amino acid and α-*D*-amino acid with side chain R, primary structure, secondary structure, tertiary structure, and quaternary structure.

2.2 Native Mass Spectrometry

2.2.1 What is Native Mass Spectrometry?

In native mass spectrometry (MS) biomolecules are transferred gently into the gas phase using non-denaturing conditions. The aim is to maintain the tertiary and quaternary structure and to transfer non-covalent complexes intact. The term *native MS* was coined in 2004,[8] but first experiments have been performed already in the early 1990s.[11-14] The strictest definition of the native state of a protein implies the fully folded and biologically functional form in its natural environment. As mass spectrometry is a gas-phase method and biomolecules are investigated outside the cell in a high-vacuum environment, *native MS* seems impossible. Certainly, it cannot fully reflect the native state present *in vivo*, but only imitate biological conditions as best as possible. Using the term *native* in the context of this thesis thus refers to the biological status of the biomolecules in solution before ionization and detection, which is attempted to be maintained.[15]

Native MS does not yield high-resolution molecular or atomic structural information like NMR spectroscopy or X-ray crystallography. However, it can provide information on the composition, stoichiometry and on subunit assemblies of proteins and protein complexes (Figure 2.2). Combining these puzzle pieces, the topology of a protein complex can be reassembled. Furthermore, native MS allows the direct detection of ligand binding and *fishing* for the best binding compound (Figure 2.2D), which will be discussed in more detail in Chapter 2.4. In contrast to other methods, MS has the advantage of high sensitivity, selectivity, and speed, and a very low amount of sample (pmol to fmol range) is required for an experiment.[16] Another huge benefit is the ability to select and investigate one species in the presence of others. The different species can be assigned unambiguously by their mass-to-charge (m/z) ratio and thus no additional labeling is necessary.

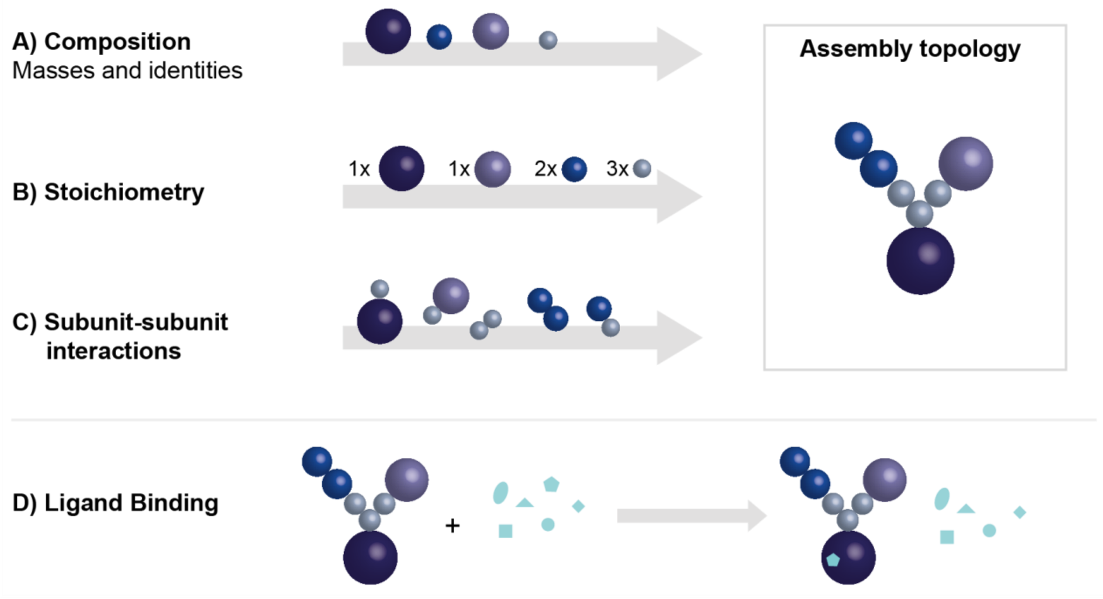


Figure 2.2: Schematic overview of the information native MS can provide. The topology of a protein assembly can be recomposed by using **A)** composition, **B)** stoichiometry, and **C)** information on subunit assemblies. **D)** Native MS can furthermore detect ligand binding and determine the best binding compound.

2.2.2 How to Make Large Proteins Fly - a Brief Historical Overview

One of the most important milestones that paved the way for the investigation of intact biomolecules in the gas phase was the introduction of two soft ionization methods, electrospray ionization (ESI)[17] and matrix-assisted laser desorption/ionization (MALDI),[18] in the 1980s. In MALDI the analyte is embedded in a matrix and irradiated with short UV and IR laser pulses, which leads predominantly to singly charged ions. MALDI has been applied successfully to preserve non-covalent complexes,[19, 20] however, there are still some critical factors to overcome. These are for example possible interactions with the matrix, the dissociation of weak non-covalent interactions during the ionization process, or the challenge to detect high-molecular weight ions. For these reasons ESI is still the preferred ionization technique in the field.[21]

With ESI usually multiply charged ions are formed. Starting from atmospheric pressure, an analyte solution is injected into a metal capillary and an electric potential is applied. This potential is positive *versus* ground for positive ion mode and vice versa for negative ion mode.[22, 23] In the following, the procedure in the positive ion mode will be discussed as it is more commonly used for peptides and proteins.

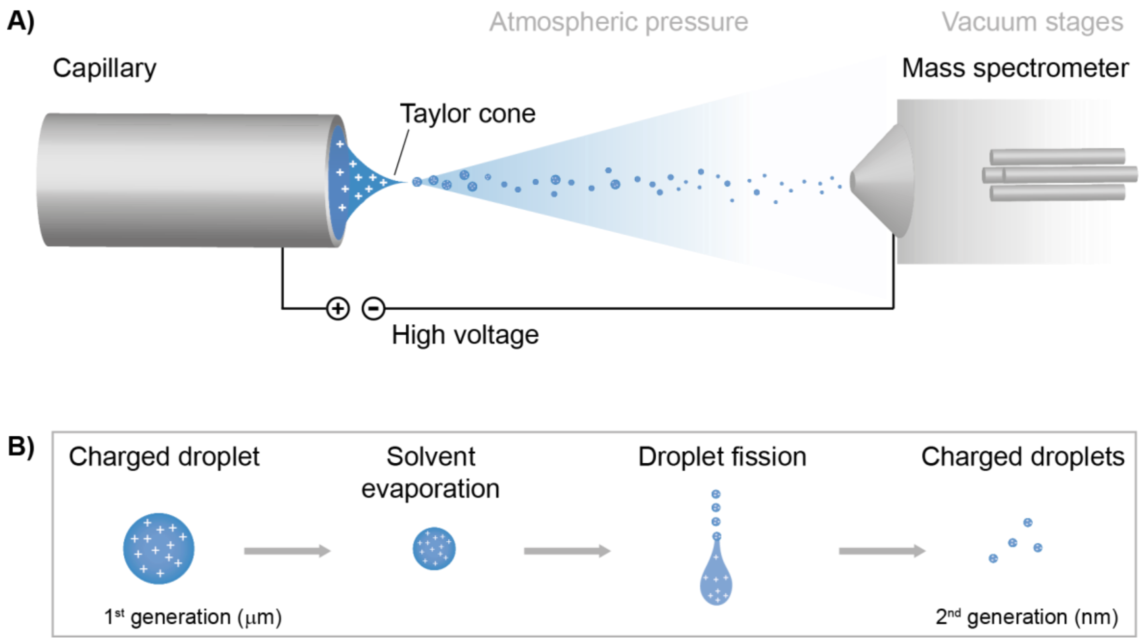


Figure 2.3: A) Schematic depiction of the electrospray ionization process. **B)** Schematic representation of the desolvation process.

A schematic representation of the ESI mechanism is depicted in Figure 2.3A. Due to the electric potential, positively charged analytes are enriched at the capillary tip and a so-called Taylor cone is formed. From this cone a fine jet of initial ESI droplets filled with the analytes is emitted with radii in the µm range (Figure 2.3B). The droplets undergo rapid solvent evaporation and shrink until the Rayleigh limit is reached.[24] This limit describes the threshold where the surface tension of the droplet is able to compensate Coulomb repulsion and the droplet is still stable. Droplets at the Rayleigh limit produce even smaller and highly charged offspring droplets *via* jet fission. The repeated evaporation yields ESI droplets with radii of a few nm.

Three different mechanisms are accepted for the ion release and they are schematically represented in Figure 2.4: (A) Small molecules are released by the ion evaporation model (IEM).[25] In this model the charged ion travels to the droplet surface and departs in form of a small gas-phase cluster, which consists of the ion and a few solvent molecules. The remaining solvation shell is lost when traversing the sample interface and colliding with background gas.[26] (B) Large globular species, like natively folded proteins desolvate *via* the charged residue model (CRM).[22] Here it is assumed that the generated nanodroplets contain a single analyte molecule and *dry out* with the charge of the droplet being transferred to the analyte.[22, 23]

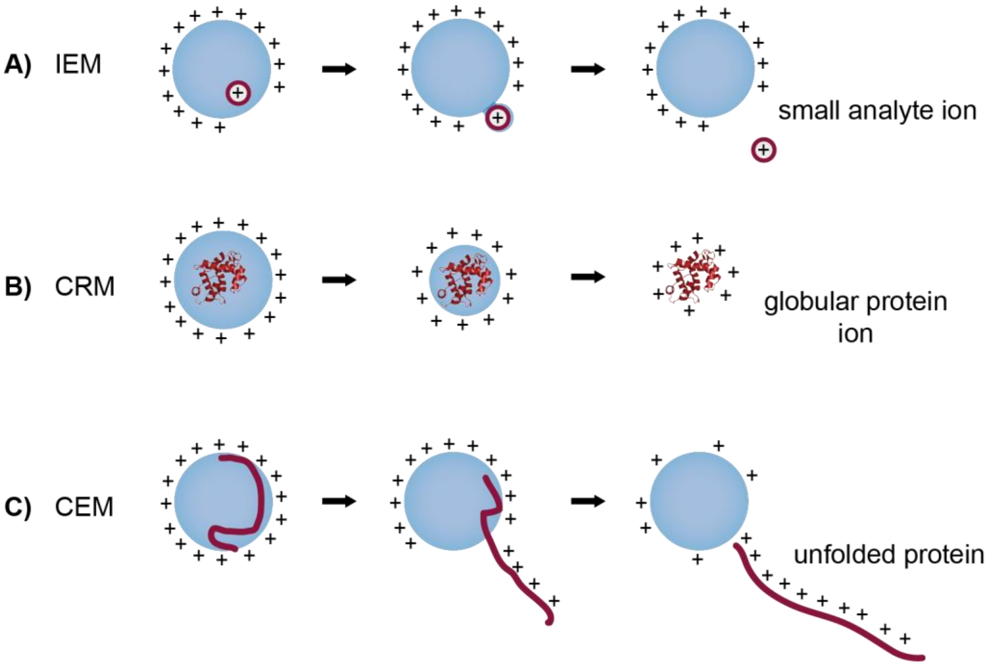


Figure 2.4: Schematic representation of three different ion release mechanisms in ESI: **A)** ion evaporation model (IEM), **B)** charged residue model (CRM), and **C)** chain ejection model (CEM).

(C) In case of unfolded proteins, the ion is assumed to be released by the chain ejection model (CEM).[27] While natively folded proteins are compact and hydrophobic residues are embedded in the core, unfolded proteins adopt an extended structure with solvent-accessible hydrophobic side chains. Due to the large hydrophobic character, it is not favorable for the unfolded protein to stay in the droplet. Thus, it migrates to the droplet surface and a stepwise ejection occurs.

A miniaturized version of ESI is nanoflow ESI (nanoESI, nESI).[28] In contrast to conventional ESI, where tip openings are ~100 µm, nESI uses metal-coated borosilicate capillaries with smaller tip openings of around 1-5 µm. This results in smaller initial droplets, significantly reduced flow rates and reduced sample consumption as well as increased sensitivity and enhanced ionization efficiency.[23, 28] Moreover, nESI is usually more tolerable to nonvolatile salts and buffers.[29] With the commercial introduction of the TriVersa NanoMate (Advion®) in 2005, the automatization of nESI was promoted, which facilitates applications in high throughput screening. This chip-based nano electrospray ionization technology with a robotic inlet system allows fully automated direct infusion of individual samples and can also be coupled to nano liquid chromatography.[30]

Not only the introduction of soft ionization methods, but also major developments in mass analyzers led to the high success and exorbitantly increasing applications of native mass spectrometry for the analysis of biomolecules. To date the most commonly used combination of mass analyzers in native MS is quadrupole time-of-flight (Q-ToF) mass spectrometry.[31] While time-of-flight mass analyzers have theoretically no upper mass limit and stand out by a high sensitivity and high resolution on a fast timescale,[32] conventional quadrupoles allow an acquisition up to only m/z 2000 - 4000. This is however too low for measurements of intact large biomolecules. With quadrupoles running at a lower radiofrequency or different dimensions, this problem can be overcome and the m/z-selection range can be increased up to m/z 32,000.[33, 34] Moreover, in contrast to small molecules, higher pressures and voltages have to be applied to keep large biomolecules focused on their trajectory (collisional focusing or collisional cooling) and to guarantee successful transmission. This is enabled by a pressure increase in several pumping stages of the mass spectrometer.[35] The Q-ToF technology has been used with great success on many protein complexes with high m/z ions,[36, 37] and molecular weights of even several million Daltons.[38] Alternative mass analyzers with high sensitivity and high resolution are the Fourier transform-ion cyclotron resonance (FT-ICR)[39, 40] analyzer and the Orbitrap mass analyzer.[41] Especially the latter has gained growing attention, since a modified version allows the transmission of high m/z ratios and measurements with a significantly higher resolving power compared to conventional Q-ToF instruments can be performed.[15] A high resolution in the high mass range has led to some excellent studies on protein complexes, where small differences could be mapped even in megadalton assemblies.[42, 43]

2.2.3 Protein Structure in the Gas Phase

In the beginning of this chapter the term native MS was introduced and in this regard it was pointed out that *native* refers to the state of the biomolecules in solution before ionization and detection. The question to what extent native structure can be retained in the gas phase led to a fierce debate between experts in the field, which is partly still ongoing today.[44, 45]

In this context, the charge-state distribution in the mass spectrum can provide a first indication on the structural preservation. It is accepted that folded proteins adopt compact structures and accommodate fewer charges, while unfolded species show extended structures and can be highly charged.[46] Therefore, proteins with a compact or even native-like structure appear with a narrow charge-state distribution at lower charge in the mass spectrum, whereas significantly higher charges and broader charge-state distributions are typically formed from solutions in which the protein is denatured.[47-50] It is important to mention that the choice of solvent already plays a decisive role, as it has an influence on the protein structure in solution and thus also on the resulting charge-state distribution in the mass spectrum. In conventional ESI experiments, the sample is sprayed from an acidified organic solvent. Figure 2.5 (top) exemplarily shows the nESI mass spectrum of the muscle protein myoglobin from an aqueous solution containing 50 % methanol (MeOH) and 1 % formic acid (FA). Under these conditions the protein receives numerous positive charges with two broad charge-state distributions centered on 17+ and 9+. The experimentally determined molecular weight of the main species of 16,937 Da identifies the protein as the apo-form (non-heme bound myoglobin) and the free heme unit is detected at *m/z* 654. Besides, holo-myoglobin is also formed with low intensity under these conditions. Instead of organic solvents, aqueous buffer solutions at neutral pH are employed in native MS. Unfortunately, the buffers that are usually used for purification and stabilization of proteins, such as tris(hydroxymethyl)-aminomethan (TRIS) or 2-(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) cannot be used, because they contain nonvolatile salts and interfere with ESI. A solvent that proved to be well compatible with ESI is aqueous ammonium acetate (10 mM to 1 M), as it is volatile and therefore desolvates easily during the ionization process.[23] When myoglobin was measured from an ammonium acetate solution at neutral pH, the nESI mass spectrum in the positive ion mode shows a narrow distribution of relatively low charge states centered around 8+ (Figure 2.5 bottom). In this case, the heme unit is still bound to myoglobin (17,548 Da) and the apo-form is not detected. Under non-denaturing conditions the protein is substantially less charged than in organic solvents, which can be attributed to a more compact surface in folded structures. Furthermore, this example illustrates nicely that non-covalent protein-ligand interactions can be maintained in the gas phase under non-denaturing conditions, whereas in organic solvents these interactions are likely to be disrupted.

Besides the charge-state distribution, another indication for an intact structure is the fully functional protein. An excellent example is presented by an early study on the tobacco mosaic virus, where it was shown that the functional structure can be retained in the gas phase.[51] Virus particles were electrosprayed from solution and subsequently captured and visualized by electron microscopy. Fascinatingly, the virus was still viable and infected tobacco plants after its passage through the mass spectrometer. This strongly indicates that the virus capsid structure is preserved in the gas phase.

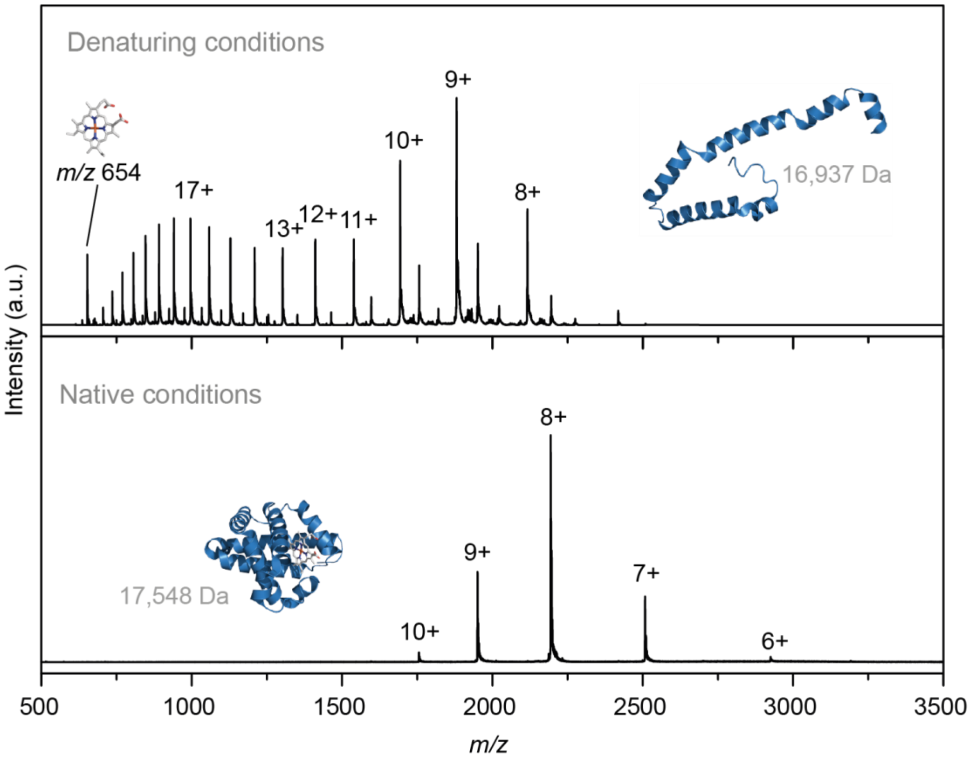


Figure 2.5: nESI mass spectra of myoglobin sprayed from (top) a MeOH: H_2O (50:49) solution with 1 % formic acid and sprayed from (bottom) 10 mM aqueous ammonium acetate solution at neutral pH. Binding of the heme unit is observed when sprayed from ammonium acetate, which indicates to a native-like protein structure in the gas phase. Sprayed from acidified organic solvent the apo-protein is the primarily detected species. PDB entry: 4FWZ was used for the exemplary protein structure of myoglobin.

However, in a different study on green fluorescent protein (GFP) the functional structure could not be preserved.[52] Although the mass spectrum showed a narrow charge-state distribution and a large stability against dissociation, fluorescence emission was not observed for gas-phase GFP. This study exemplifies that the charge-state distribution is not always a reliable criterion and that native MS data has to be interpreted carefully.[53]

Furthermore, it should be considered that changing the environment from solution to the gas phase also entails alternating binding forces. In the absence of water hydrophobic contributions are substantially weakened or get lost, whereas ionic attractions and van der Waals interactions are strengthened.[54] Depending on the nature of the proteins, reorganization or even unfolding might be the consequence.[44, 45] However, ionic interactions such as salt bridges can also stabilize the native fold in the gas phase.[55] For the majority of proteins it can be assumed that at least structural motifs are retained in the gas phase. This is strongly indicated by several studies over the last decade on various different biomolecules such as whole virus particles,[38] the proteasome,[56, 57] membrane proteins,[58] or very complex molecular machines like the V-type ATP synthase.[59]

Up to date, proteins and protein complexes are the most commonly investigated analytes by native MS.[15] To further increase the information content, native MS can be readily combined with bioinformatics[60] or other structural methods, such as hydrogen/deuterium (H/D) exchange,[61] cross-linking,[62] infrared spectroscopy[63] or ion mobility spectrometry.[64] Using these hyphenated methods usually adds another, orthogonal, dimension to *m/z* and thus can give further insights on structural or mechanistic details.

2.3 Ion Mobility-Mass Spectrometry

2.3.1 Development and Applications

Ion mobility spectrometry (IMS) is an analytical technique in which ions are separated based on their mobility in a gaseous medium. The basic principles were already described more than one century ago when the French scientist Langevin studied the motion of ions in an electric field.[65, 66] Since then substantial technical developments have been achieved and during several decades IMS has been used for a wide range of applications. It plays for example a prominent role for military and security purposes, where it is employed to detect chemical weapons, explosives or illegal drugs.[67, 68] In industry, IMS is used as quality control and monitoring device to detect hazardous side products.[66] Moreover, it is utilized in environmental research for air quality analysis and the detection of microbiological organisms.[69] IMS does not only allow studying the structure of gas-phase ions, but it also enables the separation of structural isomers and therefore it can be used additionally as a gas-phase separation method.[66]

In recent years, IMS systems found an increasing demand for new applications of complex samples like biological tissues in medicine. To improve the analysis of these challenging samples, hyphenated methods are used. Ion mobility-mass spectrometry (IM-MS), which combines IMS with MS, emerged as a valuable tool for many different applications over the last decades. While mass spectrometry separates ions based on their mass-to-charge (*m/z*) ratio, the separation in IMS is based on the overall size and shape. As these two aspects of separation occur on different timescales, the two methods are well compatible for an orthogonal analysis in a single instrument.[70]

2.3.2 General Principles

A typical IM-MS experiment starts with the ionization of the analyte for example *via* nano electrospray ionization. Ion packages are then pulsed into the ion mobility cell, which is filled with an inert buffer gas, such as helium or nitrogen. Guided by a weak electric field, the ions traverse the cell and undergo collisions with the buffer gas. Thereby, ions are separated according to their mobility K, which is dependent on charge, size, and shape of the species. This principle is particularly helpful when analyzing molecules with the same mass (*e.g.* isomers) that cannot be differentiated by the m/z ratio. Larger ions collide more often with the gas atoms and have a lower mobility than smaller, more compact, ions (Figure 2.6A). Furthermore, highly charged ions travel faster through the cell than lower charged ions (Figure 2.6B). The velocity v of an ion is the product of ion mobility K and the electric field E (Eq. 2.1). The drift time t_D that a specific ion species needs to travel through the ion mobility cell with a known dimension d is measured and can also be used to determine the velocity of an ion.[71]

$$v = KE = \frac{d}{t_D} \tag{2.1}$$

The measured drift time is instrument-dependent and therefore hardly comparable between different instruments. Using a classical drift-tube IMS setup, the drift time of a particular ion can be further converted into a rotationally-averaged collision cross section Ω (CCS). The CCS corresponds to the area that collides with the drift gas. It is a molecular property that is instrument-independent and thus can be compared to theoretical values obtained from X-ray crystallography, NMR spectroscopy, or model structures. To calculate the CCS, the mobility K is normalized against temperature T and pressure P, which results in the reduced mobility K_0 (Eq. 2.2).

$$K_0 = \frac{d}{Et_D} \frac{273}{T} \frac{P}{760} \tag{2.2}$$

Subsequently, the reduced mobility K_0 is inserted in the Mason-Schamp equation[72] (Eq. 2.3) to obtain the collision cross section Ω (typically in Å^2 or nm^2):

$$\Omega = \frac{3\,ze}{16\,N} \left(\frac{2\,\pi}{\mu\,k_B T}\right)^{1/2} \frac{1}{K_0} \tag{2.3}$$

with ze being the ion's charge, N the density number of gas, μ the reduced mass of the ion and buffer gas, and k_B the Boltzmann constant.

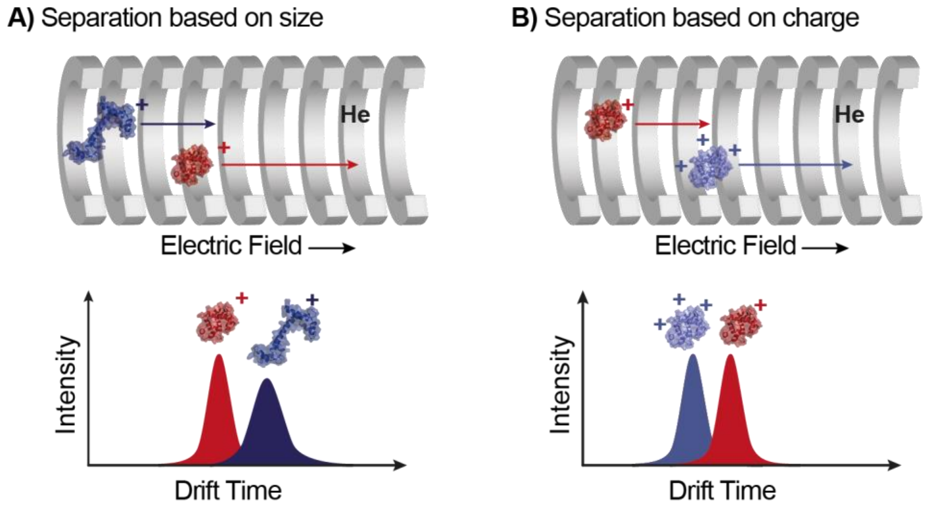


Figure 2.6: Principle of an ion mobility separation (IMS). The ion mobility cell is filled with an inert buffer gas, such as helium, and ions are guided through the cell by a weak electric field. Ions of the same *m/z* can be separated according to their mobility, which is dependent on **A)** the size, shape, and **B)** charge of the ionic species. The drift time of each ion package is recorded and plotted for all analyzed species as an arrival time distribution (ATD).

Experimentally determined CCSs can also be compared with theoretically predicted values. The three approaches that are most commonly used to obtain theoretical CCSs from three-dimensional model structures are the trajectory method, the projection approximation (PA), and the exact hard sphere scattering (EHSS) method. Details of the individual approaches can be found in the literature.[71, 73, 74] Briefly, PA is the most reduced and computationally cheapest method, but it is not necessarily suitable for large biomolecules. EHSS is more advanced compared to PA. In general, PA often underestimates CCS values of native protein structures, while EHSS overestimates them. The most reliable results can be obtained using the trajectory method, but it is also the most expensive and elaborate computational approach. In all three methods helium is used as standard gas for the calculations, which should be considered when comparing with experimental values. In general, nitrogen CCS values are higher in magnitude compared to helium CCS values.

2.3.3 Instrument Types

In the following section, different ion mobility types are briefly discussed and a schematic representation of the characteristics is illustrated in Figure 2.7. Drift tube ion mobility spectrometry (DT-IMS, Figure 2.7A) is the oldest and simplest form of ion mobility and its basic principles were explained in detail in the previous section. Together with traveling wave ion mobility spectrometry (TW-IMS) it presents one of the two most important and most frequently used types of IMS. In DT-IMS, a homogeneous, linear electric field is used, which is generated for example by a stack of ring electrodes.[75] The ions are pulsed in packages into the drift cell and move against the gas flow. Due to the uniform electric field, the mobility of the ions is inverse proportional to Ω and thus absolute CCS values can be determined directly from the measured drift times and the applied experimental conditions (Eq. 2.3). The IMS resolution can be improved with the length of the drift cell. In this regard, drift tube lengths up to three meter have been reported[76] and even a cyclic layout is possible.[77]

In traveling wave ion mobility mass spectrometry (TW-IMS, Figure 2.7B) the ion mobility cell consists of a series of stacked ring electrodes. It uses traversing pulses that guide the ions through the ion mobility cell. As a result of this inhomogeneous electric field, the ions follow complex trajectories during their migration through the IMS cell.[78] They virtually *surf* on the traveling wave, and ions with higher mobility are carried further, while ions with lower mobility roll over the wave. The velocity of the wave as well as the amplitude of the voltage pulse can be adjusted to optimize separation and ion transmission.[66, 78] This leads to a good resolution and duty cycle, but does not allow the determination of absolute CCS values. Instead, CCS values can be estimated using species with similar physicochemical properties and known CCSs as calibrants.[79] While such a calibration is relatively straightforward in the positive ion mode by now, it is still challenging for larger biomolecules in the negative ion mode, as suitable calibrants are often not available. With the introduction of the commercially available Synapt HDMS in 2006,[80] TWIMS has gained high popularity entailed by a great and still increasing number of applications.

Another form of ion mobility is high-field asymmetric waveform ion mobility spectrometry (FAIMS, Figure 2.7C).[81] In this technique, an asymmetric waveform is applied between two electrodes. The ions move with the gas flow through the cell and experience two alternating electric fields, which are applied perpendicular to the flow. The mobility of the ions is different depending on whether they are exposed to high or low electric field strength. This entails a radial movement of the ions with the risk of neutralization at the electrode walls.

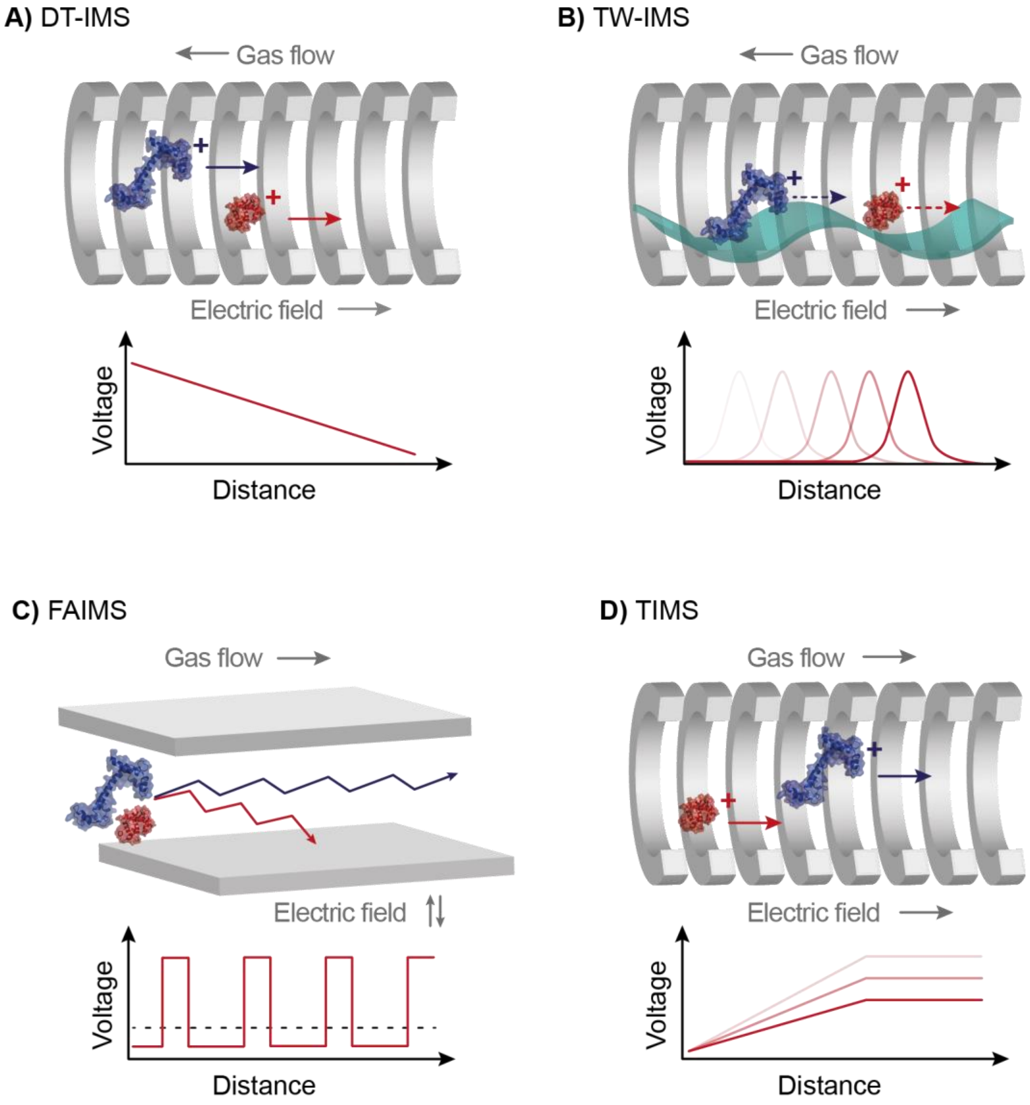


Figure 2.7: Schematic representation of the characteristics of different ion mobility types: **A)** Drift tube IMS (DT-IMS), **B)** Traveling wave IMS (TW-IMS), **C)** High field asymmetric waveform IMS (FAIMS), and **D)** Trapped IMS (TIMS). For each type the characteristic motion of the ions in the cell and the corresponding potential gradient diagram are shown.

To refocus the ions on their pathway, a compensation voltage is applied, which differs for ions of different charge and size. Instead of the drift time, the compensation voltage for a given species is generally reported. With FAIMS only one species can be analyzed at a time and CCSs cannot be determined. Furthermore, under high-field conditions, proteins can be activated and undergo conformational changes. Thus, it is primarily used as a filter technique to dispose of unwanted species and increase the sensitivity of the ions of interest.[82]

Another very recent development is trapped IMS (TIMS).[83] Here, the ions are trapped in the ion mobility cell by an electric field and separated according to their size-to-charge ratio.

Subsequently, the ions are eluted as packages by manipulation of the electric field. This leads to a very high IMS resolution and efficient separation. As the ion mobility resolution can be adjusted, the duty cycle can be increased by up to 100 %. CCS values can be estimated with a calibration procedure.

2.3.4 Ion Mobility-Mass Spectrometry of Proteins

After the successful coupling of IMS and MS in the early 1960s,[84] primarily atoms and small molecules, such as gases or metal ions, were studied. The first peptides and small proteins were investigated at the end of the 1990s,[85] however, it still took a few more years until larger proteins and their assemblies could be analyzed successfully. One of the first studies of large intact complexes revealed that the 11-membered ring assembly of an RNA-binding protein is mostly maintained.[86] This work provided one of the first pieces of direct evidence that IM-MS can contribute essentially to elucidate whether native structure can be retained in the gas phase. When measuring the drift time of a protein through the ion mobility cell, the overall shape of each charge state in the mass spectrum can be displayed in form of an arrival time distribution (ATD). Figure 2.8 shows the ATDs for charge states 4+ to 9+ exemplarily for the protein ubiquitin as a function of collision cross section. It can be observed that low charge states (4+) adopt more compact structures with smaller CCS values, whereas species in higher charge states (8+, 9+) show more extended structures, which is often attributed to Coulomb repulsion within the molecule. Intermediate charge states (~5+ to 7+) can adopt a multitude of conformations, as reflected in very broad ATDs with several features. From these findings it can be assumed that lowly charged species resemble a native-like structure of a protein. In the case of ubiquitin a comparison of IM-MS data and solution phase data revealed that ions with a charge <7+ are of similar size as the native fold of the protein.[87, 88]

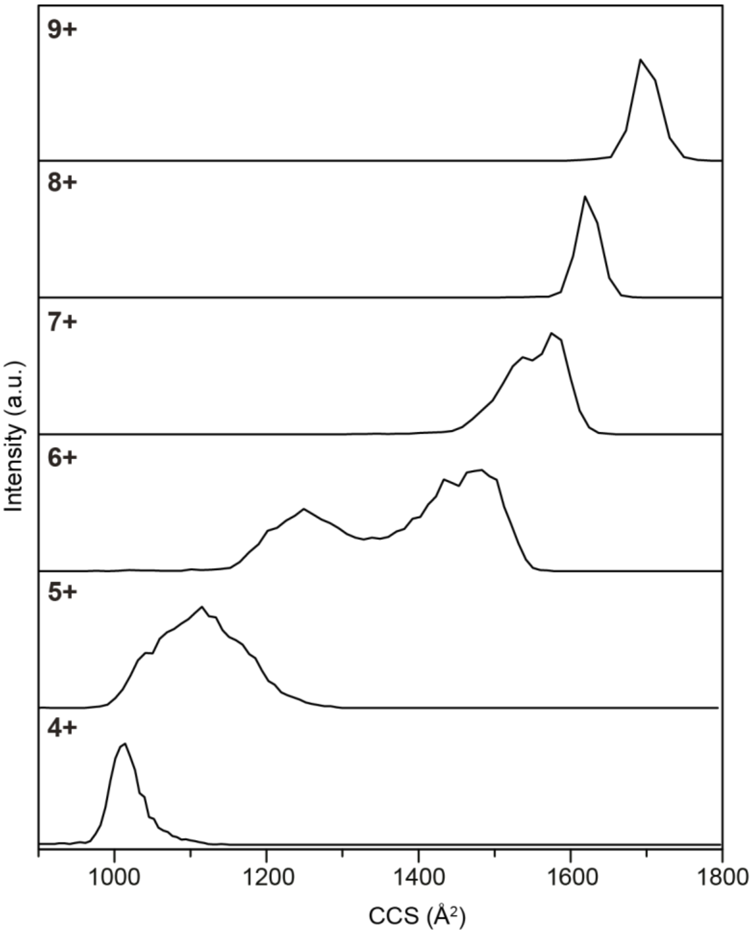


Figure 2.8: Measured arrival time distributions (ATDs) for the protein ubiquitin in different charge states (from 4+ to 9+). The x-axis was transformed from drift time to absolute CCS in Å². The figure illustrates that the CCS increases for higher charge states. Moreover, lower charge states (4+) and higher charge states (8+, 9+) appear in narrow ATDs with only one conformational family, whereas intermediate charge states (~5+ to 7+) can adopt a multitude of conformations, which is reflected in a broad ATD.

2.4 Protein-Ligand Binding

2.4.1 Complex Stoichiometry and Ligand Affinity

To fulfill their roles, proteins interact with each other or with other molecules in specific complexes. Such molecules, referred to as ligands, are for example peptides, nucleic acids, carbohydrates or lipids, which bind with a high specificity and affinity to the target. Binding is typically reversible and occurs through non-covalent intermolecular forces, such as hydrogen bonds, ionic attractions or hydrophobic forces. In turn, irreversible covalent binding is rather uncommon in biological systems.[89]

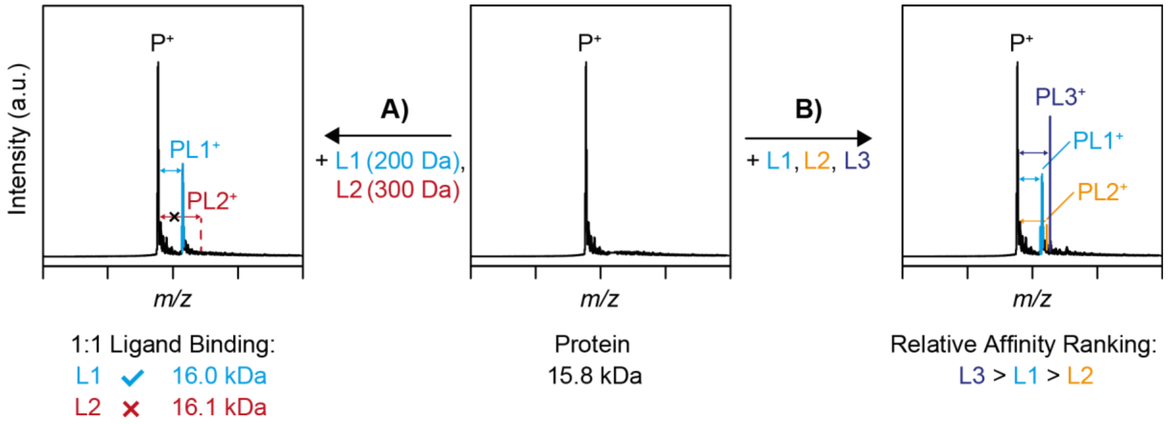


Figure 2.9: Deconvoluted mass spectra of an exemplary protein P. **A)** Binding of ligand L1 (blue) to P is detected by a discrete shift in mass (blue arrow). Ligand L2 (red) does not bind to P. **B)** Three ligands L1 (blue), L2 (orange), and L3 (dark blue) bind to protein P with different affinity, which is reflected in the corresponding intensities of detected PL complexes in the mass spectrum.

The characterization of protein-ligand interactions and the elucidation of their underlying mechanisms are essential for a deeper understanding of molecular recognition processes. However, this knowledge is also important with respect to the design and development of novel therapeutics. In this context, native MS provides some advantages, as several ligands can be measured simultaneously in a short time scale, without the need for labeling, and very low amount of sample is required. Binding is detected by a discrete shift in mass. By determining the accurate molecular weight of the protein-ligand complexes, the attached ligand and the complex stoichiometry can be identified. This is shown in Figure 2.9A with an exemplary protein P (15.8 kDa). While ligand L1 (blue, 200 Da) binds to P and forms a 1:1 complex PL1 (16.0 kDa), binding with ligand L2 is not observed. In addition to stoichiometry and ligand identity, information on the affinity can be obtained with native MS. This is realized qualitatively by competition experiments, where the sample contains the target protein and several ligands.[90] The resulting mass spectra show the formation of complexes with the different ligands, (ideally) deviating in their intensities as it is illustrated in Figure 2.9B with protein P and the ligands L1, L2 and L3. This gives a qualitative ranking, in which the best binding ligand is identified (L3) based on the corresponding intensities of detected PL complexes.

Protein-ligand interactions can also be quantified by determining the association or dissociation constant (K_a, K_d). In general, it is assumed that the equilibrium between the free protein P and protein-ligand complex PL in solution is also reflected in the gas phase (Eq. 2.4). Therefore, the ratio R of the total intensities I(PL)/I(P) is equivalent to the ratio of the concentrations of $[PL]_{eq}/[P]_{eq}$ in solution at equilibrium (Eq. 2.5) and can be used to quantify the binding strength (Eq. 2.6).[91]

$$P + L \rightleftharpoons PL \tag{2.4}$$

$$\frac{[PL]_{eq}}{[P]_{eq}} = \frac{I(PL)}{I(P)} = R \tag{2.5}$$

$$K_d = \frac{[P][L]}{[PL]} \tag{2.6}$$

The direct ESI-MS assay relies on the direct detection of the gas-phase ions of free and ligand-bound protein P and PL. There are two commonly applied ways for the determination of binding affinities by native ESI-MS. In the following they are discussed for a complex with a 1:1 stoichiometry: (i) With the single-point method, the dissociation constant can be obtained from a single mass spectrum using equation 2.9 under the following two assumptions:[91, 92]

$$[P] = [P]_0 - [PL] \tag{2.7}$$

$$[L] = [L]_0 - [PL] \tag{2.8}$$

$$K_d = \frac{[P][L]}{[PL]} = \frac{([P]_0 - [PL])([L]_0 - [PL])}{[PL]} \tag{2.9}$$

$[P]_0$ and $[L]_0$ are the initial protein and ligand concentrations.

(ii) The titration method involves the recording of several mass spectra, during which the protein concentration is held constant and the ligand concentration is varied. For each of the spectra the ratio R = [PL]/[P] = I(PL)/I(P) is calculated by reading out the intensities I(PL) and I(P). Subsequently, R is plotted against the initial ligand concentration $[L]_0$ and the K_d is obtained from the following expression:[91, 93]

$$R = \frac{1}{2}(-1 - \frac{[P]_0 + [L]_0}{K_d} + \sqrt{4\frac{[L]_0}{K_d} + \left(\frac{[L]_0 - [P]_0}{K_d} - 1\right)^2}) \tag{2.10}$$

In practice R is limited between 0.05 and 20, which results in binding affinities in the µM to mM range.[94] Depending on the protein-ligand system, the fit function can be modified according to several different (isomeric) ligands[95] or to a different binding stoichiometry.[96]

In order to obtain reliable binding affinities with both methods, several assumptions have to be made: First of all, the protein-ligand interaction needs to be specific. Secondly, the free protein and the protein-ligand complex need to have uniform ESI response factors, meaning a similar ionization and detection efficiency. This criterion is usually fulfilled when the ligand is very small and its molecular weight is lower than ~10 % of the protein.[94] However, in the case that the ligand is for example another protein, the ionization efficiencies certainly differentiate and response factors have to be introduced as a parameter in the binding model.[97] Furthermore, collisional heating of gaseous ions may occur at various stages during the passage through the mass spectrometer and lead to dissociation of the protein-ligand complex. Thus, conditions have to be optimized in such a way that in-source dissociation does not occur. Otherwise, this will falsify the magnitude of K_d and in extreme cases even lead to false negatives.[94]

Besides the above-mentioned challenges, non-specific binding can introduce errors into the K_d values. Especially when high ligand concentrations are used, ligand molecules can attach (in addition to the specific binding site) non-specifically to the protein in the final stages of ESI. The simplest way to circumvent non-specific binding is to lower the ligand concentration. However, this is not always possible, in particular not for low affinity ligands where higher concentrations are required to detect reasonable binding. Against all expectations non-specific interactions tend to be quite stable and it is not necessarily possible to remove non-specifically bound ligands by selective in-source dissociation.[94] Diverse strategies have been introduced to correct for non-specific binding, such as the reporter molecule method[98] or the reference protein method,[99] in which a non-interacting reference protein is added and thus the fraction of non-specific binding can be determined from the mass spectra. In addition to these direct approaches, sophisticated models for data analysis have been proposed for deconvolution of specific and non-specific binding.[100-103]

In many studies the experimentally determined dissociation/association constants coincide with values from other methods.[90, 104, 105] However, it is important to note that gas-phase affinities do not always match affinities from solution and that there are likewise several discouraging studies. In one example, K_ds for DNA nucleotides were determined in the low μM to nM range by ESI-MS and were found to be significantly lower than the values obtained by isothermal titration calorimetry (ITC).[106] While ITC in general is able to produce K_d values in this range, it is rather challenging with native MS due to sample handling or instrumental limitations.[107] The comparison of K_d values can also be critical for hydrophobic systems. As hydrophobic contributions are substantially weakened or even vanish completely in the gas phase, a rather low affinity or no formation of these complexes is expected. However, even complexes being largely held together by hydrophobic interactions can remain intact in the gas phase under certain conditions.[108, 109] In conclusion, the success and the reliability of the ligand affinity strongly depend on the system and the mode of interaction between protein and ligand. Nonetheless, native MS has already greatly contributed to gain precious insights on various different systems such as protein-DNA,[110] protein-RNA,[111] protein-carbohydrate,[112] protein-lipid,[113, 114] protein-protein[115, 116] or protein-drug[117] complexes.

2.4.2 Tandem Mass Spectrometry on Protein-Ligand Complexes

In many cases the mass spectrum alone does not yield satisfying information on the investigated biomolecule. This applies in particular for larger proteins, as the peaks broaden and binding to small ligands cannot necessarily be identified accurately. To gain additional information, such as ligand stoichiometry or complex stability, tandem mass spectrometry (MS^2 or MS/MS) is applied. The general procedure involves the isolation and subsequent activation of a precursor ion species and finally the detection of the product ions. The activation of the ion can be induced either in small steps (low energy activation) or in a single fast event (high energy activation).[118] To date, a multitude of tandem MS techniques are available, which are based on different principles, each with its own strengths and drawbacks. Often, two or more techniques are combined to obtain complementary data, and thus increase the information content.

The most commonly used technique is collisional induced dissociation (CID), which is based on collisions with gas molecules. The precursor ion is usually selected in the quadrupole and subsequently guided to the collision cell, where it collides with a nonreactive gas, such as helium, nitrogen or argon (Figure 2.10A). At each collision a fraction of the kinetic energy of the ion is transformed to internal energy (low energy activation), which leads to the dissociation of the ion after a certain number of collisions.[119] The most labile bond, usually a non-covalent interaction, is broken first. This leads to a disassembly of the protein constituents and can shed light on the subunit composition of multimeric proteins or provide stoichiometry information on ligands.[120] The resulting mass spectra of an exemplary CID experiment for a protein-ligand complex are shown in Figure 2.10B.

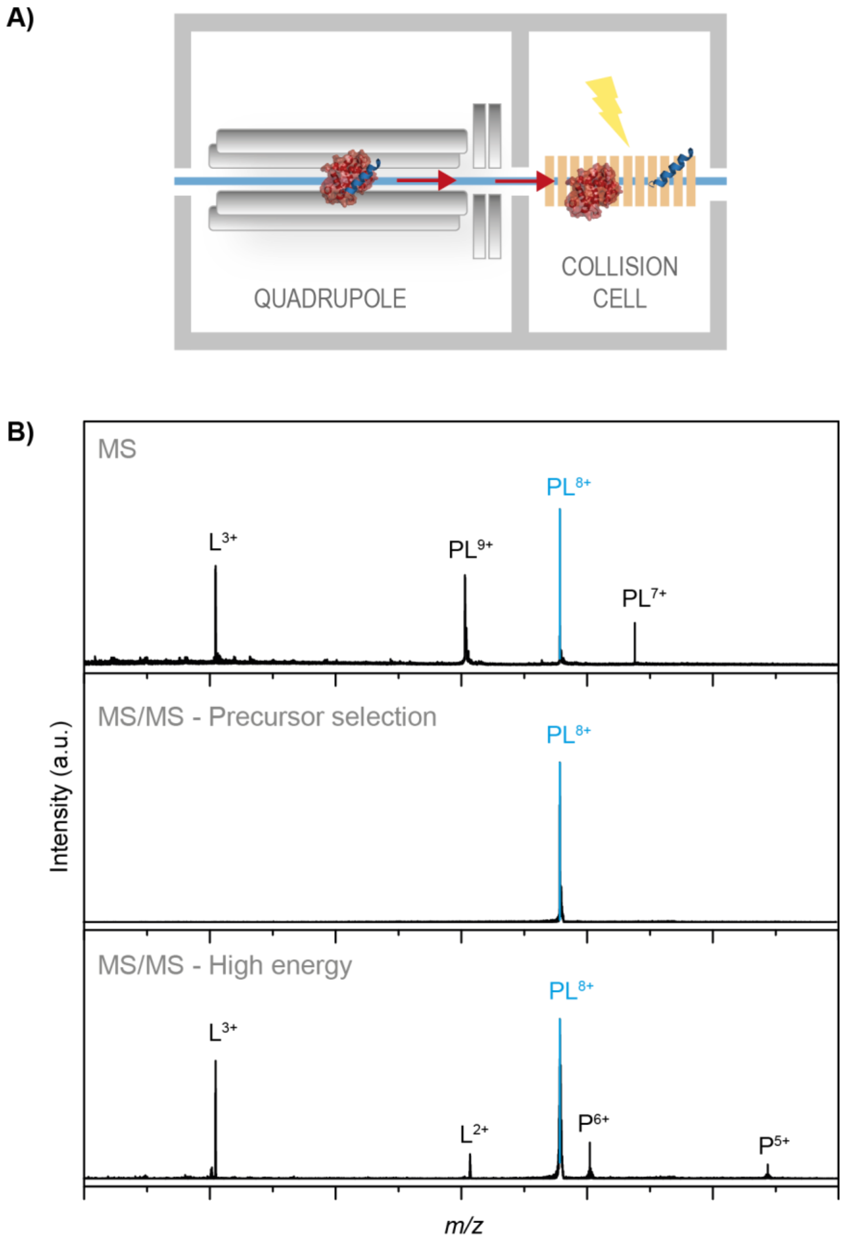


Figure 2.10: Schematic procedure of collision-induced dissociation (CID) of a protein-ligand complex. **A)** A precursor ion species is selected in the quadrupole and subsequently activated by collisions with gas molecules in the collision cell to induce dissociation. **B)** Exemplary mass spectra of a CID experiment. The top spectrum shows the complete mass spectrum of a protein-ligand complex in charge states (P^{9+} to P^{7+}) and the free ligand (L^{3+}). The precursor in charge state 8+ is selected (middle) and dissociation is induced, which results in different product ions of free protein and ligand (bottom).

The full mass spectrum shows a protein-ligand complex in different charge states (P^{9+} to P^{7+}) and the free ligand (L^{3+})(top). The protein-ligand species in charge state 8+ is selected as precursor ion species in the quadrupole (middle) and dissociates at higher energies to different product ions, such as the free protein and free ligand (bottom). By stepwise increasing the energy in the collision cell and determining a CE_{50} value (CE_{50} = collision energy where 50 % of the complex has dissociated) relative binding strengths can be analyzed with CID (this will be covered in more detail in Chapter 5 and 6). Thus, CID can also be used for qualitative affinity screening.

Another tandem MS technique that is based on collisions and that has gained reasonable attention during the last years is surface induced dissociation (SID), where the precursor ions collide with a surface. SID is also used to disassemble protein complexes in native MS and causes dissociation prior to unfolding of the proteins.[121]

Electron capture dissociation (ECD)[122] and electron transfer dissociation (ETD, Chapter 5)[123] are based on interactions with electrons. The capture/transfer of electrons leads to charge reduction and fragmentation, which usually results in a protein backbone cleavage, while non-covalent interactions are retained. These techniques are especially advantageous when posttranslational modifications or ligand-binding sites are investigated.

Ions can also be activated and dissociated by the absorption of photons (photo dissociation), whereby the dissociation efficiency strongly depends on the photon wavelength. Infrared light has only very low energy per photon (~0.1 eV) and consequently multiple photos are required for dissociation (infrared multi photon dissociation, IRMPD).[124] In contrast, ultra violet photo dissociation (UVPD)[125] uses higher energy photons (~3 to 8 eV) to activate and dissociate ions. UVPD has gained increasing attention in recent years, as it enables the characterization of protein primary sequence and the identification of ligand binding sites in one experiment. Furthermore, it produces much more assignable fragments than other tandem MS techniques.[125]

2.4.3 Catch and Release ESI-MS*

The binding of metal ions[126] or other ligands [127, 128] often induces a conformational change of the protein. This can be readily detected by IM-MS, which makes it an attractive tool for the investigation of protein-ligand interactions. Although IM-MS has not yet found its way into routine application in pharma industry, it has gained considerable attention in recent years and several studies in this field demonstrated its enormous potential. Membrane proteins and their complexes with lipids have for example been investigated on various occasions and unique

---

*This subchapter is based on the publication "*Ion mobility-mass spectrometry as a tool to investigate protein-ligand interactions*", published in M. Göth, K. Pagel, *Anal. Bioanal. Chem.* **2017**, *409*, 4305-4310, http://dx.doi.org/10.1007/s00216-017-0384-9. Copyright 2017 Springer Nature.

insights about the effect of lipid binding on the protein stability or the role of the detergents for structural preservation have been obtained.[129, 130] In another example, a screening scaffold with IM-MS was established to identify small molecules, which inhibit the formation of amyloid fibrils, as occurring in neurodegenerative disorders such as Alzheimer's or Parkinson's.[131]

A particular powerful approach to identify and quantify PL binding is *catch and release* electrospray ionization-mass spectrometry (CaR ESI-MS).[40, 132-134] Initially used as a MS-only experiment, Klassen and co-workers developed this approach further and used it in combination with IMS for detailed and conclusive screening studies on carbohydrate libraries.[135] The target protein is first incubated with a compound library and the formed protein-ligand complexes are subsequently analyzed using ESI-MS (Figure 2.11A). The attached ligands are *caught* by the protein and can be in principle identified due to the increase in molecular weight of the corresponding protein-ligand complex. In case of isomeric species or if ligands do not differ significantly in mass, however, this approach often does not provide an unambiguous result. CaR ESI-MS solves this problem by a controlled *release* of the ligands from the proteins in the mass spectrometer using CID. MS or IM-MS, often in combination with further fragmentation, can subsequently be used to unambiguously identify the dissociated ligands.

A recent study on the norovirus P particle demonstrates impressively how the CaR ESI-MS assay can be applied to identify new inhibitors in a fast and straightforward fashion and even provides a qualitative ranking of their binding affinities.[112] Within a 146 compound carbohydrate library, 28 binders were identified, including several milk oligosaccharides as well as bacterial oligosaccharides, which have not yet been known to bind to noroviruses. Isomeric ligands were successfully distinguished based on their drift time (Figure 2.11B, L1 and L5; L11 and L51, respectively). Remarkably, the obtained intensity of each individual species in the CID spectra was in good qualitative agreement with the binding affinities of the corresponding ligands.

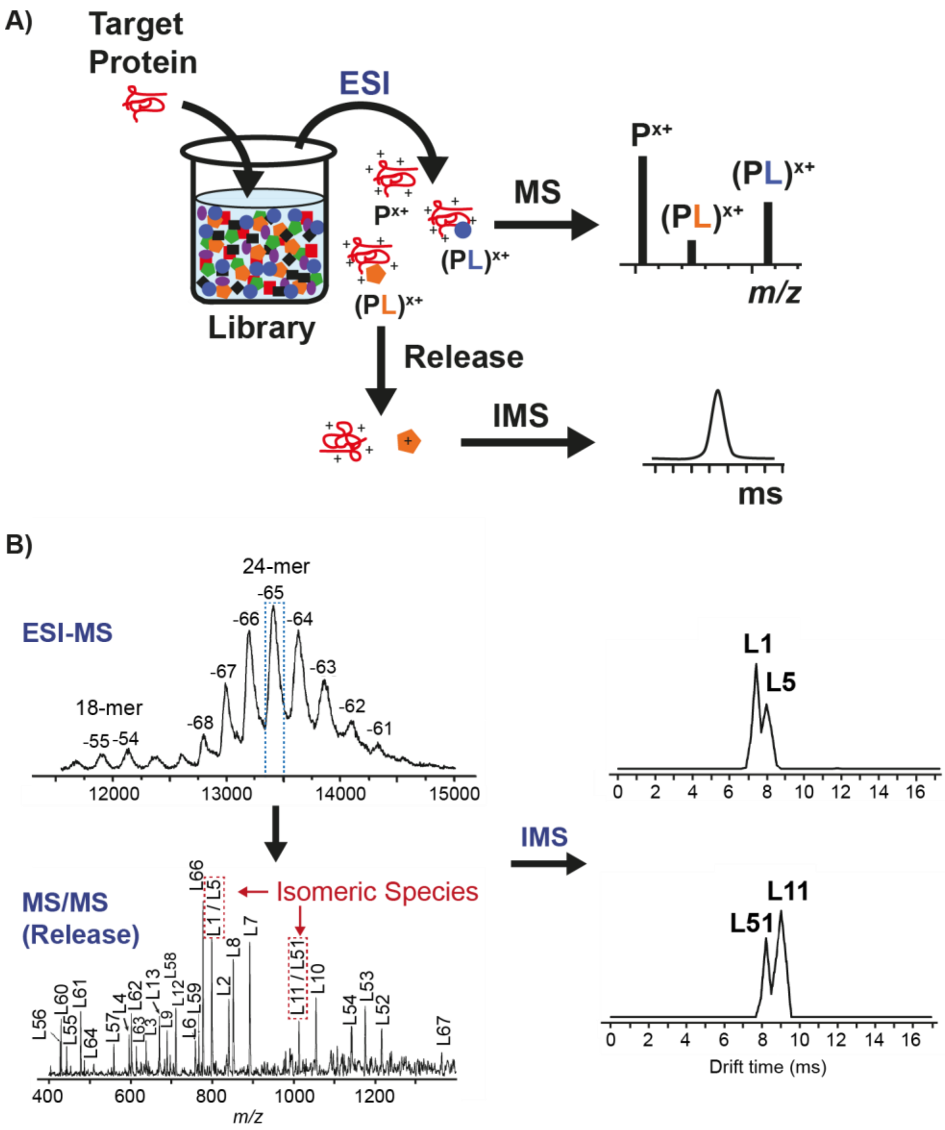


Figure 2.11: A) Scheme of the catch and release electrospray ionization-mass spectrometry assay (CaR ESI-MS). Adapted with permission from [135]. Copyright 2012 American Chemical Society. **B)** Mass spectrum of the norovirus P particle incubated with a compound library. Charge state -65 is selected and subjected to collision-induced dissociation (MS/MS). Released ligands are assigned by their molecular weight. Ligands L1 and L5 as well as L11 and L51 are isomeric and can only be assigned using IMS. All four ligands bind to norovirus P particle, L1 and L11 reveal a higher affinity compared to their isomeric analog. Adapted with permission from [112]. Copyright 2013 American Chemical Society

2.4.4 Collision-Induced Unfolding*

Another remarkable approach for the characterization of PL interactions is collision-induced unfolding (CIU),[136] which can be regarded as the gas-phase analog of calorimetry experiments in solution. Instead of measuring the heat that is released upon binding of the ligand, here, the heat that is required to unfold the complex by collisional activation is measured. Following ionization and transfer into the gas phase, one particular charge state of the protein-ligand complex is *m/z*-selected, for example using a quadrupole mass filter (Figure 2.12A). This precursor is subsequently activated in a collision cell by a stepwise increase of the acceleration voltage, which induces unfolding and eventually dissociation of the complex. The applied voltages are plotted against the drift time or CCS of the species of interest, which leads to a unique fingerprint for each PL complex.

By comparing the fingerprints of apo-proteins to protein-ligand complexes, different aspects can be investigated. For example, CIU showed a great potential to be used for the characterization of functional protein domains[137] and stability analysis.[138-140] Recently, CIU was used in combination with CID to establish an assay for the distinction of two Abelson kinase inhibitors.[141] While type I inhibitors bind to all kinase conformations, type IIs favorably interact with the inactive (closed) state, which makes it advantageous for disease treatment. In the resulting CIU fingerprints significant differences in specific sections were observed for both inhibitors (Figure 2.12B). The corresponding averaged IMS spectra and the number of conformations in these sections are unique for each of the two inhibitor types. Thus, analogous experiments with another data set of unknown inhibitor candidates resulted in a clear assignment to one of the two types of binding.

In a very recent work, CIU was furthermore successfully applied to elucidate the subunit topology of human albumin. In particular, domain-specific binders and different multiprotein constructs were studied and showed for the first time that CIU can be used for the systematic analysis of the unfolding pathway of a multiprotein complex.[142]

---

*This subchapter is based on the publication "*Ion mobility-mass spectrometry as a tool to investigate protein-ligand interactions*", published in M. Göth, K. Pagel, *Anal. Bioanal. Chem.* **2017**, *409*, 4305-4310, http://dx.doi.org/10.1007/s00216-017-0384-9. Copyright 2017 Springer Nature.

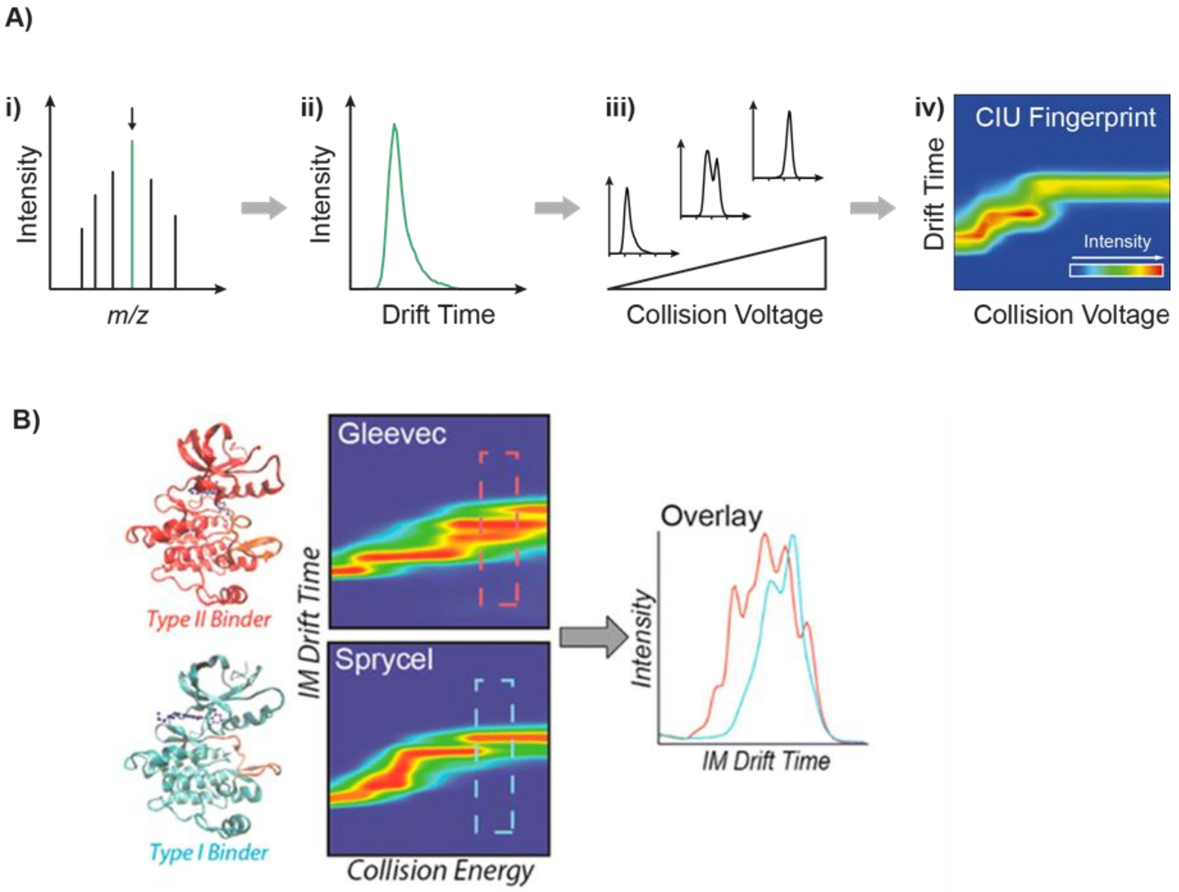


Figure 2.12: A) Schematic description of collision-induced unfolding (CIU). i) ii) One charge state of a protein-ligand complex with its corresponding arrival time distribution (ATD) is selected as precursor from the mass spectrum. iii) The precursor species is activated by gradual increase of the collision voltage, which induces unfolding and finally dissociation of the complex. iv) Collision voltages are plotted as a function of the measured drift time resulting in a characteristic CIU fingerprint. **B)** Application of CIU to discriminate two inhibitors that bind in a conformation-selective fashion. Depending on the type of binding, significant differences in specific sections of the CIU fingerprints are observed. Reprinted with permission from [141]. Copyright 2013 American Chemical Society.

2.5 Instruments

2.5.1 The Ultima nESI-Q-ToF Mass Spectrometer

The Ultima mass spectrometer (Micromass/Waters Corporation, Manchester, UK) is a high-mass quadrupole time-of-flight (Q-ToF) instrument, equipped with a Z-spray nESI source. A schematic overview is depicted in Figure 2.13. The instrument was upgraded for high-mass measurements and is therefore well suitable for the analysis of large biomolecules up to several 100 kDa and ~*m/z* 100,000. A detailed description of a similar high-mass modification can be found elsewhere.[36] Essential modifications involve a pressure increase in several pumping stages of the instrument (indicated as P2 and P3 in Figure 2.13), which is required for collisional focusing (collisional cooling) and successful transmission of large biomolecules. Table 2.1 gives an overview of typical pressure values in standard and high-mass mode.[36] Furthermore, the quadrupole operates in lower frequency (~300 kHz) compared to a standard setup (~832 kHz) so that precursor ions up to *m/z* 32,000 can be selected. In addition, the instrument is capable of performing tandem MS experiments. Due to the upgrade very high collision voltages up to 400 V can be applied to induce dissociation of *m/z*-selected precursor ions.

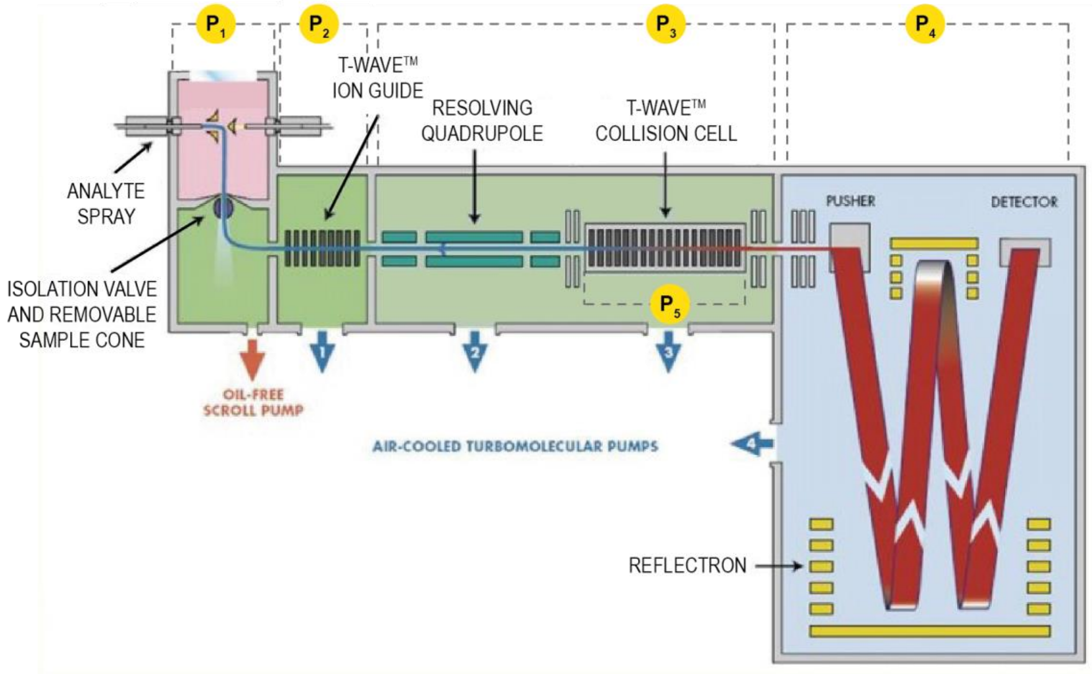


Figure 2.13: Schematic representation of an Ultima nESI-Q-ToF instrument (Micromass/Waters). P1-P5 indicate pressures in different compartments of the mass spectrometer. P2-P4 are modified for high-mass measurements (Table 2.1). Illustration from Waters Corporation.

Table 2.1: Overview of typical pressure values in standard and high-mass mode.

Pressures	Standard mode	High-mass mode
P_2	10^{-4} mbar	10^{-2} - 10^{-3} mbar
P_3	10^{-6} mbar	10^{-4} - 10^{-5} mbar
P_4	10^{-7} mbar	10^{-6} - 10^{-7} mbar

The general workflow of a native MS experiment using the Ultima tandem mass spectrometer begins by spraying the analyte solution from home-fabricated Pt/Pd-coated borosilicate capillaries to ionize the sample molecules *via* nESI. The ions are then transferred into the high vacuum and guided through the ion guide to the quadrupole, where *m/z*-selection of a precursor species can be performed. Mass-selected ions are transferred further to the collision cell. Here, the ions collide with argon gas atoms and collision-induced dissociation (CID) experiments can be carried out. Subsequently, the ions are guided to the time-of-flight analyzer, where they are separated according to their flight time and finally reach the detector.

2.5.2 The Synapt G2-S nESI-Q-IMS-ToF Mass Spectrometer

The Synapt G2-S HDMS (Waters Corporation, Manchester, UK) instrument is a commercially available traveling wave ion mobility-mass spectrometer equipped with a Z-spray nESI source. The principles of traveling wave IMS (TW-IMS) were discussed in Chapter 2.3.3 and detailed technical descriptions of the instrument can be found elsewhere.[80, 143] Here, only the special features of the instrument as well as the general workflow are briefly described.

A schematic overview of the Synapt is presented in Figure 2.14 with the general workflow as follows: The analyte solution is sprayed from home-fabricated Pt/Pd coated borosilicate capillaries and ionized *via* nESI. The ions are then transferred into high vacuum and guided through stepwave ion guides to the quadrupole, where *m/z*-selection of a precursor species can be performed. Subsequently, the ions reach the heart of the machine, the Triwave™, which consists of three compartments: a trap cell, the ion mobility cell and a transfer cell. Trap and transfer cell are each filled with argon at a pressure of ~$1\text{-}5\cdot10^{-2}$ mbar. The ion mobility cell operates with at a pressure of 2-3 mbar. CID can be performed in both, the trap (trap dissociation) and the transfer cell (transfer dissociation). In the original setup, the ion mobility cell is preceded by a helium cell, which provides a soft transition (~10^{-1} mbar) for the ions into the IM cell. The last compartment of the instrument is the high-resolution time-of-flight mass analyzer, where the ions are separated according to their flight time and finally reach the ion detection system.

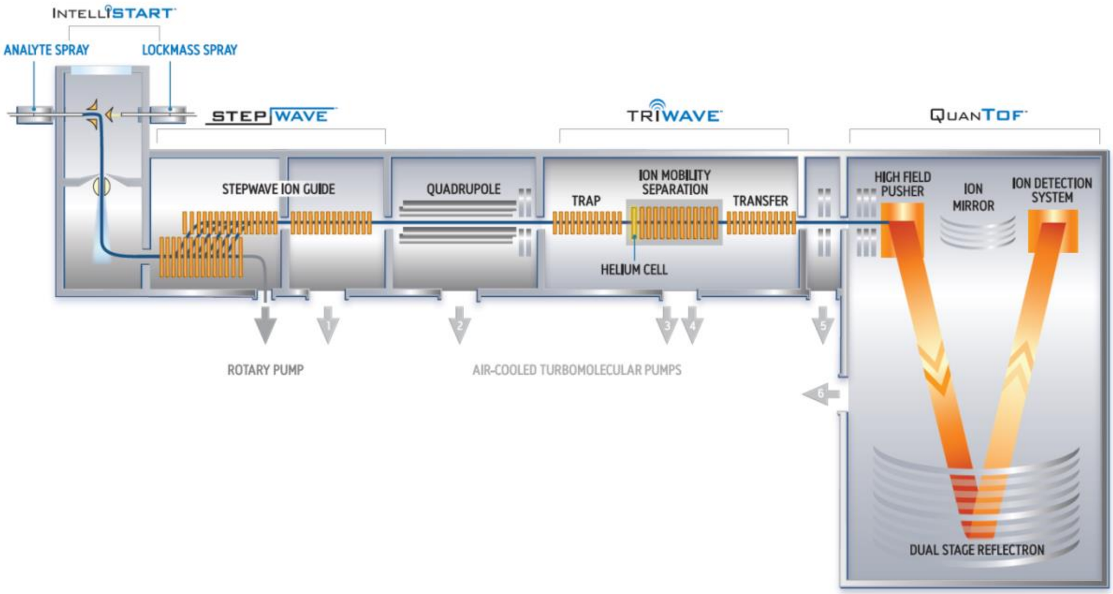
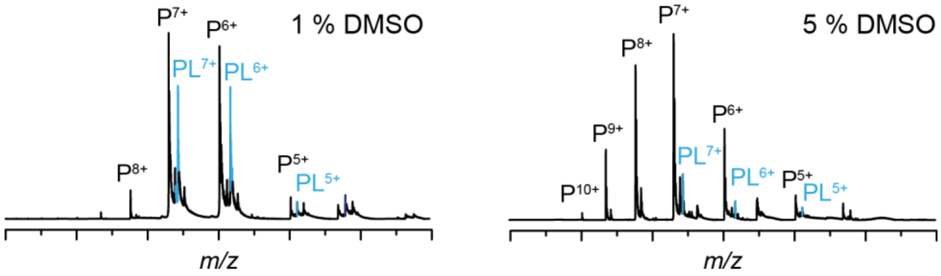


Figure 2.14: Schematic representation of the Synapt G2-S HDMS nESI-Q-IMS-ToF instrument (Waters). Illustration from Waters Corporation.

Great benefits of the Synapt G2-S are the high *m/z*-resolving power and sensitivity, which enable the detection of even very low concentrated samples. In addition, CID-fragmentation of *m/z*-selected ions can be performed prior to and after ion mobility separation. Thus, either conformer-specific fragments or the size and shape of fragments can be investigated depending on the aim of the experiment. As discussed in Chapter 2.3.3, TW-IMS does not allow the direct determination of CCSs due to a non-uniform electric field within the ion mobility cell. Therefore, calibrants with similar physicochemical properties and well-known CCSs have to be measured for CCS estimations. In 2015, the original traveling wave ion mobility cell of the instrument was replaced by a RF-confined drift tube.[144] With this DT-IMS setup, CCS values can be directly determined from the measurements using the Mason Schamp equation (Chapter 2.3.2). Furthermore, an additional gas valve enables also the use of helium as an IM drift gas, while with the original setup only measurements in nitrogen were possible. This allows an easier comparison with literature CCS values, as these are still predominantly in helium. However, the modification also leads to a decrease in the resolution. While the original TW ion mobility cell with a length of ~25 cm had a resolving power of $\Delta\Omega/\Omega(N_2) \approx 40$, the DT ion mobility cell of the same length only reaches a resolution of $\Delta\Omega/\Omega(N_2) \approx 20\text{-}25$ (estimations based on group internal experiments). Data presented in Chapter 5 was measured with the original TW setup, whereas data from Chapter 6 was recorded with the modified instrument.

3 The Role of DMSO in Protein-Ligand Interactions



3.1 Introduction

One key step in drug discovery is the identification of chemical starting points. Protein-ligand interactions have to be studied in detail and a compound that binds with very high affinity to the target protein and is able to compete with the natural binding partner is required. Usually, numerous substances have to be evaluated to find the best fitting compound. Therefore, high-throughput screening (HTS) techniques are employed, in which thousands of potential binders are screened in libraries automatically and relatively fast against the target(s) of interest. As these compound libraries include many different molecules with different properties, a standard solvent is required to simplify and accelerate the sample preparation and measurement process. Commonly, the solvent of choice is the dipolar dimethyl sulfoxide (DMSO), because of its numerous favorable properties: It has a very high dissolving ability, low chemical reactivity, low toxicity and a low vapor pressure.[145] In HTS DMSO concentrations vary usually between 1 and 10 %, depending on the solubility of the compounds.[145, 146] However, in other structural biology type studies the DMSO concentration is likely to add up to 70 %.[146] Gas-phase studies show that high concentrations of DMSO lead to supercharging of proteins and to a disruption of the native structure.[147, 148] Although it was demonstrated that small amounts of DMSO can have protective effects on labile protein interactions,[149] higher amounts of DMSO can lower the binding affinity of protein-ligand complexes.[146] This may lead to deceptive conclusions in protein analysis where high concentrations of DMSO are used. Only a few research groups have addressed to this problem though and the effect itself as well as suitable approaches to overcome it, have to be further investigated. In this chapter the effect of DMSO on proteins and protein-ligand interactions is studied with native mass spectrometry on the basis of bromodomain protein 4 (BRD4).

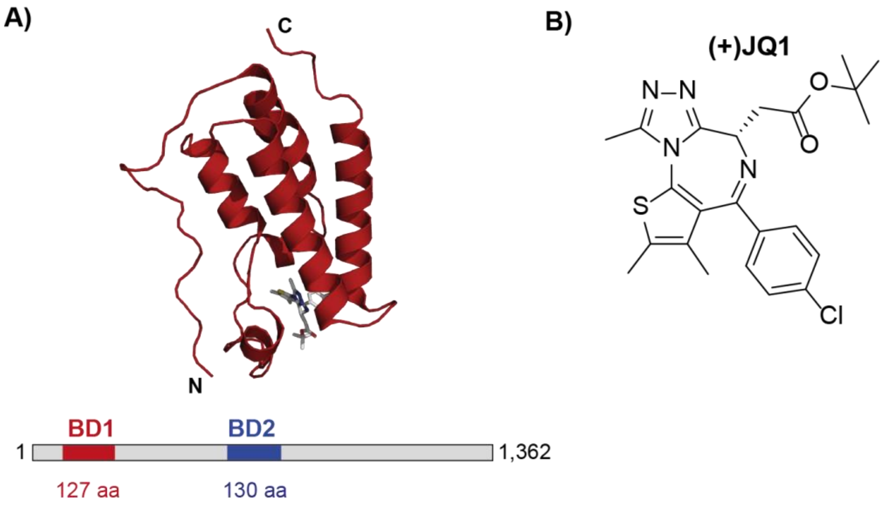


Figure 3.1: A) Schematic sequence of full length bromodomain protein 4 (BRD4) with its two bromodomains BD1 (red) and BD2 (blue). The crystal structure depicts BD1 in complex with the bromodomain inhibitor (+)JQ1 (PDB entry: 3MXF). **B)** Chemical structure of the triazolodiazepine (+)JQ1, which binds to bromodomains with a nanomolar affinity.[150]

BRD4 is a member of the bromodomain and extra-terminal domain (BET) family and the so far best studied protein of this family. With its two bromodomains, BD1 and BD2, BRD4 recognizes acetylated lysine residues in histones and other proteins and thus plays an important role in the regulation of chromatin architecture and transcription.[151] Generally, bromodomains consist of ~110 amino acids that form a common left-handed helical bundle fold with four antiparallel helices (Figure 3.1A).[152] Acetylated lysines bind in the deep hydrophobic cavity, which is formed by the four helices and in addition anchor *via* a hydrogen bond to an asparagine residue. In recent studies the role of BRD4 in cell proliferation and cancer growth has been reported, which makes this protein a valuable target for cancer research.[153-155]

One widely-known and highly potent bromodomain inhibitor is the triazolodiazepine (+)JQ1 (Figure 3.1B).[150] The binding affinity to bromodomains has been investigated with several methods and the dissociation constant was determined in the nanomolar range by isothermal titration calorimetry (ITC).[150, 151] While (+)JQ1 is a very strong binder, its isomer, (-)JQ1 has no perceptible affinity to bromodomains.[150] The following study on the impact of DMSO is performed with the inhibitor (+)JQ1 and the two BRD4-bromodomains BD1 and BD2.

3.2 Experimental Details

Sample Preparation and Mass Spectrometry

Bromodomain 1 (BD1, MW= 15,083.26 Da) and bromodomain 2 (BD2, MW= 15,036.28 Da) from BRD4 human were obtained from Bayer AG. The inhibitor (+)JQ1 (MW= 456.99 Da, 1 mg crystalline solid) was purchased from Biomol. Prior to analysis the proteins were buffer exchanged into ammonium acetate (Sigma Aldrich, 7.5 M, pH = 7.0 - 7.5). Therefore, the corresponding protein solution was dialyzed twice for 2 h and overnight in a 150 mM (or 50 mM) ammonium acetate solution (Slide-A-Lyzer MINI dialysis devices, MWCO 2 kDa, Thermo Scientific). After dialysis the concentration was determined by measuring the UV absorbance at 280 nm (Specord 40, Analytik Jena AG and later using a NanoDrop (Thermo)) using molar extinction coefficients of 26030 $L\cdot mol^{-1}\cdot cm^{-1}$ and 13610 $L\cdot mol^{-1}\cdot cm^{-1}$ for BD1 and BD2, respectively. The proteins were either analyzed immediately or divided into aliquots and stored at -26 °C until usage.

The ligand (+)JQ1 was dissolved in DMSO to a 10 mM stock solution and further diluted to less concentrated solutions (with DMSO and/or H_2O) individually for each experiment. Different amounts of DMSO from 0 to 50 % (*v*/*v*) were added to the protein-ammonium acetate solution with a final protein concentration of 10 μM in each sample. For binding studies with (+)JQ1, stock solutions with different inhibitor concentrations were prepared (100 μM to 10 mM, 100 % DMSO) and further diluted with H_2O. Within one titration experiment varying inhibitor concentrations were added to the protein with a final protein concentration of 5 μM, and a constant DMSO level in all samples. Ligand stock solutions and the final protein samples were freshly mixed before each measurement.

ESI mass spectra were acquired with a Synapt G2-S mass spectrometer (Waters) and with an Ultima high-mass quadrupole time-of-flight (Q-ToF) mass spectrometer (Micromass/Waters), both equipped with a nESI source. Details to both instruments are discussed in Chapter 2.5. The electrospray source was operated in positive ion mode. Ions were formed using nanoflow borosilicate electrospray capillaries, which were pulled and Pd/Pt coated in-house. Later titration experiments with BD1 were performed using a chip-based TriVersa NanoMate (Advion®) as electrospray source. The spectra were externally calibrated with a 20 μg/μL CsI solution (water/2-propanol 1/1). Typical parameters for the measurements were as follows: (Synapt) capillary voltage 0.9 – 1.2 kV, sample cone 50 V, source temperature 20 °C, trap collision voltage 2 V, transfer collision voltage 8 V, ion mobility cell switched on, helium cell gas flow 160 mL/min, IMS gas flow 70 mL/min, trap DC entrance 3 V, trap DC bias 38 V, trap DC exit 3 V. (Ultima) capillary voltage 1.5 kV, sample cone 60 V, source temperature 40 °C, RF lens 1 40 V, collision energy 0 V, collision cell pressure 4.0 – 4.6 x 10^{-3} mbar. (NanoMate) voltage 1.7-1.9 kV, pressure 0.7 psi, positive ion mode, manual mode.

Data Analysis

The mass spectrometer was controlled by the software MassLynx (Ultima: version 4.0, Synapt version 4.1). Manual data analysis was performed with Excel (Microsoft Corporation) and OriginPro 8.6 (OriginLab Corporation) software. The average charge state (cs_{av}) of the protein ions was calculated using the intensity $I(P^{n+})$ of every detected charge state n with the following equation:

$$cs_{av} = \sum_n [I(P^{n+}) * n] / \sum_n I(P^{n+}) \tag{3.1}$$

The general procedure for the determination of dissociation constants (K_d) for protein-ligand complexes from ESI mass spectra was discussed in detail in Chapter 2.4.1 and the assumptions and equations therein were applied for the data analysis in this chapter. Instead of plotting R =[PL]/[P] as a function of initial ligand concentration $[L]_0$, [PL] was plotted directly against $[L]_0$ in this chapter and Eq. 2.10 (Chapter 2.4) was transformed accordingly to yield Eq. 3.2[156] as the fit function for titration experiments:

$$[PL] = 0.5 * [([L]_0 + [P]_0 + K_d) - \sqrt{([L]_0 + [P]_0 + K_d)^2 - 4[P]_0[L]_0}] \tag{3.2}$$

3.3 Results and Discussion

3.3.1 The Influence of DMSO on the Gas-Phase Structure of BRD4

The first part of this chapter examines the effect of DMSO on the two bromodomains of BRD4, BD1 and BD2. The study is conducted using native (non-denaturing) mass spectrometry, which was introduced in Chapter 2.2. Proteins are sprayed from volatile solutions at near neutral pH to mimic physiological conditions and preserve their tertiary and quaternary structure. By using organic solvents and acids it is possible to induce unfolding of the protein. This is reflected in the resulting mass spectra and can be observed for BD1 and BD2 when comparing Figure 3.2A with Figure 3.2B.

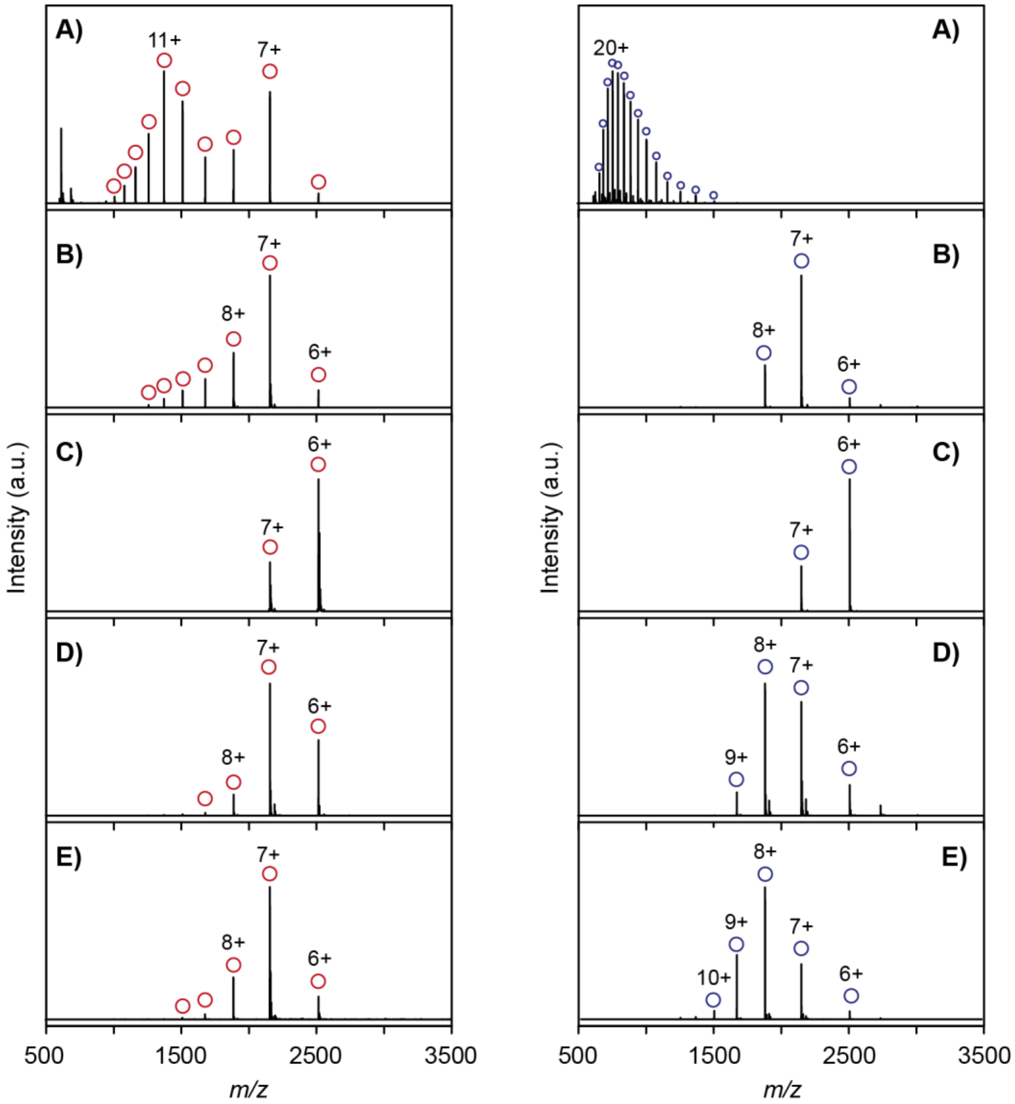


Figure 3.2: nESI mass spectra of BRD4 BD1 (left column, red open circles, MW_{calc}= 15,083.26 Da, MW_{exp}= 15,083.30 Da) and BD2 (right column, blue open circles, MW_{calc}= 15,036.28 Da, MW_{exp}= 15,036.08 Da) sprayed **A)** under denaturing conditions from a 1:1 MeOH:H2O, 0.1 % formic acid (FA) solution and under non-denaturing conditions sprayed from a 150 mM ammonium acetate solution (pH = 7.0 - 7.5) with **B)** 0 % DMSO, **C)** 1 % DMSO, **D)** 10 % DMSO, and **E)** 20 % DMSO. The protein concentration was 10 µM.

It becomes apparent that both proteins adopt a broad(er) charge-state distribution with higher intensities at lower *m/z* values when sprayed from a MeOH/H_2O solution (1/1, with 0.1 % FA, Figure 3.2A), whereas when sprayed from an ammonium acetate solution, a narrow(er) charge-state distribution is observed with a generally lower average charge. The effect is more obvious for BD2 (right column in Figure 3.2). Under denaturing conditions the protein is presumably unfolded with a charge state maximum at 20+, while under native-like conditions the charge-state envelope consists of only three charge states with a maximum intensity on 7+. In case of BD1 the mass spectrum in Figure 3.2A (left column) points to a partially unfolded protein structure with two charge-state distributions centered around 7+ and 11+. The experimentally determined molecular weights MW_{exp}= 15,083.30 Da and MW_{exp}= 15,036.08 Da for BD1 and BD2, respectively, are in good accordance with the calculated values MW_{calc}= 15,083.26 Da and MW_{calc}= 15,036.28 Da.

In order to determine the influence of DMSO on the two protein domains, varying amounts of DMSO (*v*/*v*) were added to the protein solutions. Figure 3.2B-E show the resulting mass spectra of 10 µM protein solutions of BD1 (red open circles, left column) and BD2 (blue open circles, right column) with 0, 1, 10, and 20 % DMSO (*v*/*v*). In general, the mass spectra of both proteins show similar trends: At low DMSO concentrations (1 % (C), also at 2 % (data not shown)) the charge-state envelope becomes narrower and shifts toward lower charge states (higher *m/z* values) in comparison to the spectra in ammonium acetate. With increasing DMSO concentration the charge-state distributions broaden again and the intensities shift towards higher charge states (lower *m/z* values). This process starts at ~10 % DMSO (D) in case of BD1 and already at ~5 % DMSO in case of BD2. The trend becomes more obvious when plotting the average charge state of every mass spectrum against the amount of DMSO in the sample (Figure 3.3). With 1 % and 2 % DMSO the average charge is basically identical for both protein domains (~6.3+), whereas with increasing DMSO concentrations the average charge of BD2 is always higher than for BD1 (*e.g.* at 20 % DMSO 7.1+ *vs.* 8.1+, BD1 *vs.* BD2, respectively). Analog measurements on the Ultima Q-ToF MS lead to a similar trend (data not shown). With both instruments it is challenging to spray from solutions with high DMSO levels (> 20 %) and a stable spray could not be obtained for more than a few seconds with a DMSO concentration of 50 %.

These first results show that DMSO has an influence on the gas-phase structure of the investigated bromodomains BD1 and BD2. A low amount of DMSO in the spray solution causes a shift to lower charge states in the native protein spectrum. This charge-state shift at low DMSO concentrations has been previously observed with other proteins.[145-147] With increasing level of DMSO in the protein solution, however, the signals in the mass spectra shift towards higher charge states (lower *m/z* values). This supercharging effect can be induced also with other agents such as *m*-nitrobenzyl alcohol (*m*-NBA) or sulfolane, referred to as supercharging agents (SCAs).[157, 158]

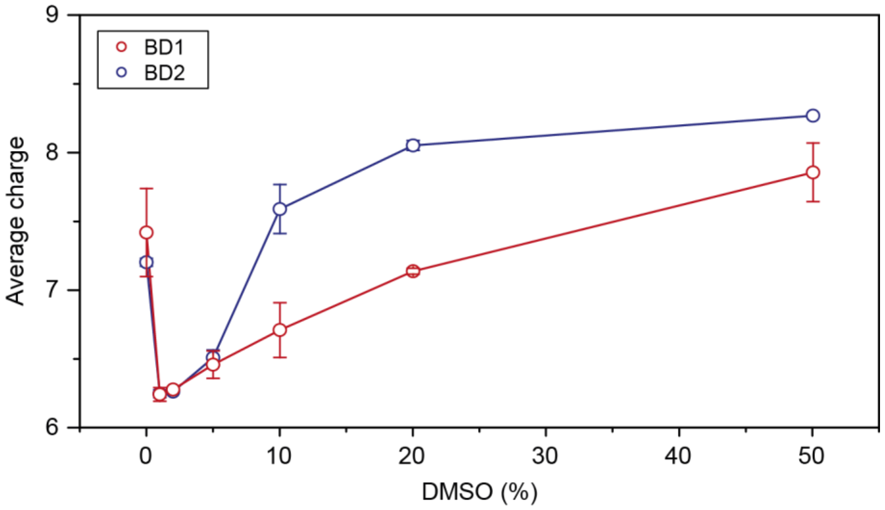


Figure 3.3: Average charge states of BD1 (red open circles) and BD2 (blue open circles) (each 10 µM) as a function of DMSO concentration (*v/v*). Data and error correspond to the average of two independent measurements recorded on the Synapt G2-S mass spectrometer.

Understanding the mechanism of DMSO is not trivial, as several different factors play a role. On the one hand the charge-reducing effect of DMSO in the electrospray process has been accredited to the basic nature of DMSO as well as compaction of the protein structure.[145, 147] A recent ion mobility study supports these findings by reporting a small but reproducible decrease in collision cross sections when low concentrations of DMSO are present in the spray solution.[159] Another important factor is the surface tension of the electrospray droplets. Adding DMSO to an aqueous solution decreases the surface tension of the droplets (44 mN/m *vs.* 72 mN/m).[160] Therefore they can hold fewer charges before reaching the Rayleigh limit and consequently, reduced-charge species are detected in the mass spectrum. However, due to its low vapor pressure and high boiling point (189 °C at 1 bar),[161] DMSO is enriched in the droplet during evaporation. As high amounts of DMSO can lead to protein denaturation, protein unfolding is likely to occur at later stages of the droplet lifetime, which results in significant increase in charge at higher DMSO concentrations. A comparison between unfolding in solution and supercharging in the gas phase points to a 3 to 5 % (mole fraction) enrichment of DMSO in the droplet.[147]

Recently, the supercharging of proteins by *m*-NBA and sulfolane was investigated using molecular dynamics simulations.[162] The study strongly indicates that the mechanism is not related to surface tension nor chemical or thermal denaturation, as proposed before, but can be attributed to a charge trapping within ESI nanodroplets. While in pure water droplets, charge carriers (such as Na^+) can be ejected during evaporation, this does not occur in a similar extent in droplets with SCAs. On the one hand SCA accumulation in the outer layers of the droplet hampers charge carriers to access the surface of the droplet, where they can be ejected. On the other hand, the remaining charge carriers are bound to the protein deep in the core of the droplet after all water has evaporated. Thus, the protein adopts higher charge states. So far it has not been investigated if this mechanism also applies for DMSO, however, it would be a convenient explanation.

3.3.2 The Influence of DMSO on the Interaction of BRD4 and (+)JQ1

The influence of DMSO on protein-ligand affinities has only been addressed by a few mass spectrometric studies so far. While there is indication that small amounts of DMSO can not only stabilize the protein's quaternary structure, but also stabilize protein-ligand interactions,[149] other studies show that DMSO can decrease the binding affinity of protein-ligand interactions even at low DMSO levels.[145, 146] In this section the influence of DMSO on the interaction between the two bromodomains of BRD4 and the inhibitor (+)JQ1 (457 Da) is investigated.

Complexes of BD2 with (+)JQ1 are observed in the gas phase in different charge states and this is exemplarily shown in Figure 3.4. The top spectrum shows the bare protein measured under native-like conditions in presence of 1 % DMSO (blue open circles). Upon adding 10 μM of inhibitor the formation of a protein-ligand complex (PL) with charge states ranging from 7+ to 5+ is observed (blue filled circles). Complex formation with (+)JQ1 is also observed with the first bromodomain BD1 (spectra not shown).

To determine whether DMSO influences the protein-ligand affinity, titration experiments were performed by measuring several spectra with constant protein concentration and DMSO level, and varying ligand concentrations (0 to 50 μM). Subsequently, the peak heights of the bare protein and protein-ligand complex are extracted for every charge state of every spectrum to calculate the amount of formed complex [PL]. Figure 3.4B shows [PL] in dependence of the ligand concentration $[L]_0$. Curve fitting with Eq. 3.2 leads to a dissociation constant K_d of 0.30 ± 0.20 μM. This value agrees with values determined by isothermal titration calorimetry.[150] However, the error is relatively large and despite the fact that the fit seems to match well, it is to be questioned whether the here determined dissociation constant is reliable.

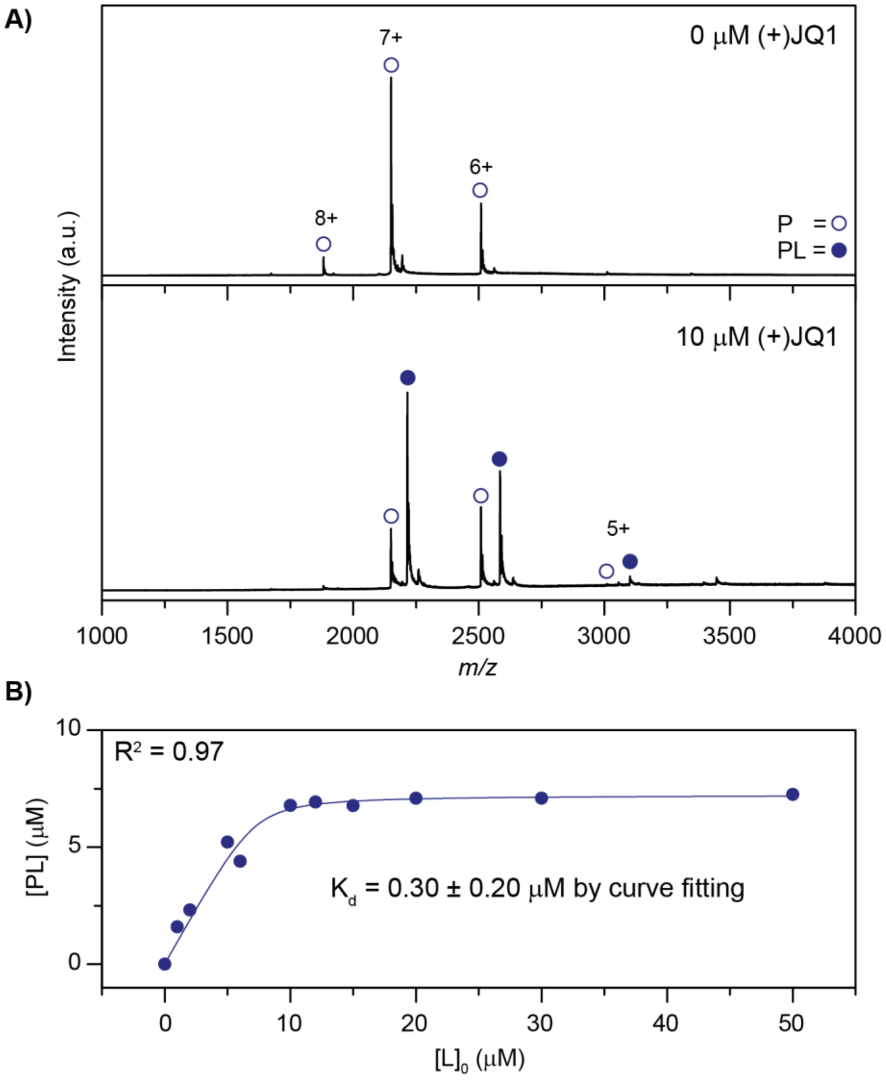


Figure 3.4: A) nESI mass spectra of BD2 (10 µM) in 10 mM ammonium acetate and in presence of 1 % DMSO without ligand (top) and with 10 µM (+)JQ1 (bottom). The complex forms in charge states 5+ to 7+ (filled blue circles). **B)** Amount of protein-ligand complex [PL] determined by ESI-MS titration plotted as a function of increasing ligand concentration $[L]_0$. The dissociation constant K_d was determined by curve fitting (Eq. 3.2). Data were recorded on the Synapt G2-S instrument.

Although in some studies K_ds in the nanomolar range were determined successfully,[93] it is challenging to generate reliable K_ds with mass spectrometry for such high-affinity binders. Furthermore, hydrophobic interactions often play a substantial role in protein-ligand complexes. Although there is strong evidence that hydrophobic complexes can be transferred intact into the gas phase, these forces are theoretically weakened or get lost completely and the success strongly depends on the system.[54] Therefore, differences in the dissociation constant between condensed phase and gas-phase methods do not have to necessarily imply doubts on the reliability of the technique. However, care has to be taken on data interpretation and admittedly mass spectrometry might not be suitable for some systems. Interestingly, complete binding of JQ1 is not even observed with high ligand concentrations. Instead, spectra with high concentrations show multiple binding. This indicates that the ligand concentration in the sample is too high and as a consequence the ligand additionally attaches non-specifically to the protein during the electrospray process when the droplet is shrinking. One further question in this regard is, if the soft conditions used here are still too harsh to enable quantitative complex formation or if there is some other reason why the complex is saturated already at a protein-ligand: protein ratio of ~70 %.

Similar titration experiments as in Figure 3.4B were performed again for BD2 and JQ1 in presence of either 1 % or 5 % DMSO using the Ultima Q-ToF instrument. Exemplary mass spectra for a ligand concentration of 0, 3, 10, and 15 μM are presented in Figure 3.5. With this instrument, charge state 5+ appears in a slightly higher intensity in the spectra with 1 % DMSO than measured with the Synapt mass spectrometer (Figure 3.5A). Binding is also observed in charge states 7+ to 5+. Comparing the two titration sets, two issues become immediately apparent: i) The charge state distribution is much broader with 5 % DMSO in the sample solution (Figure 3.5B), as also observed in the previous section. ii) A comparison of mass spectra for similar ligand concentrations shows significantly lower binding for the 5 % DMSO samples in contrast to the 1 % DMSO samples.

In analogy to the previous experiment, the peak intensities of bare protein (open circles) and protein-ligand complex (filled circles) were extracted and the fraction of formed complex [PL] was plotted against the applied ligand concentration $[L]_0$. The graph in Figure 3.6 illustrates that a higher amount of complex is formed with only 1 % DMSO (blue data) than with 5 % DMSO (grey data). Unfortunately, only ligand concentrations up to 25 μM and 50 μM could be measured with 1 % and 5 % DMSO, respectively. Since at these concentrations the protein was not yet saturated, it was not possible to determine a reliable dissociation constant by curve fitting, as in Figure 3.4.

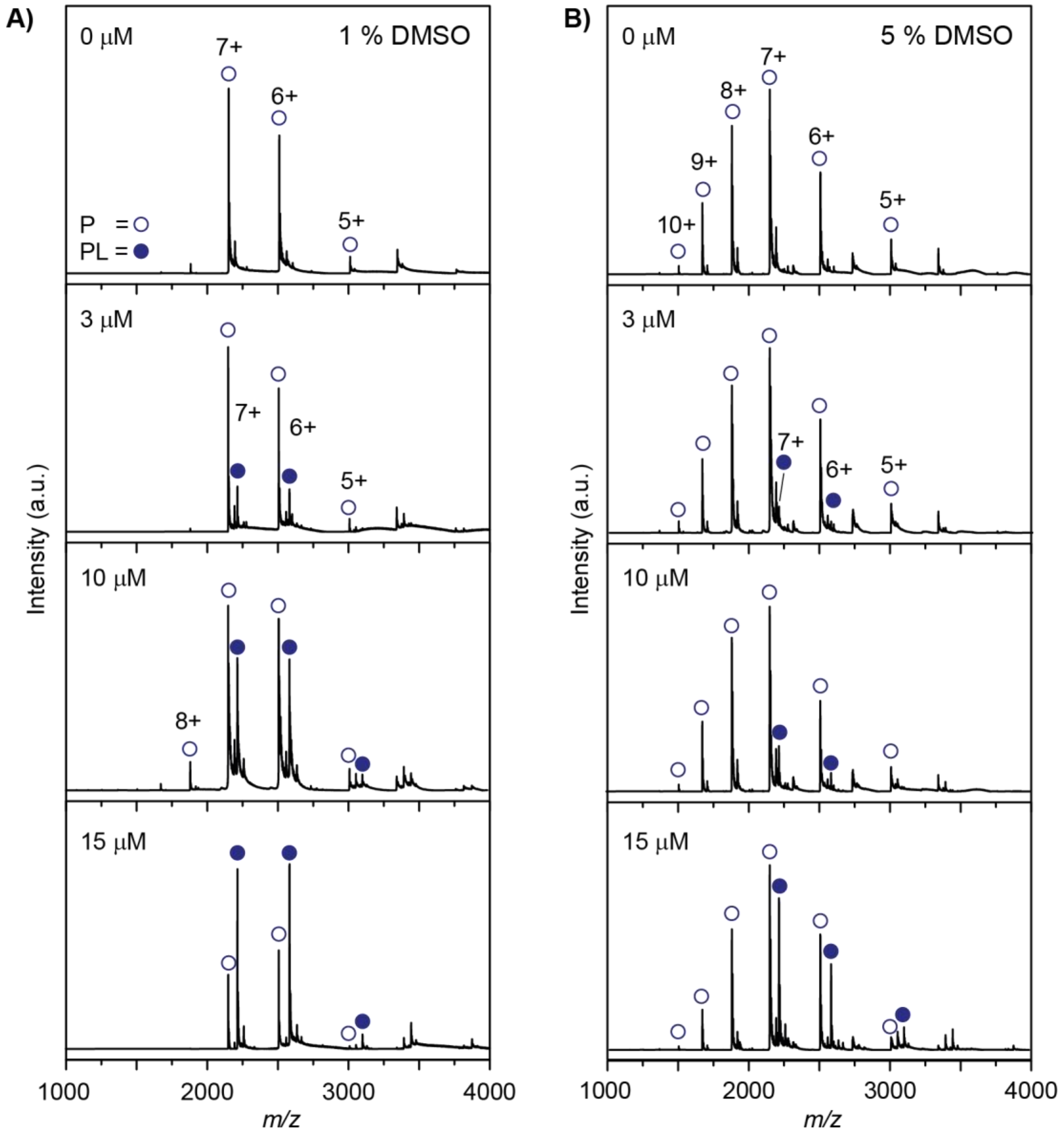


Figure 3.5: nESI mass spectra from a titration experiment with BD2 in presence of **A)** 1 % and **B)** 5 % DMSO. For the titrations a constant protein concentration of 5 μM and varying ligand concentrations were used. The spectra show ligand concentrations of 0 μM, 3 μM, 10 μM, and 15 μM from top to bottom. Open circles correspond to signals of the bare protein; filled circles indicate the protein-ligand complex. Experiments were performed using the Ultima Q-ToF instrument.

The direct ESI-MS assay allows determining the dissociation constant from only one mass spectrum of a protein-ligand complex (single-point method, Chapter 2.4.1, Eq. 2.9).[94] Although this value serves well for qualitative comparisons, it is error-prone, as it represents only one specific protein-to-ligand binding ratio and therefore not necessarily the absolute dissociation constant. Table 3.1 shows the obtained K_ds calculated from the mass spectra for a ligand concentration of 15 μM and 25 μM by the single-point method.

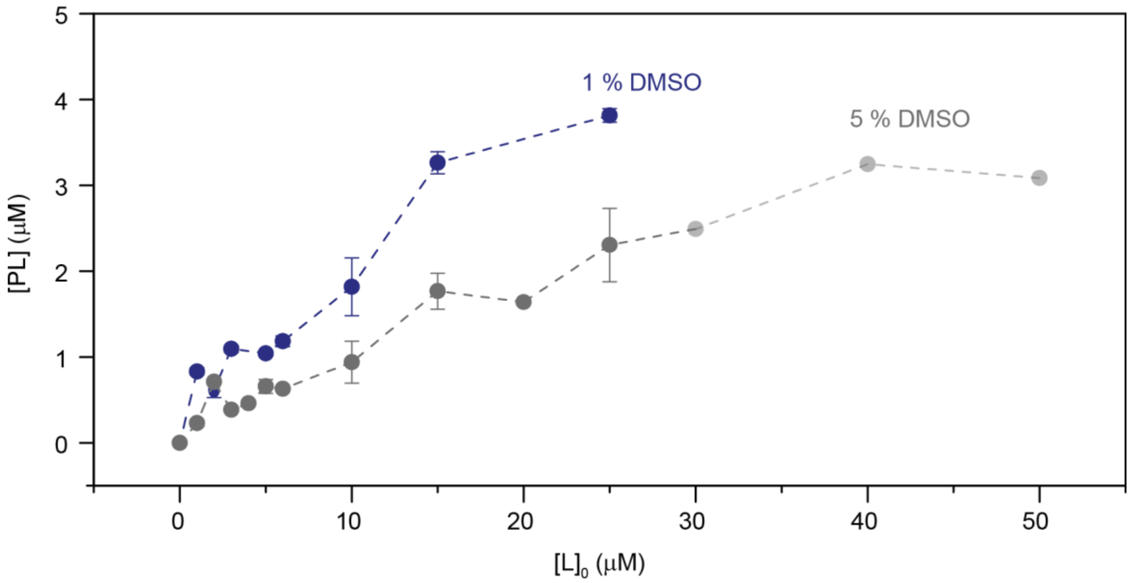


Figure 3.6: Dependency of ligand-bound BD2 [PL] on ligand concentration $[L]_0$ for titration experiments with 1 % (blue) and 5 % (grey) DMSO in the spray solution and a protein concentration of 5 μM. Data represent the mean of two independent measurements with the standard deviation. The samples with 30 μM, 40 μM, and 50 μM were only measured once (light grey data).

Table 3.1: Dissociation constants K_d for the BD2/(+)JQ1 complex determined by the single point method.

$[L]_0$ (μM)	K_d (μM) 1 % DMSO	K_d (μM) 5 % DMSO
15	6.3 ± 0.6	24.5 ± 3.4
25	6.6 ± 0.4	27.5 ± 6.9

In presence of 1 % DMSO a K_d of ~ 6 μM was determined, whereas in presence of 5 % DMSO the value is ~25 μM. K_d values for both ligand concentrations match within the error. Inspecting the mass spectra, the [PL] dependences as well as the calculated dissociation constants by the ESI-MS assay, all data strongly indicate that DMSO decreases the protein-ligand affinity of BD2 and (+)JQ1. A similar trend is also observed in case of the complex between BD1 and (+)JQ1 and the corresponding data are illustrated in Figure 3.7. Titration experiments with this bromodomain were performed for samples with 1 % (red data), 5 % (grey data), and 10 % DMSO (light grey data). In analogy to Figure 3.6 the fraction of formed complex [PL] was plotted against the applied ligand concentrations $[L]_0$. The graphs for the different titration experiments show again that in general more complex is formed when the spray solution contains lower amounts of DMSO.

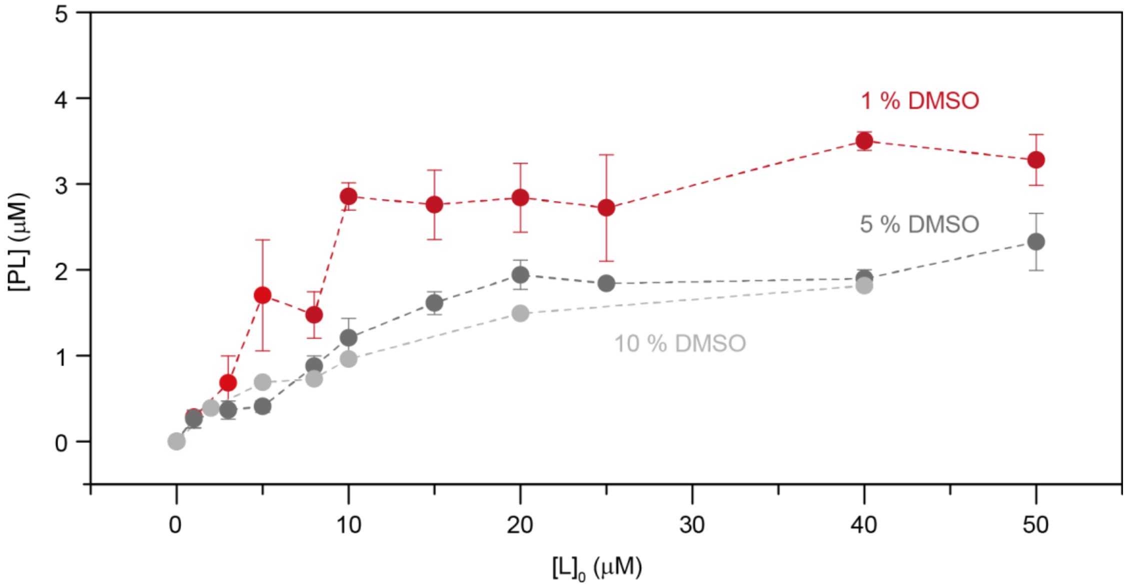


Figure 3.7: Dependency of ligand-bound BD1 [PL] on ligand concentration [L]$_0$ (0-50 µM) for titration experiments with 1 % (red), 5 % (grey), and 10 % (light grey) DMSO in the spray solution and a protein concentration of 5 µM. Data represent the mean of three independent measurements with the standard deviation. In case of the titration with 10 % DMSO the measurement was only performed once.

The difference is significant when comparing the trend for 1 % and 5 % DMSO (red and grey data), even though the deviations are quite large for some data points in the 1 % trend line (for example for [L]$_0$ 5 µM and 25 µM). Increasing the DMSO level even higher to 10 %, results in a further decrease of complex formation (light grey data). This difference to the 5 % titration data is meaningful, but not as large as the difference between 1 % and 5 % data. Furthermore, this experiment was only measured once, thus no standard deviation could be determined. Table 3.2 shows the dissociation constants for the three titration experiments at four different ligand concentrations (5 µM, 10 µM, 20 µM, and 40 µM), which were also determined by the single-point method. Two major conclusions can be drawn from of these values: (i) Within a titration experiment with one specific DMSO concentration, the K_d values determined at different ligand concentrations significantly differ from each other. This is probably in an acceptable range when considering the 1 % titration, where the determined values lie between ~9 and ~16 µM.

Table 3.2: Dissociation constants K_d for the BD1/(+)JQ1 complex determined by the single point method.

$[L]_0$ (μM)	K_d (μM) 1 % DMSO[a]	K_d (μM) 5 % DMSO [a]	K_d (μM) 10 % DMSO [b]
5	9.1 ± 6.7	53.7 ± 11.0	26.9
10	5.4 ± 0.8	29.4 ± 9.2	38.1
20	13.8 ± 4.8	28.8 ± 4.1	43.6
40	15.7 ± 1.6	62.6 ± 5.4	67.0

[a] K_d values and errors correspond to the average of three measurements. Values were determined for every titration experiment individually and subsequently averaged. [b] As the titration experiment with 10 % DMSO was only carried out once, no standard deviation could be determined.

However, inspecting the values for the 5 % DMSO titration the K_d values fluctuate between ~29 μM ($[L]_0$ = 10 μM, or 20 μM) and ~63 μM ($[L]_0$ = 40 μM). These results illustrate that the single-point method does not yield reliable constants for every system. (ii) K_d values increase for increasing DMSO level. Applying for example a ligand concentration $[L]_0$ of 20 μM, dissociation constants of ~14 μM (1 % DMSO), ~30 μM (5 % DMSO), and ~44 μM (10 % DMSO) have been obtained. These values support the observed trend in Figure 3.7 and strongly indicate a lower binding affinity with increasing DMSO concentration. One exception can be found in case of the K_d values determined for a ligand concentration of 5 μM. At this concentration the value for the 10 % DMSO titration set points to a higher binding affinity than with 5 % DMSO in the spray solution. However, as stated before, the set with 10 % DMSO has only been measured once and therefore this deviation could also be an exception. It would also be interesting to investigate samples with less than 1 % DMSO in order to reveal how much influence already minor DMSO levels have on this system. Unfortunately, this could not be investigated with this system, due to limitations in ligand solubility and sample volume.

Taken together, titration experiments with both bromodomains, BD1 and BD2, in complex with the inhibitor (+)JQ1 clearly show that DMSO has an influence on the binding affinity of this system. In this matter the question arises whether the decrease of binding affinity with increasing DMSO concentration can be attributed to a less compact and therefore less native-like protein structure that cannot form the binding pocket properly or if the DMSO competes with the inhibitor. Mass spectra that were recorded with high DMSO concentration in the spray solution show DMSO adducts (spectra not shown), which supports the second assumption. However, a change of charge-state distribution towards higher charge states with higher DMSO concentrations is also observed, indicating unfolding and a less native-like protein structure. Thus, probably both factors play a role in this protein system. To unravel the effect of DMSO, further experiments also with a variety of other protein systems have to be performed.

3.3.3 Substitution of DMSO by Other Solvents

In the previous subchapter it was shown that higher DMSO concentrations can lead to a decreased protein-ligand affinity. One approach to overcome this effect consists in the substitution of DMSO by other solvents, which is investigated in the following. In order to find a suitable solvent, the compound solubility has to be considered first. Secondly, the substitute solvent should not alter the charge-state distribution in the mass spectrum considerably, nor lead to immediate unfolding (or precipitation) of the protein. Ideally, the solvent of choice has similar properties to DMSO, which means it is polar and/or aprotic. Four different solvents were selected for a primary investigation with BD1: dimethylformamide (DMF), tetrahydrofuran (THF), acetonitrile (ACN), and ethanol (EtOH), which are all polar, but not all aprotic. (+)JQ1 is well soluble in DMF (16 mg/mL) and EtOH (14 mg/mL), thus these two solvents are regarded as the most promising DMSO substitutes. Sample solutions with the bare protein BD1 in ammonium acetate and 1 % of each substitute were produced and the resulting mass spectra sprayed from different solvent compositions are shown in Figure 3.8. Comparing spectra B) to D) with the control A) in ammonium acetate, the spectrum sprayed from 1 % DMF (B) resembles most likely the charge-state distribution in A). In contrast, THF (C), ACN (D), and EtOH (E) lead to a slight shift towards higher charge states and are therefore rather not suitable to substitute for DMSO.

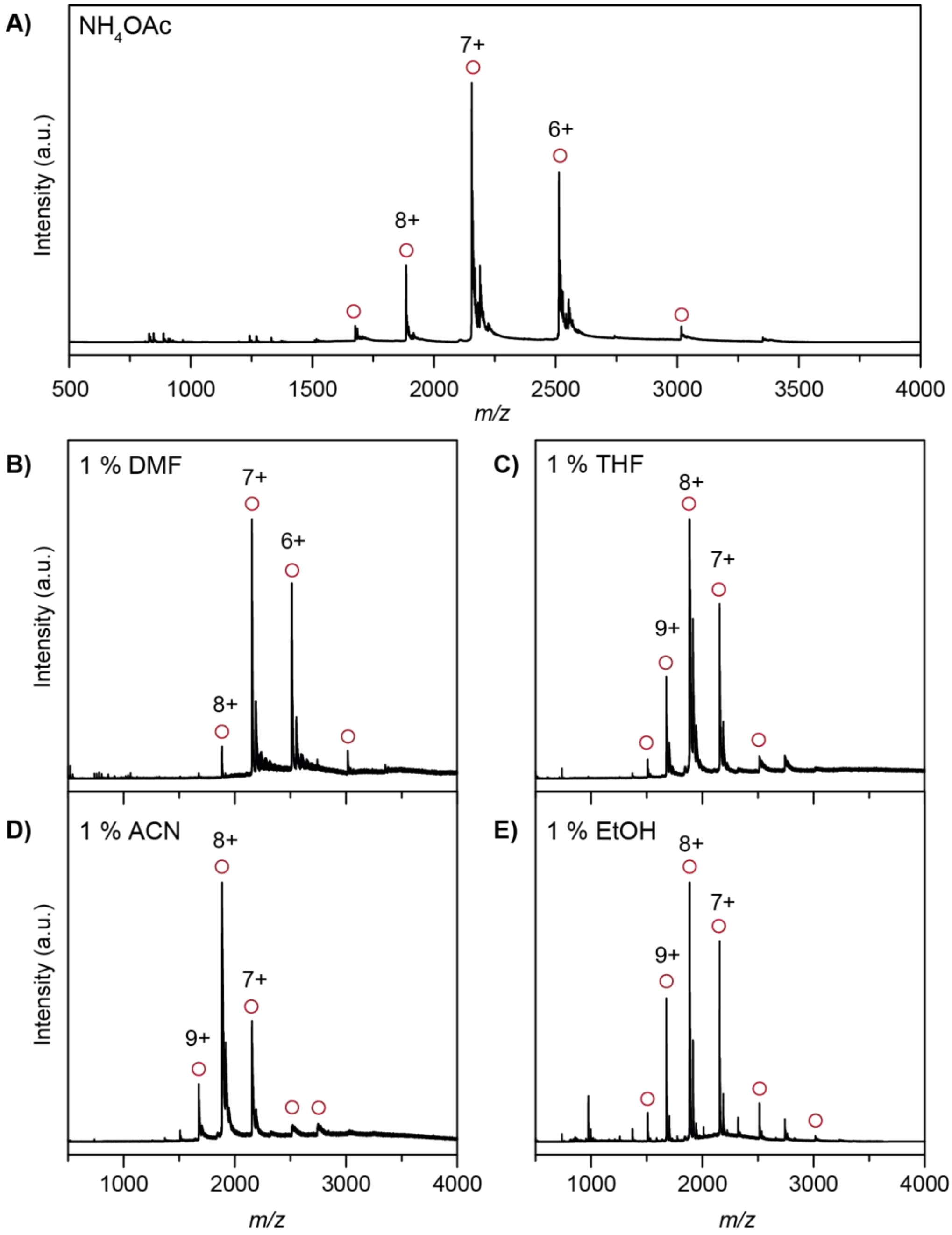


Figure 3.8: nESI mass spectra of BD1 (10 µM) in **A)** aqueous ammonium acetate with **B)** 1 % DMF, **C)** 1 % THF, **D)** 1 % ACN, and **E)** 1 % EtOH. The charge-state distribution with 1 % DMF resembles spectrum A) in ammonium acetate, whereas with the other investigated solvents, the protein adopts also higher charge states.

Consequently, DMF was selected as a first test candidate and mass spectra were recorded for protein-ligand complexes with different ligand concentrations (0, 5, 20, and 60 µM) and a DMF level of 5 % (Figure 3.9).

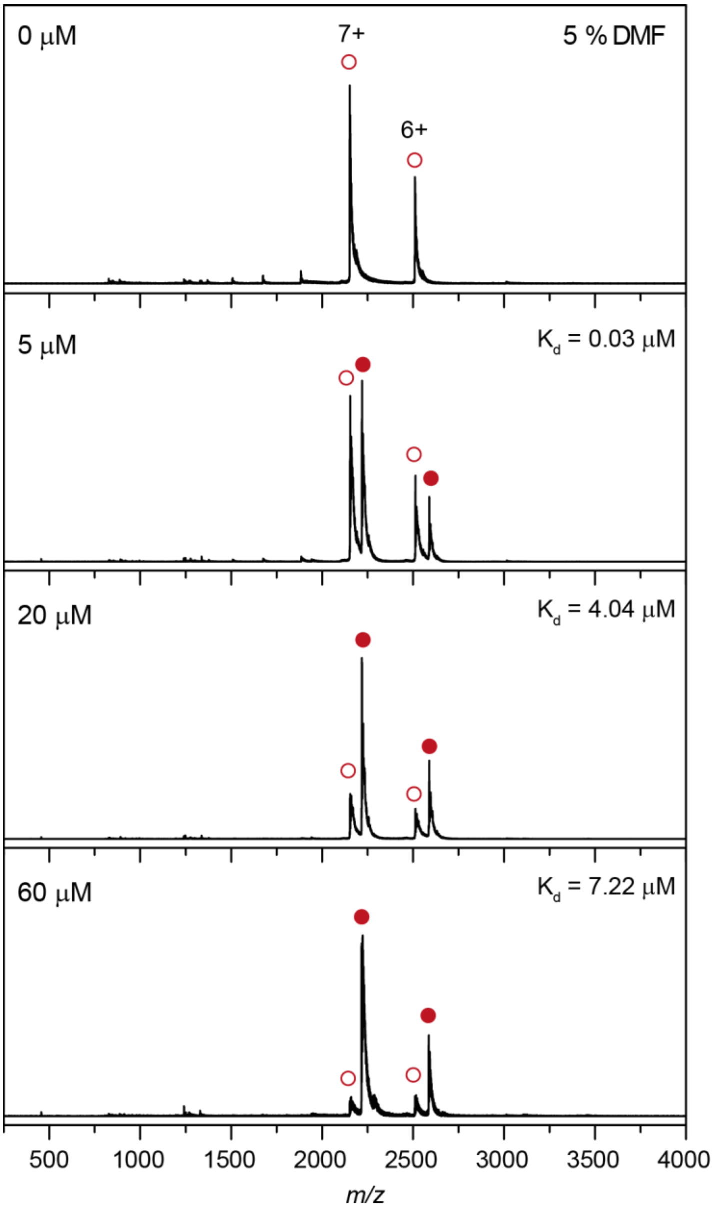


Figure 3.9: nESI mass spectra of BD1 (10 μM) with 0, 5, 20, and 60 μM (+)JQ1 in aqueous ammonium acetate and in presence of 5 % DMF. Dissociation constants K_d were determined by the direct ESI-MS assay and are indicated in every spectrum.

Charge states 6+ and 7+ of free protein (red open circles) and protein-ligand complex (red filled circles) are the main detected species in all mass spectra. A 1:1 ratio of protein-ligand complex to free protein is already reached at a ligand concentration of $[L]_0$ = 5 μM. At a ligand concentration of 60 μM nearly quantitative binding is observed, however, the complex is at this PL/P ratio saturated and higher concentrations of 100-150 μM do not result in quantitative binding (spectra not shown). Dissociation constants K_d were determined using the direct ESI-MS assay and the values are annotated in the corresponding mass spectra. For a ligand concentration of 5 μM the K_d determined by single-point measurement is 0.03 μM, which lies within the expected range (77 nM).[150] Even though quantitative binding cannot be observed, the ratio of formed complex is much higher, compared to solutions with similar DMSO concentration. This is also reflected in the K_d value: For 20 μM ligand with 5 % DMF in the spray solution a K_d of ~4 μM was determined, whereas with the same amount of DMSO a seven times lower affinity was calculated (K_d ~28 μM).

These first test experiments show clearly that the substitution of DMSO can be one possibility to prevent the effect of DMSO on protein-ligand interactions. In general, this approach is promising, if the study involves only a few ligands and properties like the solubility are well known and ideally also very similar for all ligands. However, for studies with a large number and variety of ligands, as in screening assays in the pharmaceutical industry, this approach is certainly not feasible. In this regard, the initial tests whether DMF is able to dissolve a great variety of different compounds as good as DMSO, present already an elaborate task. The easier solution at the moment is to keep the DMSO level in the final sample as low as possible. This means that the ligands either have to be diluted with buffer or water to reduce the total DMSO volume in the ligand stock solutions or very concentrated ligand solutions have to be produced, so that the added volume of ligand to the protein solution is comparatively small.

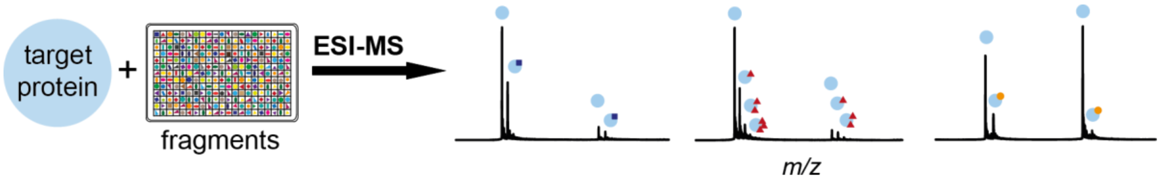
3.4 Conclusions

In this chapter the two bromodomains of BRD4, BD1 and BD2, were investigated under native conditions in the gas phase. Mass spectra in aqueous ammonium acetate solution show mainly low charge states, which indicates a native-like conformation. Addition of increasing amounts of DMSO to the protein solutions shows that DMSO has an influence on the charge-state distribution of the two proteins. While small amounts of DMSO cause a shift to lower charge states and thus indicate to a compaction of the gas-phase structure, high amounts of DMSO lead to the formation of higher charge states. The effect of DMSO has not yet been fully understood but there are several approaches to elucidate it. Certainly, factors such as the basic nature or the low vapor pressure and high boiling point contribute to the enrichment of DMSO in the electrospray droplet, which can lead to protein unfolding at higher DMSO levels. Another approach to explain the effect of DMSO could lie in the investigation of charge carrier ejection from the electrospray droplet by molecular dynamics simulations.

Protein-ligand complexes were detected with BD1 and BD2 and the ligand (+)JQ1 under native conditions with different amounts of DMSO. Here, a comparison of titration experiments with 1 and 5 % DMSO (and 10 % DMSO in case of BD1) show that ligand binding decreases significantly with increasing DMSO concentration. K_d values determined by the single-point method support this trend. Thus, the obtained data strongly indicate that DMSO has an influence on the binding affinity of BRD4-(+)JQ1 complexes. A suitable approach to overcome this effect can consist in the substitution of the solvent, in which the ligand is dissolved. For the here investigated protein-ligand system dimethylformamide was selected as an adequate solvent to substitute DMSO. With a concentration of 5 % DMF in the spray solution, the protein does not show a significant shift to higher charge-states and therefore it can be assumed that the native-like structure is maintained under these conditions. Furthermore, DMF does apparently not bind to BD1 and thus does not compete with the inhibitor. Consequently, almost quantitative binding is observed for higher ligand concentrations in the mass spectrum. The substitution of DMSO by DMF is a promising approach in case of the BRD4-(+)JQ1 system, and further tests with more compounds could evaluate the potential of this solvent for screening. However, if using DMSO for compound handling and storage, it is important to keep the concentration as low as possible. Certainly, this always depends on the solubility of the compounds and the number of necessary dilution steps.

---

4 Native Mass Spectrometry towards Fragment-Based Screening*



4.1 Introduction

In recent years fragment-based screening (FBS) has become a valuable and reliable lead finding tool that offers a promising alternative to high-throughput screening (HTS).[163, 164] In HTS very large compound libraries containing up to several millions of compounds with molecular weights of ~500 Da are screened for binding to the target protein. In contrast, FBS involves the screening of lower molecular weight fragments of ~300 Da. Due to their smaller size, fragments have a better chemical tractability and therefore probe a greater fraction of the estimated chemical space compared to larger compounds.[165] As a result, FBS requires much smaller and less complex libraries, which contain typically only a few thousand fragments and hit rates are in general higher than in HTS.[166] In turn, fragment hits typically exhibit much lower binding affinities to the target protein (high µM to mM range), which makes a reliable detection of these interactions challenging.[167, 168] Especially the necessity to use high fragment concentrations in order to detect weak interactions increases the risk of false positives, for example due to aggregation, non-specific binding or assay interference.

The most commonly applied biophysical techniques in FBS are nuclear magnetic resonance (NMR) spectroscopy,[169] X-ray crystallography,[170] surface plasmon resonance (SPR),[171] and thermal shift assay (TSA).[172, 173] To achieve accurate and reliable results, usually a cascade of screens with different methods is used.[164, 174] This typically implies a primary screen involving a high-throughput method (*e.g.* SPR or TSA), followed by an orthogonal biophysical technique for hit validation and elimination of false positives (*e.g.* NMR or isothermal titration calorimetry (ITC)). Confirmed hits are subsequently studied by X-ray crystallography to determine the exact binding mode and further optimized by structure-based methods. Modern beamlines also allow the combination of primary screening and binding-mode determination

---

[*] This chapter is based on the publication "*Critical Evaluation of Native Electrospray Ionization Mass Spectrometry for Fragment-Based Screening*", published in M. Göth *et al.*, *ChemMedChem* **2017**, *12*, 1201-1211, http://dx.doi.org/10.1002/cmdc.201700177. Figures and content adapted with permission. Copyright 2017 Wiley-VCH Verlag.

by crystal soaking.[175] In order to obtain a high affinity lead compound, fragment hits are chemically extended (fragment growing) by medicinal chemists or combined in case that two fragments concurrently occupy different regions of the binding site (fragment linking).[164, 176, 177]

The analysis of small molecules, peptides and proteins by electrospray ionization-mass spectrometry (ESI-MS) is a well-established technology in pharmacologic research *e.g.* for studying drug metabolism and pharmacokinetics, for quality control as well as for target identification and validation.[178, 179] However, conventional ESI-MS utilizes denaturing buffer and ionization conditions, which cause dissociation of non-covalent interactions. In contrast, native MS, which was introduced in Chapter 2.2, preserves the native protein conformation during the ionization process, thereby making native MS an attractive tool for the analysis of non-covalent interactions between proteins and small molecules.[44, 179] Native MS has recently also gained further interest in the field of FBS and has been employed in a few studies for screening and additional support in hit validation.[117, 168, 180-184] In a recent work, a library of 70 fragments was screened independently with native MS and SPR against human carbonic anhydrase II with a good agreement of both screening methods of 80 %.[185] This study illustrates that native MS is able to fit in the fragment-screening cascade. However, in another recent example a library containing 361 entries was screened against the enzyme endothiapepsin using six different methods including native MS and yielded only a low mutual overlap with not even one common hit identified by all six methods.[186] This work emphasizes that the comparison of different methods has to be interpreted carefully, as each technique is based on different biophysical principles and conditions.

In the present study, the potential and drawbacks of native MS in fragment-based screening should be evaluated by pointing out different challenges and obstacles that have to be faced during the experiment and data analysis. In order to get a broader view on the general feasibility and also account for protein-specific differences, four proteins covering different target classes and different molecular weights are investigated (Table 4.1). All of them are associated with epigenetic regulation or metabolism of various cancer types and therefore represent important oncological targets. A small set of fragments was screened against each of these proteins by native ESI-MS (Appendix A, Table A.1). In addition, the results are compared using TSA as a reference method, since it is widely used for fragment screening and hit validation mainly due to its high throughput, low protein consumption, and general applicability.[172, 187, 188]

Table 4.1: Overview of the investigated target proteins.

Target protein	Description	Molecular weight	No. of fragments
MTH1	8-oxo-dGTPase	18.1 kDa	33
KDM5B (catalytic core)	Histone H3K4 demethylase	55.4 kDa	16
BRPF1 (bromodomain)	Acetylated lysine binding domain protein	13.7 kDa	21
UHRF1 (SRA domain)	Hemimethylated DNA binding domain protein	23.6 kDa	21

The basic principle of TSA lies in the detection of the thermal unfolding of a protein in the presence of a fluorescent dye. Binding of a ligand typically increases the conformational stability of the target protein, which results in a dose-dependent increase of its melting temperature (T_m).[189-191] Besides the general evaluation of native mass spectrometry for FBS, it is also attempted to use the MS results to clarify the meaning of different TSA behaviors (*e.g.* destabilizers) in terms of binding to the target protein.

4.2 Experimental Details

Protein Expression and Purification*

All proteins in this study were obtained from Bayer AG and were produced using recombinant technology. N-terminal His-tagged full length MTH1 was overexpressed in E.coli and affinity purified using Ni-NTA agarose beads (Qiagem®). After removal of the His-tag by tobacco etch virus (TEV) protease treatment, the protein was further purified by size exclusion chromatography in 20 mM Tris, pH 7.5, 50 mM NaCl, 1 mM dithiothreitol (DTT). A KDM5B construct comprising amino acids 26/772 with a deletion of AT-rich interaction domain (ARID) and plant homeodomain 1 (PHD1) domain (Δ aa 102-373) and insertion of a four-glycine linker containing N-terminal His-tag was cloned and expressed in Sf9 cells. KDM5B was purified using Ni-sepharose 6 fast flow (GE Healthcare Life Sciences) and subsequently treated with TEV protease over night at 6 °C. The tag-free protein was further purified by size exclusion chromatography followed by a second Ni-affinity chromatography step in order to remove still tagged protein. The final buffer was 10 mM HEPES, pH 7.5, 150 mM NaCl, 5 % glycerol, 0.5 mM TCEP. N-terminal His-tagged bromodomain (aa 626-740) of BRPF1 was expressed in E.coli and affinity purified by HisTrap

---

* Protein expression and purification were performed by Bayer AG, Berlin.

HP sepharose (GE Healthcare Life Sciences). TEV protease treatment was performed on the Ni-column over night at 4 °C. Tag-free protein was further purified by size exclusion chromatography with buffer consisting of 20 mM HEPES, pH 7.5, 150 mM NaCl. C-terminal tagged SRA domain (aa 414-617) of UHRF1 was expressed in E.coli and affinity purified using His-Talon Superflow cartridges (Clontech®) and subsequently by size exclusion chromatography with buffer consisting of 20 mM Tris, pH 8.0, 500 mM NaCl, 10 % glycerol, and 2 mM DTT.

Thermal Shift Assay*

Fragments used in this study derived from a Bayer-internal fragment library, which has been compiled based on a filtering process of Bayer's high throughput screening (HTS) library containing more than 3.2 million compounds. The resulting library consists of ~2,000 fragments. Thermal shift assays were carried out with the ThermoFluor system (Johnson & Johnson Pharmaceutical Research and Development) or ViiA^{TM}7 Real Time PCR System (Thermo Fisher Scientific). The reactions were accomplished in a 384-well plate format with 5 µL reaction volume. The fragments were added from stock solutions in 50 mM DMSO, giving a final fragment concentration of 1 mM with a DMSO concentration of 2 %. Melting curves were recorded after heating the samples from 25 °C up to 95 °C while measuring the fluorescence intensity of the dye. The protein concentration and SyproOrange (SO) (Invitrogen) concentration was for MTH1: 4.5 µM and 5 x SO; KDM5B: 1.8 µM and 5 x SO; BRPF1: 7.3 µM and 5 x SO and UHRF1: 2.4 µM and 6 x SO. TSA data were analyzed using the software Analyzer (Genedata Screener^{R}).

Sample Preparation and Mass Spectrometry

The proteins were buffer exchanged at least four times prior to native MS experiments by PD Spin Trap G-25 columns (GE Healthcare Life Sciences) with 50 mM aqueous ammonium acetate solution (pH = 6.5 - 7.0). The concentration was determined *via* NanoDrop2000 (Thermo Scientific) at 280 nm wavelength with an extinction coefficient calculated by the software GPMAW (Lighthouse data). The protein concentration in the final samples was 10 µM. Fragments were dissolved in DMSO to obtain a 100 mM stock solution and further diluted with 50 mM ammonium acetate to the desired final concentration. The fragment stock solutions and the final sample solutions (10 µL consisting of protein, fragment and buffer)

---

* Thermal shift assay was performed by Elisa Chimik and Dr. Jörg Weiske at Bayer AG, Berlin.

were freshly prepared for each measurement and pipetted into 384-well plates. The used fragment concentrations for the screening experiments were as follows: MTH1: 100 μM and diverse concentrations for titration experiments, KDM5B: 500 μM, BRPF1: 100 and 500 μM, UHRF1: 200 μM. nESI-MS was carried out on a Synapt G2-S mass spectrometer (Waters), equipped with a Triversa NanoMate chip-based nESI source (Advion Biosciences). The Advanced User Interface was used to create an automated screening protocol and data acquisition was automated by connecting the NanoMate with the MassLynx software of the Synapt using a contact closure signal. The measurements were performed in positive ion mode (resolution) with a chip voltage of 1.7 kV and a spray pressure of 0.7 psi (0.48 bar). The scan time was set to 1 s and each sample was acquired for 30 s. Parameters were carefully adjusted to minimize in-source dissociation. Typical parameters were as follows: sampling cone voltage 10-50 V; source offset 20 V; source temperature 30 °C; trap collision voltage 0-2 V, transfer collision energy 2 V; trap DC entrance 0 V, trap DC bias 2 V, ion mobility cell switched off.

Data Analysis

Manual data analysis was performed with Excel (Microsoft Corporation) and OriginPro 8.6 (OriginLab Corporation) software. The fraction of bound complex PL was calculated using the following expression[192]

$$\%\,\mathrm{PL} = [\mathrm{I_{PL}} / (\mathrm{I_P} + \mathrm{I_{PL}})]*100 \tag{4.1}$$

where I_{PL} and I_P correspond to the peak intensities (peak heights) of complex PL and bare protein P, respectively. Unless otherwise noted, all charge states of P and PL were considered, such that I_{PL} and I_P correspond to the sum of intensities of the individual charge states. Peak intensities are substituted for concentrations, as the total (initial) protein concentration is known. Automated data analysis was carried out with the BiopharmaLynx software (Waters). Mass spectra were deconvoluted using the MaxEnt1 algorithm and the fraction of bound protein was determined from the deconvoluted spectra (MassLynx 4.1). The binding affinity can be quantified either with the single-point method or with the titration method, which were discussed in Chapter 2.4.1. In both approaches it is assumed that the PL complexes have similar ionization, transmission and detection efficiencies compared to the bare protein P and thus the introduction of a response factor is not necessary.[99] For a 1:1 protein-ligand dissociation constant K_d by single-point method can be estimated using Eq. 2.9, where [P] and

[PL] are determined from the peak intensities I_P and I_{PL} and the initial concentration of protein $[P]_0$.[168, 193]

$$[P] = [P]_0 * I_P / (I_P + I_{PL}) \tag{4.2}$$

$$[PL] = [P]_0 * I_{PL} / (I_P + I_{PL}) \tag{4.3}$$

For the determination of the dissociation constant K_d *via* an ESI-MS titration experiment, the protein concentration is kept constant while the ligand concentration is varied. [PL] is then plotted *versus* the initial ligand concentration $[L]_0$ and Eq. 3.2 (Chapter 3.2) is applied for K_d determination by curve fitting.

4.3 Results

4.3.1 Fragment-Based Screening with Native Mass Spectrometry

In the first section of this chapter the focus lies on the results that were obtained from a native MS-based fragment screen against full length MTH1. This protein, also known as NUDT1 (7,8-dihydro-8-oxoguanine triphosphatase), binds and hydrolyzes oxidized purine nucleoside triphosphates and thus prevents incorporation of damaged bases during DNA replication. In normal cells MTH1 does not play a key role, but it is vital for the survival of tumor cells. In this context it has been discovered that elimination of MTH1 leads to death of tumor cells.[194] Other than eliminating the protein, inhibition of the hydrolysis of damaged bases could be an alternative way to limit cancer growth. The here investigated protein MTH1 consists of 158 amino acids and has a molecular weight of 18.1 kDa (crystal structure is shown as an inset in Figure 4.1A).

Figure 4.1A shows the mass spectrum of MTH1 (10 μM) sprayed from 50 mM ammonium acetate solution. The spectrum contains a narrow charge state distribution with charge states ranging from 7+ to 9+, indicating that the protein retains a folded, native-like conformation.[195]

Usually, small molecules in drug discovery are stored in DMSO due to its favorable physicochemical properties. Inescapably, a certain amount of DMSO remains in the final sample solution. High DMSO concentrations (starting from ~10 %) lead to supercharging, subsequent denaturation and unfolding of the protein.[147]

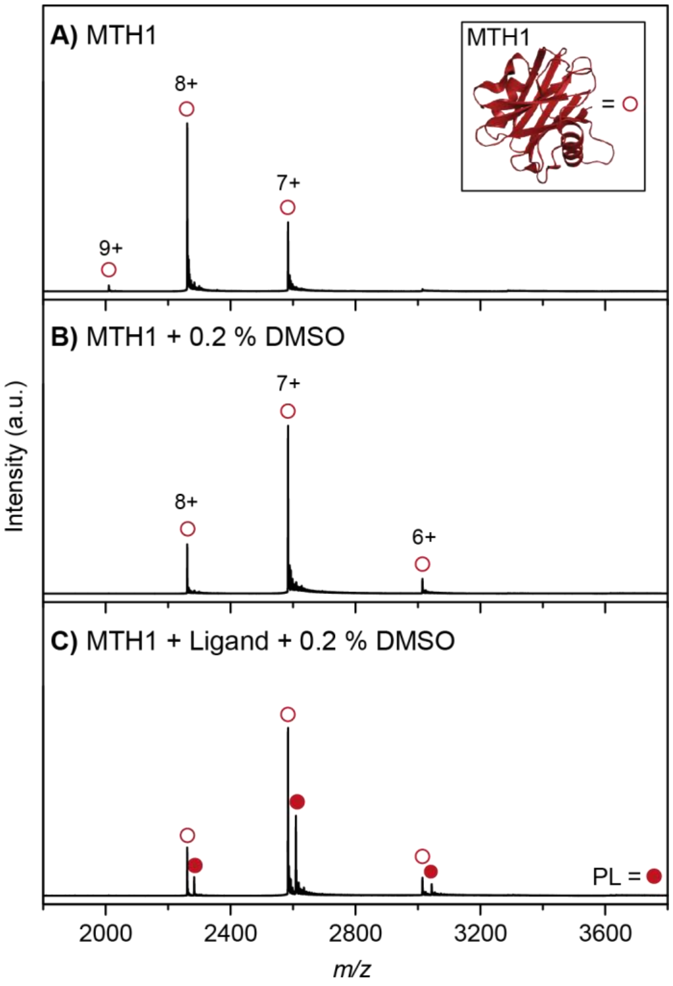


Figure 4.1: nESI mass spectra of 10 µM MTH1 (red open circle) in 50 mM ammonium acetate without DMSO (**A**), in presence of 0.2 % DMSO (**B**), and in presence of 100 µM ligand and 0.2 % DMSO (**C**), forming a protein-ligand complex (PL, red filled circle). The crystal structure of MTH1 is shown as inset in (A) (PDB entry: 4C9W).

However, as it was also observed in Chapter 1, even low amounts of DMSO might influence the binding affinity of a protein-ligand complex.[146] Hence, DMSO concentrations in this study were kept as low as possible. In presence of 0.2 % DMSO, MTH1 retains a narrow charge state distribution, indicating that the native fold remains under these conditions (Figure 4.1B). However, a strong decrease of the average charge state from 7.7+ to 7.2+ is observed upon DMSO addition. This effect is most likely caused by a DMSO induced compaction of the protein structure and has been observed previously.[147]

Next, native MS was used to screen a small set of fragments against MTH1 that exhibited stabilizing, destabilizing, or neutral effects in a TSA primary screen. A protein concentration of 10 µM and a tenfold molar excess of fragments was employed. Data acquisition was straightforward by using an automated chip-based nanoESI source,[196] and all measurements were carried out in triplicate within less than one hour (~0.5 min per sample). Figure 4.1C shows a typical mass spectrum of a fragment hit, which is easily identified by additional signals that correspond to the protein-ligand complex (PL).

When screening a large amount of fragments by native MS, a convenient way to rank hits by their binding affinity has to be established. As demonstrated by various groups, native MS in principle can be used for reliable determination of absolute binding affinities.[93, 168, 197] However, this implies a titration experiment for each ligand using a fixed protein concentration and several different ligand concentrations. Such an experiment was carried out exemplarily for MTH1 with a specific binder (fragment **M30**) and Figure 4.2A shows the corresponding mass spectra for five different fragment concentrations ($[L]_0 = 0$, 4, 10, 50, and 200 µM). Assuming that MTH1 possesses a single ligand binding site, a specific binder is expected to only form a 1:1 protein-ligand complex (PL). Accordingly, with increasing concentration of fragment **M30** in the initial solution, the signal intensity of PL gradually increases (Figure 4.2A, red filled circles). Since in this case no multiple, non-specific adducts are visible in the spectra, the peak intensities of PL can be substituted for concentrations [PL] and are plotted in Figure 4.2B against the ligand concentration $[L]_0$.

Curve fitting using Eq. 3.2 (Chapter 3.2) results in a dissociation constant (K_d) of 1.73 ± 0.16 µM. Remarkably, the amount of formed complex saturates at ~60-70 %, indicating that about one third of the target protein obviously does not take part in binding. A much easier way for fragment ranking is to calculate the dissociation constant from a single-point measurement (Eq. 2.9, Chapter 2.4). However, single-point estimation might give diverging K_d values as exemplarily highlighted for two ligand concentrations (1.69 ± 0.15 µM at $[L]_0 = 2$ µM or 116 ± 4 µM at $[L]_0 = 200$ µM) (Figure 4.2B). This illustrates that the K_d determined by single-point estimation underestimates the affinity for strong binders in case the target protein is already saturated. Weak binders, in turn, need high fragment concentrations to detect significant amount of complex and determine a reliable K_d. Since binding affinities within a large fragment library usually cover a wide range of affinities, the calculation of reliable K_d values using a fixed ligand concentration is not straightforward.

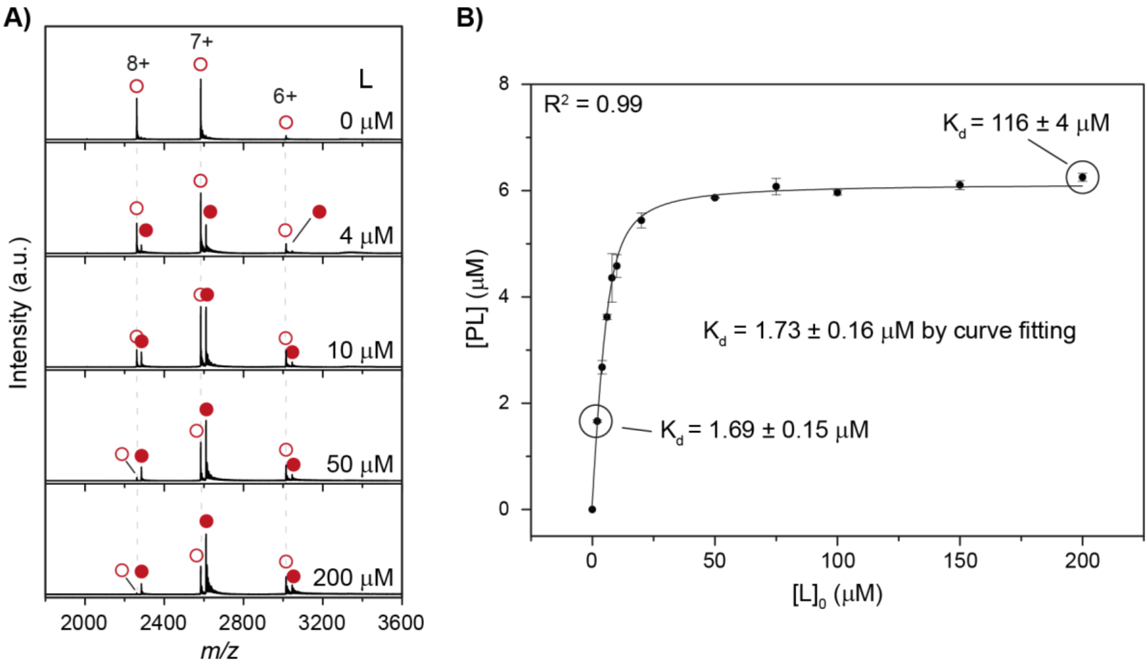


Figure 4.2: A) ESI-MS titration using a constant protein concentration (MTH1, 10 µM, red open circles) and increasing ligand concentration (**M30**, 0 to 200 µM, complex formation red filled circles) from top to bottom. **B)** The amount of protein-ligand complex [PL] is plotted as a function of increasing ligand concentration $[L]_0$. The affinity was determined by curve fitting (Eq. 3.2) and by single-point estimation (Eq. 2.9). Data represent the mean ± standard deviation of measurements in triplicate.

A third alternative that was also employed within this chapter, is relative affinity ranking, which involves the calculation of the fraction of protein-ligand complex (% PL) under the exact same experimental and instrumental conditions. It can easily be calculated by reading out the intensities (peak heights) of the bare protein (P) and protein-ligand complex (PL) in the mass spectrum (Eq. 4.1).

Following this workflow, combined with automated data acquisition, native MS appears to be a fast and well suited method for fragment screening. However, a closer inspection of mass spectra of protein-fragment complexes from the screening experiment illustrates that data interpretation is not necessarily always straightforward, since the ideal case of specific 1:1 PL complex formation (as exemplarily shown in Figure 4.2) between a protein and a ligand might not be given. Figure 4.3B shows a fragment that forms not only a 1:1 PL complex with MTH1, but also multiple, presumably non-specific adducts (PL_2, PL_3, PL_n). The formation of non-specific adducts during the ionization process is a common phenomenon in native MS,[94] especially when high ligand concentrations are used as for the detection of low affinity fragments. Since this effect is concentration-dependent, a reduction of ligand concentration can help to eliminate non-specific binding (Appendix A, Figure A.1A).

Another phenomenon that was observed during fragment screening is a fragment-induced change in the charge-state distribution of protein signals. Figure 4.3C shows a mass spectrum of MTH1 with a fragment that induces such a charge-state shift (css) in addition to formation of a direct protein-ligand complex. Compared to the bare protein (Figure 4.3A) a significant increase in the intensity of charge state 6+ and even formation of charge state 5+ is observed. An ESI-MS titration with this fragment illustrates that this effect is concentration-dependent (Figure A.1B). The css can be explained by complex formation in solution, but dissociation of the charged fragment directly after ionization during passage through the mass spectrometer, therefore reducing the protein's charge state (Figure 4.4). As shown by Blackburn *et al.*, quantifying the css can be used to calculate an apparent $K_{d.}$[198] However, using the css as a measure of binding exceedingly complicates data analysis, especially if direct complex formation occurs on top of css, as observed here. Therefore, the css was not included for calculation of % PL but the respective fragments were just annotated.

Table 4.2 summarizes the determined % PL values from the native MS screen against MTH1. Fragments that induce a css or show extensive multiple binding are annotated accordingly. Remarkably, by using a concentration of 100 μM, all tested 33 fragments show complex formation with the target protein ranging from 6 % up to 66 %. Nine out of 33 fragments induce a css (**M08**, **M11**-**M15**, **M25**, **M26**, **M31**), while three fragments exhibit multiple binding (**M10**, **M16**, **M21**). A comparison of manual and automated software-based data analysis using BiopharmaLynx shows similar results for % PL determination (Appendix A, Table A.2). Since the software uses deconvoluted spectra for determination of % PL, information about potential charge-state shifts is lost.

A major difficulty was to determine a threshold of % PL that defines a hit. Especially fragments that exhibit only low complex formation have a higher chance of being false positives, because signals that would be indicative of non-specific binding (PL_2, PL_3, PL_4) might disappear in background noise, making the discrimination between specific binding and non-specific adduct formation difficult. To find a compromise between excluding most non-specific binders and including specific, but low-affinity hits, a threshold of 20 % complex formation was set to define fragment hits. Multiple binders were not considered as hits, although two of them (**M10**, **M16**) show binding above the set threshold. Using these restrictions, 16 hits are identified by native MS (Table 4.2).

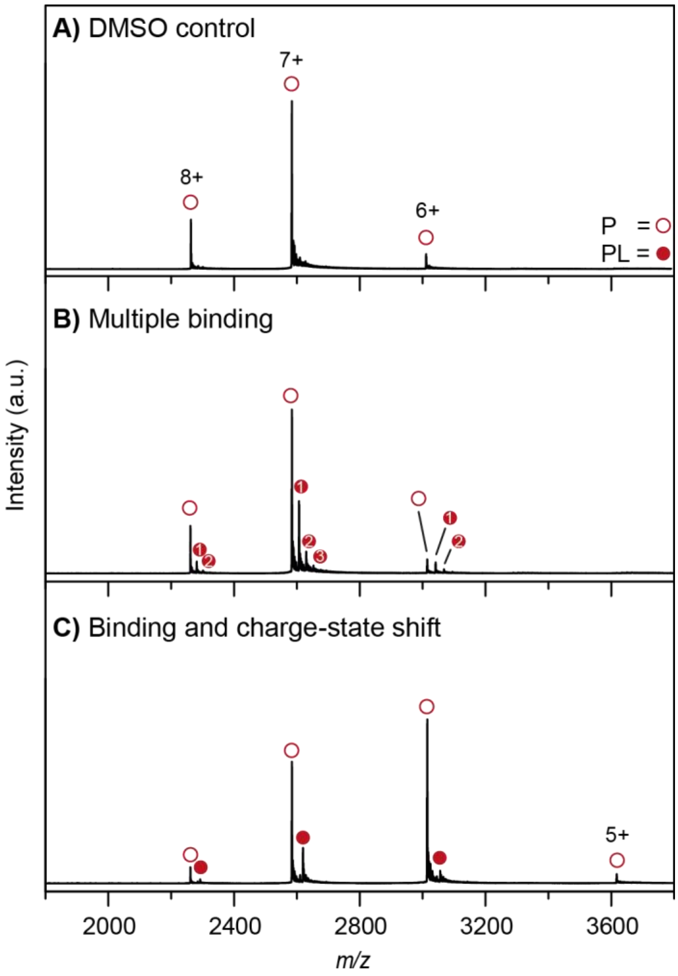


Figure 4.3: nESI mass spectra of 10 µM MTH1 (P, red open circles) in 50 mM ammonium acetate in presence of 0.2 % DMSO **(A)** and after addition of 100 µM ligand **M16** **(B)** or 100 µM ligand **M26** **(C)**. In addition to specific PL formation, other types of complex formation are observed that hamper automated data analysis. The number of attached ligands is indicated in mass spectrum B).

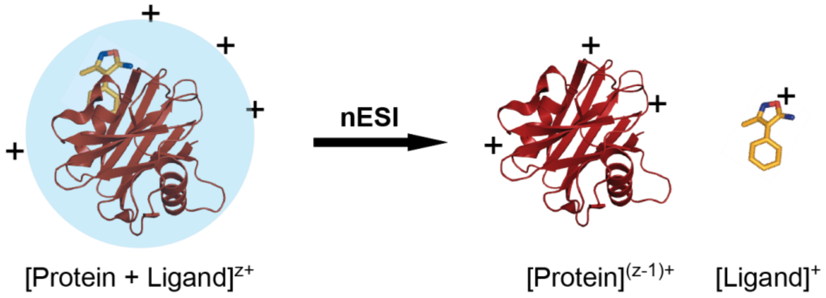


Figure 4.4: Schematic explanation for a charge-state shift in the mass spectrum. The protein-ligand complex forms in solution, but does not survive the transfer into the gas phase. As a consequence the complex dissociates and the ligand takes a charge (PDB entry: 4C9W).

Table 4.2: Comparison of MS results with the orthogonal technique TSA for a screen of 33 fragments against MTH1 (10 µM protein and 100 µM fragment).

Fragment ID[a]	MW (Da)	PL (%)[b]	charge-state shift	ESI-MS hit[c] (% PL > 20 %)	TSA hit[d] (stabilizer)	TSA ΔT_m (°C)	TSA category
M01	232.13	30	no	yes	yes	7.15	stabilizer
M02	254.78	24	no	yes	yes	5.56	stabilizer
M03	224.75	9	no	no	yes	3.27	stabilizer
M04	301.25	7	no	no	no	0.51	neutral
M05	177.24	33	no	yes	yes	3.49	stabilizer
M06	199.25	7	no	no	no	0.51	neutral
M07	212.25	8	no	no	no	-1.83	destabilizer
M08	186.25	12	yes	no	yes	8.07	stabilizer
M09	196.29	15	no	no	yes	1.63	stabilizer
M10	240.34	46	no	no	yes	7.87	stabilizer
M11	247.12	37	yes	yes	yes	9.13	stabilizer
M12	199.25	53	yes	yes	yes	9.94	stabilizer
M13	240.34	52	yes	yes	yes	10.6	stabilizer
M14	235.28	20	yes	yes	yes	2.03	stabilizer
M15	295.20	20	yes	yes	yes	5.77	stabilizer
M16	160.18	27	no	no	yes	7.83	stabilizer
M17	187.24	12	no	no	yes	10.37	stabilizer
M18	170.21	22	no	yes	yes	10.04	stabilizer
M19	222.69	12	no	no	no	-2.09	destabilizer
M20	163.22	19	no	no	yes	6.50	stabilizer
M21	246.35	17	no	no	no	-1.85	destabilizer
M22	207.23	12	no	no	no	0.52	neutral
M23	236.14	36	no	yes	yes	4.41	stabilizer
M24	244.33	7	no	no	no	0.53	neutral
M25	209.25	37	yes	yes	yes	8.53	stabilizer
M26	249.31	15	yes	no	yes	8.35	stabilizer
M27	268.11	10	no	no	no	0.53	neutral
M28	215.25	10	no	no	no	-1.77	destabilizer
M29	248.73	48	no	yes	yes	8.78	stabilizer
M30	221.71	66	no	yes	yes	13.22	stabilizer
M31	186.25	35	yes	yes	yes	13.00	stabilizer
M32	249.18	23	no	yes	yes	8.06	stabilizer
M33	228.76	34	no	yes	yes	8.09	stabilizer

[a] Underlined fragments show multiple binding. [b] % PL determined by manual data analysis, where every charge state was considered. The value corresponds to the average of measurements in triplicate. [c] Fraction of protein of the manual data analysis was used for hit definition. A threshold of 20 % PL formation was set as hit criterion. [d] Stabilizing fragments were considered as hits, if ΔT_m exceeds the protein's melting temperature by six times the standard deviation.

4.3.2 Comparison of Fragment Hits from MS and TSA

To validate the fragment hits from the native MS screen, they were compared to the data obtained from a prior TSA screen. In general, ligands can cause an increase (stabilizers) or decrease (destabilizers) of the protein's melting temperature (T_m) or have no effect (neutrals). Fragments are considered as stabilizers or destabilizers if ΔT_m exceeds or lowers the protein's melting temperature by six times the standard deviation. The set of 33 fragments selected for MTH1 consists of 24 stabilizers, five neutrals and four destabilizers (Table 4.2). While stabilizers are considered as specific binders and as such are classified as hits, the cause of destabilization is not exactly known and might result from non-specific interactions with the target protein. A comparison of the hits from TSA and native MS shows a good agreement: 16 of the 24 stabilizing fragments are also identified as hits by native MS, which corresponds to an overlap of 67 %. In accordance, the best binder in MS (fragment **M30**, PL = 66 %) also induces the greatest temperature shift in TSA (> 13 °C), with a tendency of higher stabilizers also being good binders in MS. Destabilizers show rather low complex formation (< 12 %), while only one of the four tested destabilizing fragments exhibits multiple binding to the target protein (**M21**).

Taken together, the comparison of results from native MS with the reference method TSA emphasizes that fragment screening with native MS is able to yield reliable hits up to a certain affinity threshold at which specific binding can hardly be distinguished from non-specific adduct formation.

4.3.3 Investigation of Other Target Protein Classes

Since the results from the fragment screen with MTH1 were promising, the study was expanded to three other target proteins: KDM5B, BRPF1, and UHRF1 (Table 4.1). The lysine(K)-specific demethylase 5B (KDM5B, also JARID1B) is a multi-domain enzyme that belongs to the family of histone demethylases. This protein is responsible for demethylation of tri-, di-, and mono-methylated lysines in position 4 of histone 3 (H3K4me3, H3K4me2, H3K4me). Although the specific role of KDM5B is yet unknown, it has been associated in the development of various cancer types and potentially plays a major role there.[199] Within this chapter the catalytic core of KDM5B (Figure 4.5A) is investigated using native mass spectrometry.

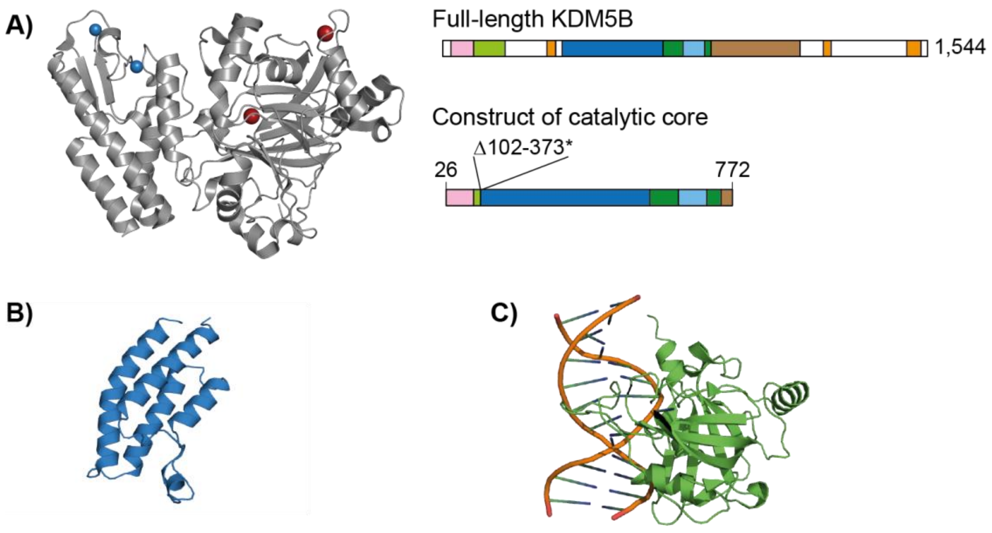


Figure 4.5: A) Crystal structure of KDM5B (PDB entry: 5FPU) with two zinc and manganese ions (blue and red spheres) and schematic sequence of the full length protein as well as the construct of the catalytic core that was investigated within this work. **B)** Right: Crystal structure of the bromodomain of BRPF1 that was investigated within this work (PDB entry: 4LC2). **C)** Crystal structure of the set and ring associated (SRA) domain of UHRF1 bound to methylated DNA (PDB entry: 3CLZ).

The bromodomain and plant homeodomain (PHD) finger containing protein 1 (BRPF1) forms multi-subunit protein complexes with two coactivators and connects these to other subunits (Figure 4.5B). Translocations of one coactivator are known to be in association with aggressive subtypes of leukemia. Whereas the bromodomain and extra terminal (BET) subfamily has been extensively studied and is well-understood, this is not likely the case for other bromodomain-containing proteins such as BRPF1. Due to their multidomain architecture, it is challenging to assign observed functions specifically to the bromodomain.[200] Within this study, fragment-complexes with the 116 amino acid BRPF1 bromodomain (blue crystal structure, Figure 4.5B) are investigated.

Ubiquitin-like PHD and really interesting new gene (RING) finger containing-domain protein 1 (UHRF1) binds to specific methylated DNA sequences (Figure 4.5C) and is involved in chromatin structure regulation and gene expression.[201] It has been identified as an oncogene in various cancers and therefore become a therapeutic target.[202]

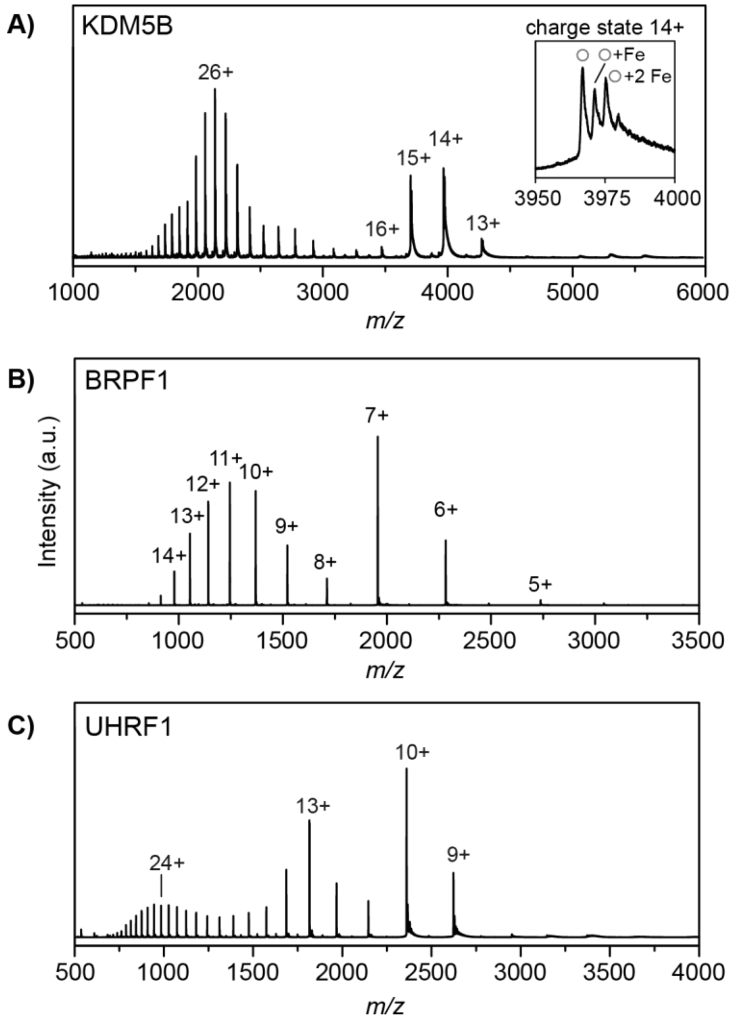


Figure 4.6: nESI mass spectra of 10 µM **A)** KDM5B, **B)** BRPF1, and **C)** UHRF1 in 50 mM ammonium acetate. All three proteins exhibit more than one charge-state distribution, which indicates that under these conditions not only native-like conformations (higher *m/z* values), but also partially unfolded structures (lower *m/z* values) are present.

For each target protein again a small set of fragments was chosen from a TSA primary screen that includes destabilizers, stabilizers and neutrals (see Table A.1 for physicochemical properties). Figure 4.6 shows the mass spectra of the three target proteins KDM5B, BRPF1, and UHRF1 in ammonium acetate. In contrast to the spectrum of MTH1 (Figure 4.1), all proteins exhibit more than one charge-state distribution. This indicates that not only native-like conformations are present (charge-state distribution at higher *m/z* values), but also partially unfolded structures (charge-state distribution at lower *m/z* values) are adopted under these conditions.[53]

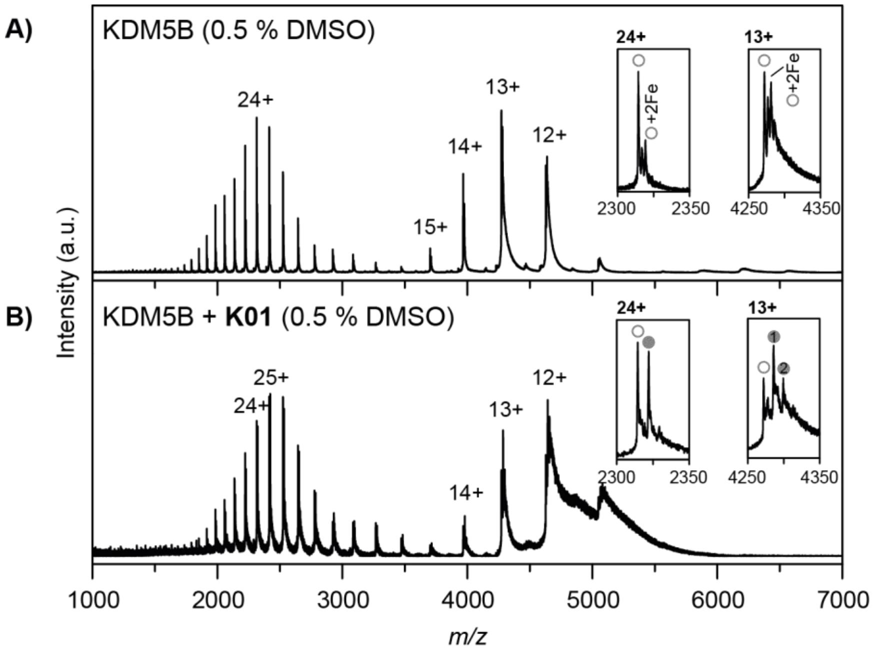


Figure 4.7: nESI mass spectra of 10 µM KDM5B in 50 mM ammonium acetate and in presence of 0.5 % DMSO **(A)** and after addition of 500 µM ligand **K01** **(B)**. Fragment addition induces extensive peak broadening and multiple binding, especially on native-like charge states 11+ to 14+ (bare protein, grey open circles; protein-ligand complex, grey filled circles).

For KDM5B - the largest protein in this study - much broader and less resolved peaks were observed, especially at native-like charge states (13+ to 16+) (Figure 4.6A). This can be explained by increased adduct formation at higher *m/z* values as well as an increased heterogeneity due to coordination of two zinc ions and one iron ion in the catalytic center of KDM5B.[203]

Addition of fragments causes multiple complex formation and extensive peak broadening, therefore impeding determination of complex formation at native-like charge states (Figure 4.7). Instead, in this particular case, a defined set of sufficiently resolved signals corresponding to the partially unfolded protein (17+ to 28+) was used for detection and quantification of fragment binding. This practice is acceptable only if the partially unfolded protein species retains its ligand-binding properties, indicating that the catalytic center is still in its functional conformation. This is likely the case for KDM5B, since the partially unfolded species is still bound to zinc and catalytic iron ions and also fragment binding is observed as for the native-like charge states (Figure 4.7). In addition, crystal structures of this construct in complex with three inhibitor chemotypes were solved in a previous study,[203] which further demonstrates that this construct in general is suitable for investigation of ligand binding.

Results for identification and quantification of fragment binding to KDM5B are summarized in Table 4.3. When using the same hit criteria as mentioned above, ten hits are identified out of 16 fragments. A comparison of the hits from TSA and native MS again shows a good agreement. Eight of the nine stabilizing fragments are also identified as hits by native MS. Again, the best binder from the native MS screen (fragment **K09**, PL = 57 %) also induces the greatest temperature shift in TSA (~5 °C). Only three fragments show multiple binding and all of them were identified as destabilizers in TSA. However, the group of seven destabilizing fragments also included two specific hits in native MS.

Table 4.3: Comparison of MS results with the orthogonal technique TSA for a screen of 16 fragments against the catalytic domain of KDM5B (10 µM protein and 500 µM fragment).

Fragment ID[a]	MW (Da)	PL (%)[b]	charge-state shift	ESI-MS hit[c] (% PL > 20 %)	TSA hit[d] (stabilizer)	TSA ΔT_m (°C)	TSA category
K01	174.20	38	yes	yes	yes	2.33	stabilizer
K02	173.17	36	yes	yes	yes	1.16	stabilizer
K03	188.18	38	yes	yes	yes	2.59	stabilizer
K04	214.26	24	yes	no	no	-3.00	destabilizer
K05	204.29	16	yes	no	no	-4.20	destabilizer
K06	207.27	26	yes	yes	yes	1.54	stabilizer
K07	203.21	10	yes	no	yes	3.55	stabilizer
K08	189.26	30	yes	yes	no	-4.19	destabilizer
K09	165.19	57	yes	yes	yes	4.98	stabilizer
K10	225.29	16	no	no	no	-3.12	destabilizer
K11	189.21	39	yes	yes	yes	1.50	stabilizer
K12	211.26	21	yes	no	no	-3.13	destabilizer
K13	197.66	23	yes	yes	no	-4.21	destabilizer
K14	191.23	32	yes	yes	yes	3.04	stabilizer
K15	179.22	39	no	yes	yes	1.94	stabilizer
K16	263.74	0	no	no	no	-4.20	destabilizer

[a] Underlined fragments show multiple binding. [b] % PL determined by manual data analysis, where only partially unfolded charge states (17+ to 28+) were considered. [c] A threshold of 20 % PL formation was set as hit criterion. [d] Stabilizing fragments were considered as hits, if ΔT_m exceeds the protein's melting temperature by six times the standard deviation.

Therefore, non-specific binding does not necessarily seem to be the cause of destabilization in TSA. Although the heterogeneity induced by complexation of iron ions within the catalytic center of KDM5B complicated data interpretation, it opened up the possibility to analyze the iron-dependency of ligand binding (Appendix A, Figure A.3). While most of the 16 fragments bind only in presence of the active site iron (potential iron coordinators), some fragments only show binding to the iron-free form (potential iron competitors). Some fragments also exhibit binding to both forms and therefore it can be assumed that they bind independent of the catalytic iron. Although the spectrum of the apo-form of KDM5B indicates binding of up to two iron ions, only one iron ion seems to be relevant for ligand binding (Appendix A, Figure A.2). This is in agreement with the observation of one active site Mn^{2+} that was used to replace the active site iron in the crystal structure.[203] Taken together, the possibility to elucidate these different binding behaviors is very interesting for creating a structure-activity relationship, however, it requires careful and manual analysis of the mass spectra.

The results of fragment screening for BRPF1 and UHRF1 are summarized in Table A.3 and A.4. For BRPF1 none of the 21 fragments reaches the hit criteria of 20 % complex formation, although eleven fragments were identified as hits by TSA including six fragments that were confirmed by X-ray crystallography (Appendix A, Table A.3). Therefore, the experiment was repeated by increasing the fragment concentration from 100 μM to 500 μM as exemplified in Figure 4.8. Although weak binding (PL < 5 %) is now observed for five of the six X-ray confirmed fragments (Appendix A, Table A.3), excessive multiple binding prevents to distinguish specific binding from non-specific adduct formation. Consequently, none of the tested fragments fulfills the hit criteria.

For UHRF1 a set of 21 fragments including ten TSA hits and eleven neutrals was screened by native MS. Only one of the tested fragments (**U01**) reaches the native MS hit criterion of at least 20 % complex formation (Appendix A, Table A.4). However, this fragment is a neutral in TSA, so that there is no mutual overlap between the two methods. In this regard it is worth mentioning that TSA hits only induced minor shifts in melting temperature ($\Delta T_m \leq 1$ K) and they therefore might constitute low affinity binders that escape detection by native MS. Noteworthy, 12 fragments induce a charge-state shift, with seven of them exclusively showing this effect without any direct binding, as exemplarily shown in Figure 4.9 with fragment **U04**. A correlation between fragments that induce a charge-state shift and TSA hits was not observed.

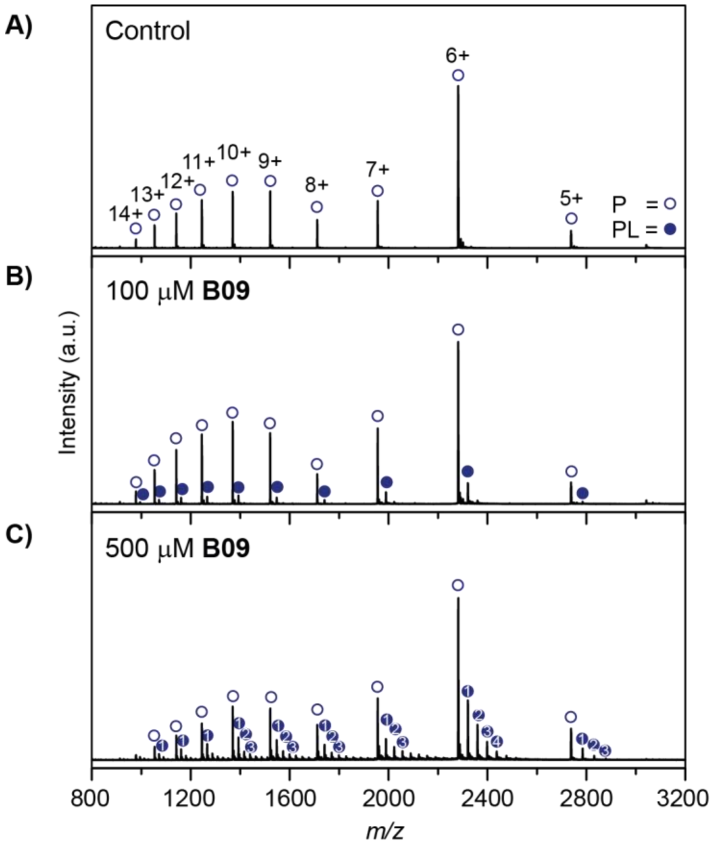


Figure 4.8: nESI mass spectra of **A)** BRPF1: DMSO control spectrum and **B)** in presence of 100 µM and **C)** 500 µM of fragment **B09**. The bare protein is indicated with blue open circles, whereas the complex is indicated with blue filled circles. Small numbers display the amount of bound ligands. Slightly more binding is observed with increased fragment concentration, but also excessive formation of multiple, presumably non-specifically bound species.

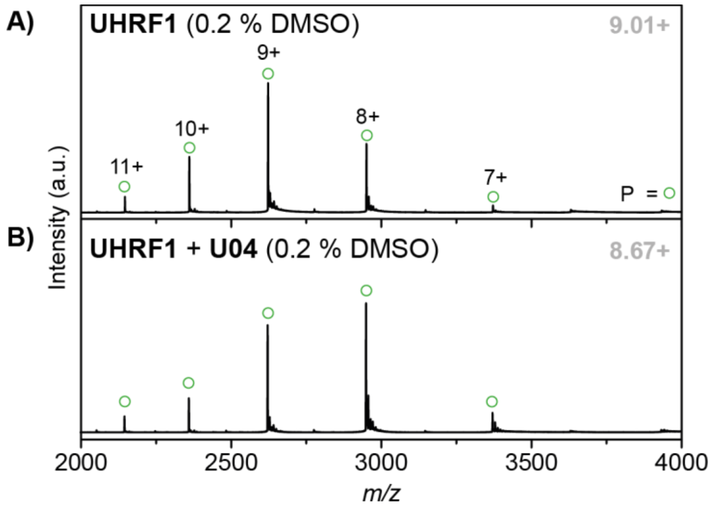


Figure 4.9: nESI mass spectra of the SRA domain of UHRF1 **A)** in presence of 0.2 % DMSO and **B)** in presence of fragment **U04** and 0.2 % DMSO. The fragment induces a charge-state shift in the mass spectrum, which likely reflects binding in solution. The calculated average charge is given as an inset in grey.

4.4 Discussion

The evaluation of native MS for FBS using four different target proteins clearly demonstrates that the applicability strongly depends not only on the nature of the target protein, but also on the experimental and instrumental conditions. In this regard, the determination of a suitable fragment concentration is a very critical factor when planning a screening campaign. Before conducting a primary screen there is often no previous knowledge about the expected binding affinities and hit rates.

Finding an optimal fragment concentration is therefore often difficult. Choosing rather low fragment concentrations is generally to be favored in order to reduce non-specific binding, but also carries the risk of missing low-affinity ligands. This might be acceptable if the overall hit rate is high enough as shown for the investigated proteins MTH1 and KDM5B. If the hit rate is low, however, increasing the fragment concentration increases the chance for detection of low-affinity binders, but also increases the risk of non-specific adduct formation. The here investigated bromodomain of BRPF1 shows exactly this problem: While no hits were identified in the initial screen, increasing the fragment concentration slightly enhances binding, but also (and to a larger extent) increases non-specific adduct formation. In this case one faces the problem of how to differentiate between specific and non-specific interactions. Diverse experimental approaches can be applied to measure the contribution of non-specific binding, such as the reference protein method.[99] In addition, several methods that use sophisticated models for data analysis have been proposed for deconvolution of specific and non-specific binding.[100-103] However, these approaches often require detailed manual analysis, since they are not yet implemented in currently available software packages for automated spectral processing. Thus, in the case that excessive multiple binding is observed during a screening campaign, native (ESI-)MS might not be the best suited method for FBS in comparison to other technologies such as SPR or NMR spectroscopy.

While multiple binding is assumed to be caused mainly by non-specific adduct formation and is a phenomenon generally observed during electrospray ionization, another major challenge for data interpretation was a fragment-induced charge-state shift. Remarkably, the tendency for multiple adduct formation as well as charge-state shifts varied significantly between different target proteins tested in this study. Charge-state shifts can be indicative of weak interactions in solution that are lost directly after ionization.[198] What does this mean for data interpretation? Simply ignoring this effect might lead to an underestimation of binding affinities or even to false negatives. This is probably the case for KDM5B, where seven of the tested fragments induce a charge-state shift without additional binding. However, as soon as charge-state shifts and specific binding is observed simultaneously (as for MTH1 and KDM5B), data interpretation is challenging. One can be optimistic that advancements in software will soon allow for the automated analysis of fragment binding also in cases where manual data inspection is currently required.

A further question that arises from our experiments is how to rank hits according to their apparent affinity. In previous native MS studies, fragments were ranked either by their dissociation constant or classified in different binding categories (A+, A, B) based on their fraction of complex formation.[168, 183, 185, 186] Of course, the most accurate way for affinity ranking would be to determine the dissociation constant K_d. Since this would imply a titration experiment for each fragment, this approach is incompatible for a primary screen and might rather be used for further characterization of a limited number of primary hits. Determination of K_d values by single-point estimation could in principle be an alternative. But as it was exemplarily shown in the titration with MTH1 (Figure 4.2), this approach is error-prone in case that potential hits cover a wide range of affinities, which is usually the case for a large fragment library. This issue is also supported in a previous study by Crosby and coworkers, in which simulated binding curves according to different affinities are compared.[168] Thus, the fraction of the complex formation (% PL) that was determined for each fragment, allows to qualitatively rank fragment hits within a defined experimental setup.

Within this study, the results of MS were compared with data from TSA. Here, the question arises how comparable these methods are in terms of detecting binders. MS is conducted in the gas phase and measures direct ligand binding, while TSA is based on a solution environment and measures protein stabilization. In theory, hydrophobic interactions are weakened in the gas phase, while electrostatic interactions are strengthened.[108, 204] However, it was shown that hydrophobically driven complexes can survive in the gas phase.[109, 205] The proteins MTH1 and KDM5B both show a good hit overlap of the two methods. 16 mutual hits are identified for MTH1, while eight common hits are obtained for KDM5B (Table 4.4). This quite good agreement is not observed for BRPF1 and UHRF1. Although the TSA primary screen resulted in several hits for both proteins, binding in the gas phase is rather low or not observed, even for six fragment hits that were confirmed by X-ray crystallography (for BRPF1).

Table 4.4: Overall comparison of MS and TSA hits.

Target protein	No. of ESI-MS hits[c]	No. of TSA hits[d]	No. of mutual hits	Total no. of fragments
MTH1	16	24	16	33
KDM5B	10	9	8	16
BRPF1[a,b]	0/2	11	0/2	21
UHRF1	1	10	0	21

[a] 100 μM fragment concentration. [b] 500 μM fragment concentration. [c] Hits in MS are fragments that show more than 20 % binding. [d] Hits in the TSA increase the protein melting temperature T_m by more than six times the standard deviation (stabilizers).

One possible explanation for missing these confirmed binders might be a low binding affinity or a binding mode predominantly based on entropically-driven hydrophobic interactions with the target protein. However, a clear correlation was not found between physicochemical properties of the fragments (Appendix A, Table A.1) and their binding behavior in native MS. This finding is in agreement with a study by Maple *et al.*, in which a larger library of 157 compounds was screened by native MS and no observable correlation between K_d (determined by native MS) and $\log_{10}(D)$ of the compounds was observed.[168]

Furthermore, the size of the protein sets the application of native MS in FBS a natural limitation, which can only partially be overcome by raising the resolution power of the mass analyzer. Adduct formation with water molecules and salt ions increases with higher molecular weight resulting in broader peak shapes (not considering natural heterogeneities induced by post-translational modifications or cofactor binding), which hampers the detection of fragment binding. Typically, the cone voltage would be adjusted in such situations in order to remove loosely bound water and salt adducts and to improve peak shape. However, fragments binding with low affinity might dissociate at increased cone voltage resulting in false negatives. This apparent problem is certainly a burden working with larger proteins with no immediate technical solution since this is an inherent issue of electrospray ionization.

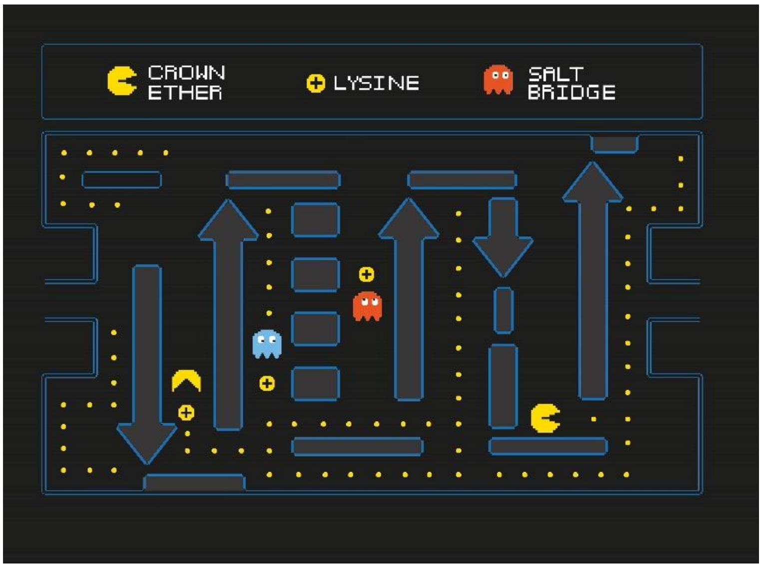
4.5 Conclusions

In this chapter four different target proteins were investigated regarding their suitability for fragment-based screening (FBS) with native mass spectrometry. The results were used to critically evaluate the method for FBS. For each protein a set of fragments was compiled that exhibited stabilization, destabilization, and neutral effects in a thermal shift assay primary screen. The fragments were ranked by their fraction of formed complex (% PL) and classified as hits using an arbitrary set threshold of 20 % PL. Generally, native MS yields a good overlap with data from TSA above a certain amount of complex formation, as it was shown for two of the investigated proteins. Stabilizers tend to be good binders in MS as well, while a clear binding type for destabilizers was not observed. Critical issues such as the optimal fragment concentration and a convenient way of ranking hits were discussed. Multiple binding and charge-state shifts were observed with all proteins and greatly hampered automated data analysis using commercially available software packages. It can be concluded that native MS may not be the method of choice as a primary screening technique. However, native MS can be considered as a highly valuable orthogonal screening method for hit validation of a restricted set of fragments.

The great advantage of the method in this regard is that it is a truly label-free method, exhibits a direct and simple readout (*i.e.* mass shift induced by ligand binding) and can also provide additional information apart from detection of pure binding. This for example includes the analysis of several ligands in a single experiment (multiplexing), which allows elucidating competitive *versus* simultaneous binding modes. In case that different fragments show simultaneous binding in different regions of a binding pocket this could serve as a starting point for fragment linking, while competing fragments can easily be ranked by their relative binding affinities in such an experiment.

Taken together, native MS is convenient for fragment-based screening, provided that the protein itself is suitable and allows automated data acquisition. This, however, bears the consequence that the potential of the method is not fully exploited, as the complete information content offered by a mass spectrum is often only obtained by manual inspection.

5 Protein Gas-Phase Microsolvation*



5.1 Introduction

Electrospray ionization-mass spectrometry (ESI-MS) has developed into a powerful and indispensable analytical tool for the structural investigation of proteins. It is still debated to what extent protein structure is retained after ESI and transfer into the gas phase and there is a growing number of publications addressing this question.[44, 206, 207] In this context it is generally accepted that the charge state presents a major determinant for the three-dimensional organization of a protein.[53] Species with a low amount of charge adopt compact, presumably native-like structures, whereas species in higher charge states adopt more extended structures.[208-210] Intermediate charge states can show many different coexisting conformations, ranging from compact to more extended structures. Furthermore, it was proposed that after transfer into the gas phase, intramolecular charge self-solvation, *i.e.* the collapse of the charged side chains onto the backbone can disrupt the intramolecular hydrogen network.[44]

---

* This chapter is based on the work „*Gas-phase microsolvation of ubiquitin: investigation of crown-ether complexation sites using ion mobility-mass spectrometry*", published in M. Göth, *et al.*, *Analyst* **2016**, *141*, 5502-5510, http://dx.doi.org/10.1039/c6an01377e. Figure and content adapted with permission. Copyright 2016 Royal Society of Chemistry.

A powerful tool to obtain information about molecular structure in the gas phase is ion mobility-mass spectrometry (IM-MS),[66, 76, 211] which was introduced in Chapter 2.3. It allows the separation of investigated molecules according to their mass, charge, size, and shape. In addition, a rotationally averaged collision cross section (CCS) [79] can be determined that yields direct information on molecular structure and can also be compared to theoretical values. The potential of IM-MS in the analysis of proteins and their assemblies has been shown for several systems.[64, 86, 212] In this regard, one extensively studied molecule is ubiquitin.[88, 207, 209, 213-219] It is a small protein (76 residues) that plays an important role in signal transduction and post-translational modification of other proteins.[220]

In this chapter, it shall be investigated how side-chain-to-backbone interactions influence the overall gas-phase structure of ubiquitin. Therefore, crown-ether molecules (CE, 18-crown 6, 264 Da; inset in Figure 5.1A) are attached non-covalently to the protein. Generally, 18-crown-6 is known to bind strongly to protonated lysines in peptides and proteins *via* three hydrogen bonds.[221-223]

A representative mass spectrum of ubiquitin and its crown-ether complexes in charge state 5+ is shown in Figure 5.1A. Up to six 18C6 adducts are observed under mild ESI conditions, each with a mass shift of *m/z* 52.8 (264/5). The corresponding arrival time distributions (ATDs) are shown in Figure 5.1B with the drift times converted to collision cross sections (CCSs). The general effect of the CE-attachment on the structure of ubiquitin can be described as follows: For ions with a low number of charges ($z = 4$) and ions in high charge states ($z \geq 8$) each additional CE leads to an increase in size and mass (Figure B.1), which in turn is reflected in an increase of the corresponding drift time and CCSs. Intermediate charge states (5+ to 7+ for ubiquitin), however, yield a rather counterintuitive result: By the addition of CEs, the conformational heterogeneity is reduced and the gas-phase structure becomes more compact. In charge states 6+ and 7+ this is reflected with CCSs being significantly smaller than those of the bare ions. In general, three major conformations seem to exist across the charge states shown in Figure B.1, with compact (~1000 Å^2), intermediate (~1200 Å^2), and extended (~1500 Å^2) structures.

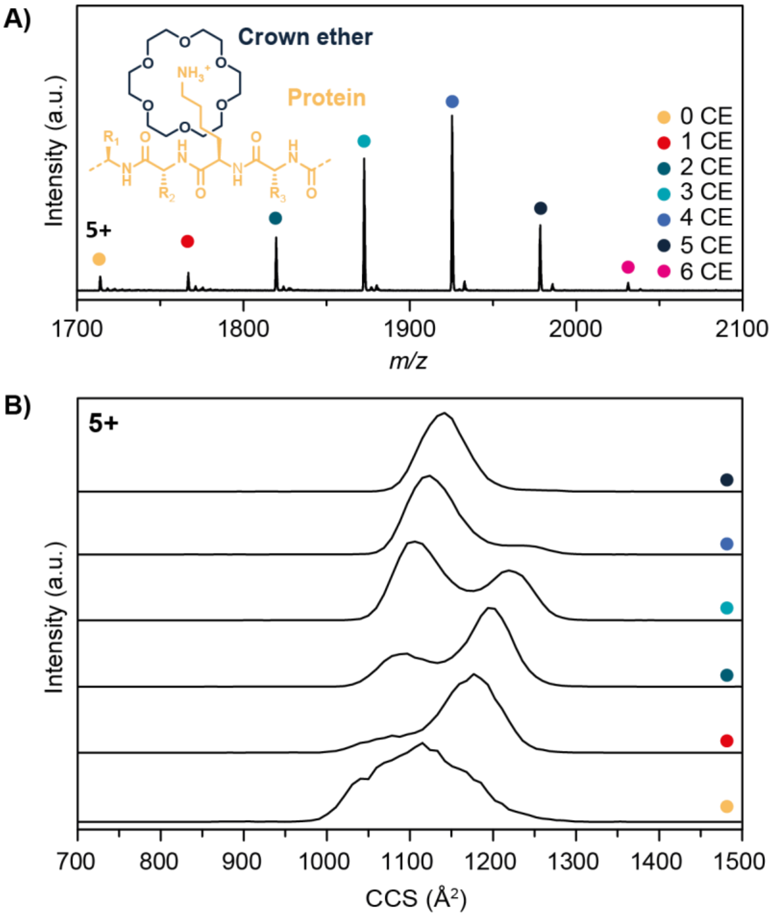


Figure 5.1: A) nESI mass spectrum of wild type (wt) ubiquitin in charge state 5+ with different numbers of 18-crown-6 (264 Da) attached. The inset structure shows in a simplified form how the crown ether (CE) binds non-covalently to protonated side chains. **B)** Arrival time distributions (ATDs) corresponding to A) with same color code for the CE complexes. With increasing number of CE attached, the overall structure becomes more compact, which is reflected by a decrease in CCS.

The population of bare 5+ ion is rather heterogeneous, showing a mix of the compact (*i.e.* presumably native-like) and intermediate states, and attachment of multiple CEs leads to the disappearance of the intermediate species. As such, the compact form becomes dominant when three or more CEs are coordinated (Figure 5.1B). The bare 6+ ion (Figure B.1) occurs predominantly in the extended state and adopts the intermediate conformation after binding a single crown ether. Binding of five crown ethers eventually results in the appearance of the compact state, but it never becomes dominant. As such, and because the 5+ ion is the lowest and most native-like charge state for which the compaction phenomenon is observed *via* IM-MS, for the rest of the analysis it will be focused on this ion. In a previous study on cytochrome c-CE complexes it was proposed that the crown ether molecules take over the role of the solvent and microsolvate the protonated lysine residues, such that the overall structure remains compact.[127] Within this chapter, this effect is investigated further by analyzing the CE-binding of ubiquitin in comparison to ubiquitin mutants using CID-MS, ETD-MS and IM-MS. In particular, the following questions are addressed: (1) Does microsolvation have the same effect when certain lysine residues are replaced by arginine and, in this context, do all lysines play an equal role for this effect? (2) Can we identify the binding site of the crown ether? (3) Does binding to the charged N-terminus lead to a similar microsolvation effect?

5.2 Experimental Details

Sample Preparation

Wild-type ubiquitin (human) was purchased from Boston Biochem (Cambridge, USA) as a colorless powder and dissolved in water to a concentration of 100 μM. The mutants were obtained from Sigma Aldrich (Taufkirchen, Germany) as phosphate-buffered saline (PBS) solutions. Prior to use, all samples were buffer exchanged in 10 mM aqueous triethylammonium acetate solution (pH = 7, Fluka Analytics) twice for 2 hours and overnight using slide-a-lyzer dialysis tubes (MWCO = 2000 Da, Thermo Scientific). After dialysis the samples were diluted to 10 μM with a 1:1 (v:v) solution of water: methanol. 18-crown-6 was purchased from Sigma Aldrich (Taufkirchen, Germany) as a colorless powder and dissolved in water to a 10 mM stock solution. For the preparation of the protein-crown ether complexes 25 to 50 equivalents of 18C6 were added to a 10 μM protein solution.

Mass Spectrometry

CID and IM-MS experiments. Measurements were performed in positive ion mode on a Synapt G2-S quadrupole-ion mobility time of flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK), equipped with a Z-spray nanoflow ESI (nESI) source. nESI tips were produced in-house from borosilicate capillaries using a needle puller (Flaming/Brown Micropipette P-1000, Sutter Instrument Company, Novato, USA) followed by Pt/Pd (80/20) coating (Sputter Coater HR 208, Cressington, Dortmund, Germany). Typical instrument parameters were as follows: 0.8-1.0 kV capillary voltage, 40 V sample cone voltage, 180 mL/min He cell gas flow (1.42e-3 mbar, He), 90 mL/min IMS gas flow (3.45e0 mbar, N_2), 2 mL/min trap gas flow (2.29e-2 mbar, Ar), 2 V trap collision voltage, 38 V trap DC bias, 40 V IMS wave height, 500-900 m/s IMS wave velocity. CCSs were estimated using an established calibration procedure[79] and absolute CCS values measured with helium as drift gas on an in-house constructed drift tube IM-MS instrument, reported elsewhere.[224] To minimize errors in the calibration, drift times were recorded for several wave velocities and CCS values were averaged. Figure 5.6 and Figure B.5 show arrival time distributions measured with a wave velocity of 700 m/s. For the CID experiments the trap collision voltage was increased stepwise from 2 to 40 V. To calculate the depletion of the precursor complex ion, the intensity of the desired precursor was divided by the sum of intensities of all CE-bound states (including the bare ion). MassLynx (v 4.1) was used to record and analyze the data. Further data analysis was performed using Origin 8.6. External *m/z* calibration was performed using CsI solution.

ETD-MS experiments.* ETD-MS experiments were performed in positive ion mode on a Synapt G2 quadrupole-ion mobility time of flight (Q-IMS-TOF) mass spectrometer (Waters, Manchester, UK), similar to the G2-S instrument described above. nESI tips were produced in house as explained above. Typical instrument parameters were as follows: 0.8-1.0 kV capillary voltage, 40 V sample cone voltage, He and IMS gas flow switched off, 20 mL/min trap gas flow (6.2e-2 mbar, He), 1 mL/min transfer gas flow (5.7e-3 mbar, Ar), 4 V trap collision voltage, 0 V transfer collision voltage, 2 V trap DC bias, 50 mL/min make up gas flow, 15 µA discharge current, 0.3 V trap wave height, 300 m/s trap wave velocity. MassLynx (v 4.1) was used to record and analyze the data. External *m/z* calibration was performed using CsI solution.

5.3 Results and Discussion

5.3.1 Crown-Ether Microsolvation of Ubiquitin and Ubiquitin Mutants

To study the impact of microsolvation of individual lysine (K) residues, wild-type ubiquitin (wt) and ubiquitin lysine-to-arginine mutants are investigated. As arginine is positively charged at physiological conditions and therefore comparable to lysine, it is assumed that the mutation does not have major effects on electrostatic interactions and protein structure. Although arginine also binds 18-crown-6 molecules, it has a considerably lower affinity (133 kJ/mol) than lysine (150 kJ/mol) and previous studies further suggested that binding to lysine is favored over arginine when both residues are present.[223]

---

* ETD-MS experiments were performed by Dr. Frederik Lermyte at the University of Antwerp, Belgium.

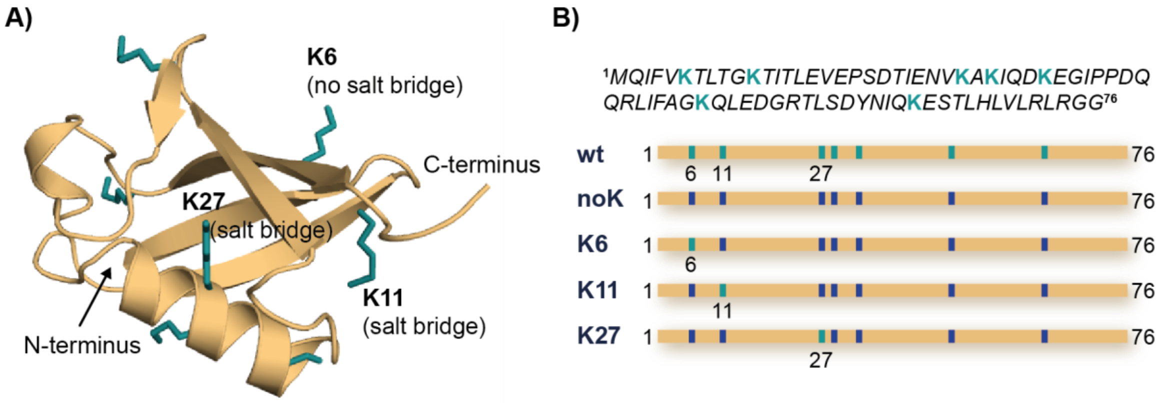


Figure 5.2: A) Crystal structure (PDB entry: 1UBQ) of wt ubiquitin with its seven lysine residues highlighted in green. While K11 and K27 are involved in salt bridges, K6 has no interaction partner in close proximity. **B)** Amino acid sequence of wt ubiquitin and schematic sequence comparing wt to the mutants noK, K6, K11, and K27. Either all lysines (green) or all but one specific lysine are replaced by arginine (blue).

In Figure 5.2A the crystal structure of wt ubiquitin is shown with the seven lysine residues highlighted in green. Figure 5.2B shows the amino acid sequence as well as a schematic overview of the mutant sequences in comparison to wt ubiquitin with arginines colored in dark blue and the retained lysines in green. In the mutants investigated here, either all lysine residues (noK-mutant) or all except one lysine at a specific position are replaced by arginine (K6-, K11-, K27-mutant). The focus specifically lies on mutants with lysines retained close to the N-terminus to be able to perform ETD experiments on ubiquitin-CE complexes in which lysine-containing N-terminal fragments (possibly with a bound CE) are produced at high abundance. Inspection of the individual residues in the crystal structure (PDB entry: 1UBQ) for a native-like structure shows that K11 and K27 are both involved in salt bridges, whereas K6 is not in close proximity to an acidic side chain.[225]

Under soft ESI conditions up to six CE adducts are observed for wt ubiquitin and up to five for the mutants, with all spectra showing a similar distribution of the complexes (Figure 5.3A). However, under slightly harsher conditions (increase of trap voltage from 2 V to 9 V), the distribution of the CE complexes changes significantly for the mutants, whereas the change is less obvious for the wt protein with a maximum intensity at the 2CE complex (Figure 5.3B). The most intense shift between soft and harsh conditions is observed for the noK-mutant, where the bare protein ion and the 1CE complex show the highest intensity with an almost 1:1 ratio. The distributions of K6-, K11- and K27-mutant complexes resemble each other with the 1CE complex as the most abundant species (Figure 5.3B).

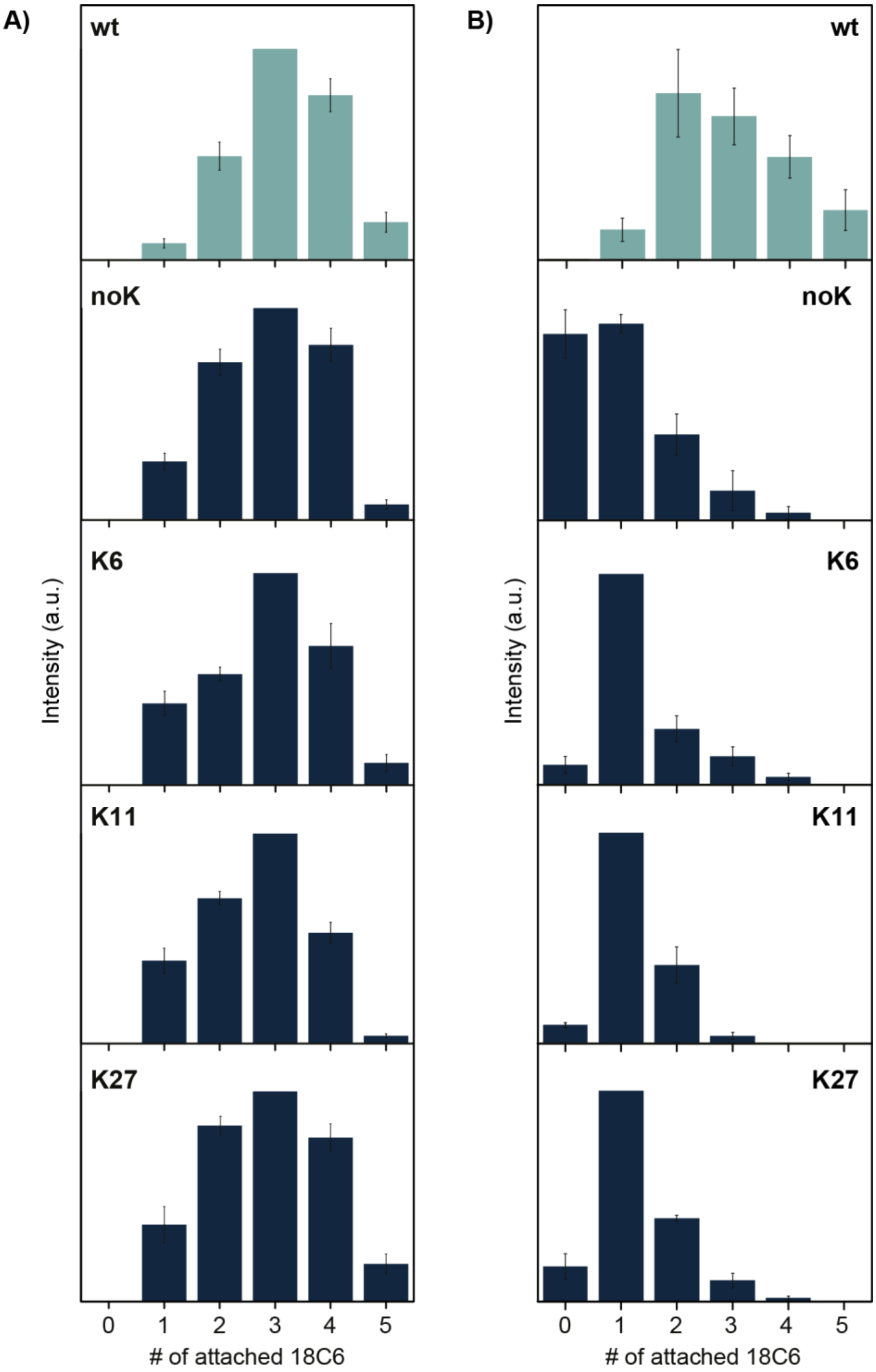


Figure 5.3: Comparison of the relative peak intensities of the protein crown-ether complexes **A)** under soft source conditions and **B)** under harsh conditions (increase of trap voltage), both at charge state 5+.

5.3.2 Energy-Resolved Collision-Induced Dissociation Experiments

In order to better understand the CE-distributions represented in Figure 5.3 and to analyze differences in CE-binding energies between the mutants, energy-resolved CID experiments were performed. To do so, either the 1CE or the 3CE complex was isolated as precursor and subsequently fragmented *via* CID. The selection of the 3CE complex is based on the assumption that the third CE is neither attached to a lysine nor to the N-terminus in case of the mutants. Representative MS/MS spectra, with the respective precursor ions shaded in grey, are shown in Figure 5.4A and Figure 5.4B for the noK-mutant at a collision voltage of 9 V. Spectra for the other investigated mutants are shown in Figure B.2. Isolation and subsequent dissociation of the 1CE complex results in the formation of the bare protein ion; dissociation of the 3CE complex, on the other hand, leads to the formation of three different product ions at increased collision voltage. To obtain a more quantitative picture for the different mutants, the relative intensity of the corresponding precursor ions was plotted *versus* the collision voltages for all samples. The resulting CID depletion curves are of typical sigmoidal shape (Figure 5.4C and Figure 5.4D). In addition, the voltage that is necessary for a 50 % depletion of the corresponding precursor ion is marked by a dashed line in both graphs.

Comparing the graphs of the 1CE depletion amongst each other, three distinct trends become apparent: (1) The noK-species needs the lowest energy to dissociate. This was expected, as there are no lysine residues for the CE to coordinate to. (2) Slightly more energy is required for a 50 % fragmentation of the K6-, K11- and K27–mutant complexes. (3) The highest energy has to be applied to dissociate the wt-1CE complex. These results show that a distinction between wt and the mutants and even between the two types of mutants is generally possible based on the relative dissociation energy. The resulting gap in effective binding energy between wt and the other three mutants is somewhat unexpected, as the wt as well as the mutants offer similar binding sites to the CE. In case of the wt the CE coordinates either to the N-terminus or to one of the seven lysine residues, which are believed to be in close spatial proximity in charge state 5+. Three considerations could explain the observed stability of the wt-1CE complex: (1) The CE-binding sites are statistically distributed. Thus, one particular site might be favored, but the CE can in principle be attached to other sites as well. These other sites would still be lysines in case of wt and likely arginines in the mutants. (2) The first CE is neither attached to K6, K11 or K27 nor to the N-terminus, but to one of the four remaining lysines that are not investigated here, and which somehow has a much greater affinity than the three analyzed mutants.

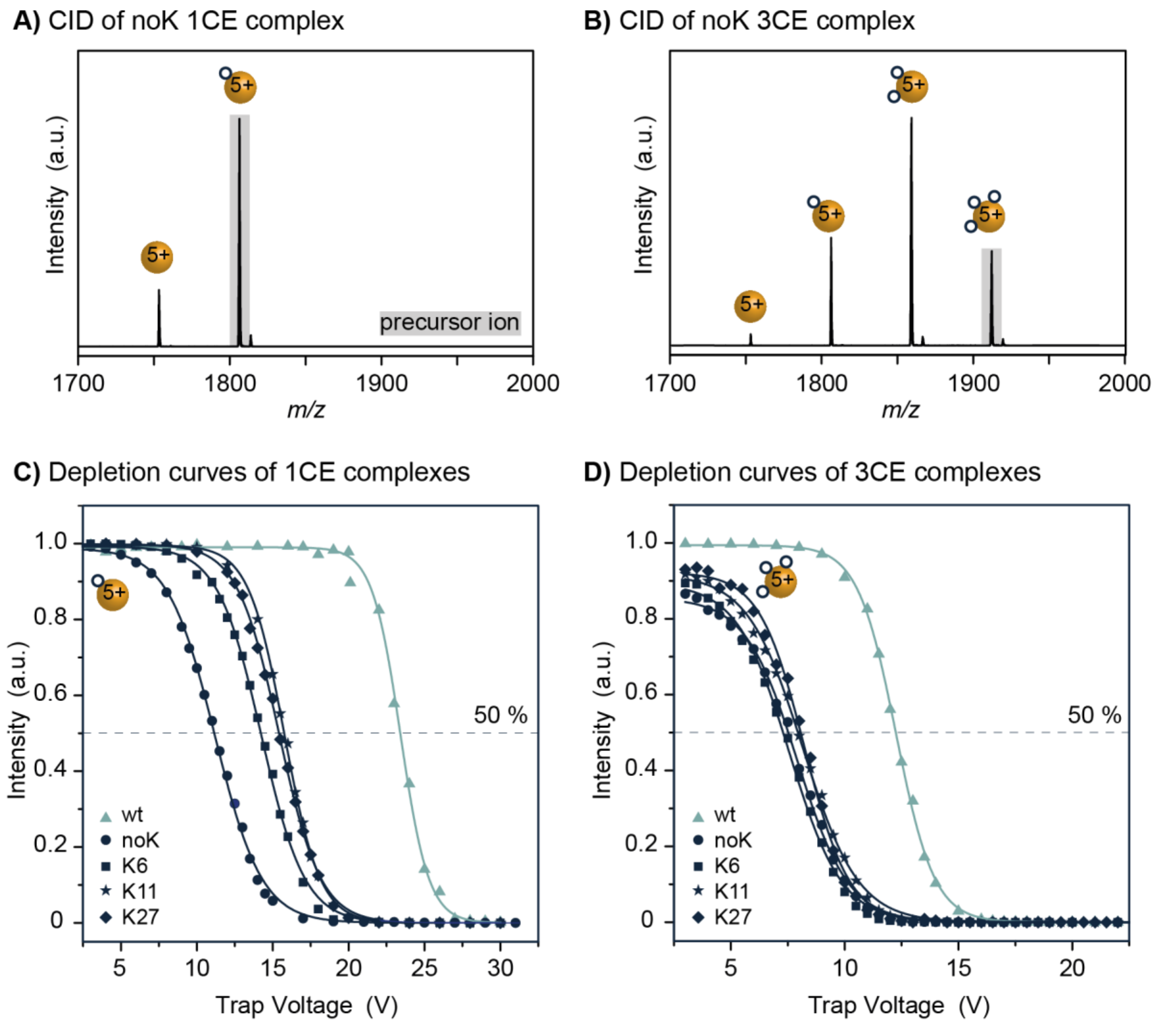


Figure 5.4: CID spectra of noK CE complexes with **A)** 1CE and **B)** 3CE as precursor ions (shaded in grey) at a trap collision voltage of 9 V. For noK the bare protein ion is already observed at low collision voltages. Depletion curves of **C)** 1CE complexes and **D)** 3CE complexes of all proteins. Collision voltages are plotted *versus* the relative abundance for the decrease of the corresponding CE complexes.

(3) The CE is not permanently bound to one specific site, but can "walk" between them when they are close enough to each other. This phenomenon was reported earlier[226, 227] and might offer an explanation for the increase in the effective binding strength of wt.

Next, the 3CE complex of each variant was isolated, fragmented and its depletion plotted as described above (Figure 5.4D). The collision voltage for a 50 % fragmentation is roughly similar for all mutant species and in addition significantly lower than the energy necessary to dissociate the wt complex. In case of the wt, the third CE is likely coordinated to a lysine, whereas the mutants presumably offer arginine as a binding site. Furthermore, a comparison of the wt depletion in Figure 5.4C and 4D shows that the third CE is bound significantly weaker than the first CE. This might be attributed to the fact that the 1CE and 3CE complexes of the wt exhibit different structures as shown in Figure 5.1. Alternatively, due to the two already occupied sites the third CE may also not be able to "walk" between different residues as freely as the first CE and is therefore not bound similarly strong. Nevertheless, the CID experiments and the resulting depletion graphs show that it is possible to differentiate between a coordination of the CE to arginine and lysine on basis of the relative dissociation energy.

5.3.3 Electron Transfer Dissociation Experiments

MS and CID experiments provide insights into the relative binding energies of the investigated protein-CE complexes and thus allow conclusions on the CE-binding sites. However, these experiments do not reveal the role of the N-terminus and to what extent it is involved in the CE-microsolvation.

To answer this question, electron transfer dissociation (ETD) experiments were performed. This method provides the advantage of preserving non-covalent interactions while cleaving covalent bonds.[123, 228, 229] Thus, fragments of different size can be formed with a CE still being attached to the corresponding lysine. It should therefore be possible to identify the binding sites from the ETD fragment mass spectra. To minimize loss and/or migration of the CE ligand in the gas phase, acceleration voltages were kept minimal in ETD experiments. While this limited the fragmentation efficiency and sequence coverage, fragmentation in the first 5 to 15 residues (depending on precursor charge state) yields insight into CE binding and compaction in wt ubiquitin, as well as the K6, K11 and noK-mutants.

Protein-CE complexes with different numbers of CEs were selected as precursor ions in charge state 5+. Generally, high sequence coverage and signal-to-noise ratio can be obtained in ETD for precursor ions with a high number of charges.[230] The reason for this is twofold: First, the increased Coulomb attraction between ETD reagent and precursor will result in an increased ion/ion reaction rate.[231] Second, increased intramolecular electrostatic repulsion will reduce the degree to which the protein structure is stabilized by non-covalent interactions, facilitating fragment release after cleavage of the backbone N-Cα bond.[232]

In this case, sufficiently resolved fragment spectra were also obtained for the lowly charged 5+ ions investigated in this study and the resulting mass spectra of all complexes show a vast variety of fragments (Figure B.3). In addition to shorter N-terminal fragments, also charge-reduced species are observed.

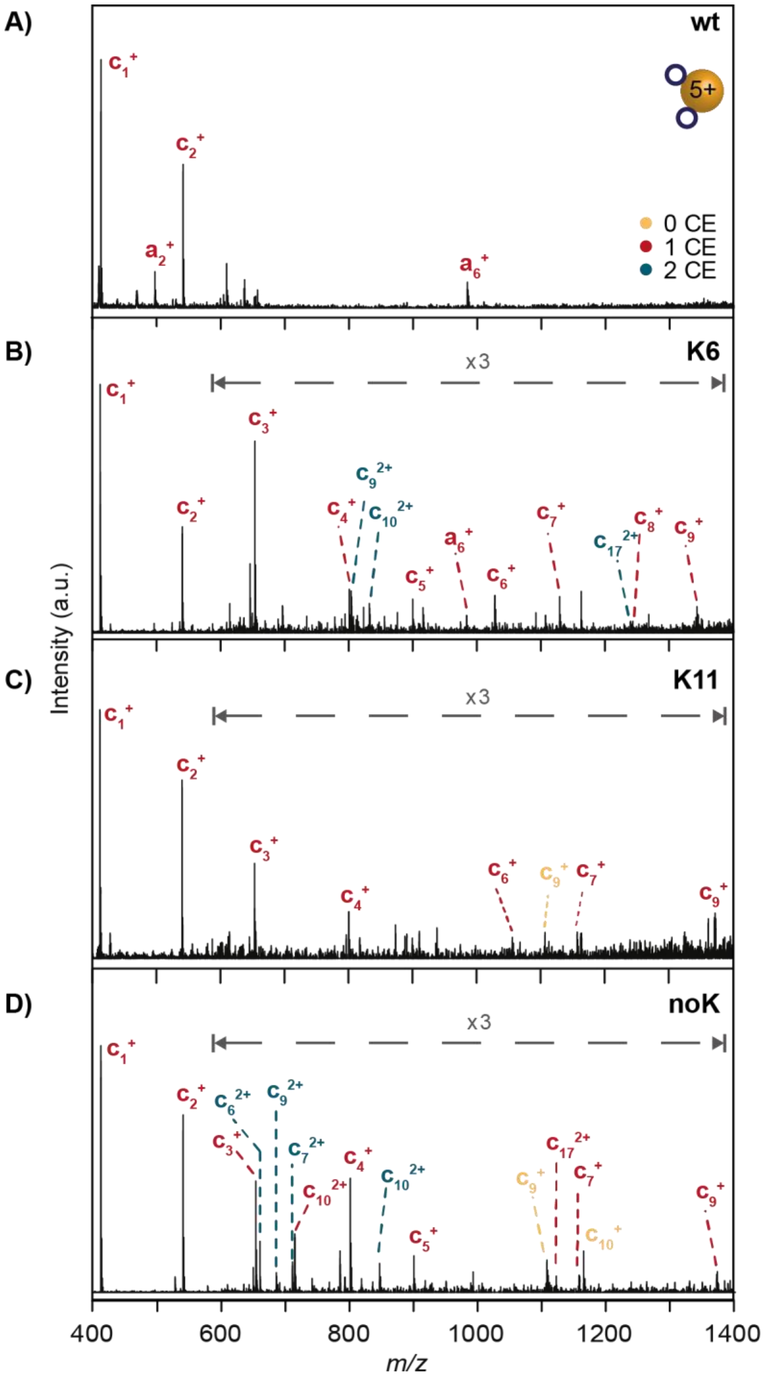


Figure 5.5: ETD fragmentation spectra of 5+ complexes with two crown ethers attached. **A)** Wild-type ubiquitin, **B)** K6-mutant, **C)** K11-mutant, and **D)** nok-mutant.

Those species are formed either by proton or electron transfer, or gas-phase adduct formation with the ETD reagent as reported earlier.[233, 234] Figure 5.5 shows the lower *m/z* region of the ETD fragment spectra of the wt, K6, K11 and noK complexes for a 5+ precursor ion with two CEs. The color code for the different amount of CEs attached is the same as in Figure 5.1.

In general, a- and c-type fragments with no CE (yellow), one CE (red) and two CEs (blue) are observed. Despite the rather low fragmentation yield of the lowly charged precursor ions, a few intense N-terminal fragments (c_1^+, c_2^+ and c_3^+) carrying one CE are observed for all four proteins. As these fragments contain neither a lysine nor an arginine residue, they indicate binding of the CE to the protonated N-terminus. In addition, for the wt complex an a_6^+ fragment was identified with a single CE bound (*i.e.* either attached to the N-terminus or to K6). However, no further fragments with more than one CE were formed for wt 5+ ions. In charge state 6+, a higher fragmentation yield was achieved, where nearly all of the fragments carried one or two CEs (Figure B.4). These spectra show sequences with up to 17 residues and indicate that the second CE binds preferentially to K11, but can also bind to K6. In turn, fragment spectra of the K6-mutant complex (Figure 5.5B) show a slightly better sequence coverage for charge state 5+. In addition to clear N-terminal binding (c_1^+, c_2^+, c_3^+, c_4^+, and c_5^+), larger fragments (a_6^+, c_6^+, c_7^+, c_8^+, and c_9^+) are formed with one CE attached (*i.e.* either to the N-terminus or to K6). Furthermore, the occurrence of the fragments c_9^{2+}, c_{10}^{2+} and c_{17}^{2+} with two CEs suggests binding of the second CE to lysine at position six.

For the K11-mutant, ETD experiments were also performed under similar conditions (Figure 5.5C). However, at charge state 5+ only fragments up to a length of nine amino acids with one or no CE attached are formed (Figure 5.5C). These fragments suggest that the first CE is likely to bind to the N-terminus. Since fragments were observed only up to c_9, it is not possible to reveal further binding sites. However, based on known affinities and ETD of the K6-mutant, the second CE likely coordinates to K11. Given the limited sequence coverage observed under these soft conditions, ETD experiments were not performed with the K27-mutant.

ETD spectra of the noK-mutant complexes (Figure 5.5D) show in addition to the intense N-terminal fragments with one CE also fragments with two CEs attached (c_6^{2+}, c_7^{2+}, c_9^{2+} and c_{10}^{2+}). These short sequences contain the N-terminus and one arginine residue at position 6 as possible coordination sites for the two CEs. Furthermore, a larger fragment is observed with c_{17}^{2+}. In this case the single attached CE is able to bind either to the N-terminus, R6 or R11.

From the ETD experiments, it can be concluded that N-terminal CE-binding likely occurs for all investigated species, as the corresponding fragments were among the most intense signals in the fragment spectra. These fragments indicate that it is the first CE that binds to the N-terminus. However, there are also fragments observed where it is not explicitly clear whether the first CE is coordinated to the N-terminus or to K6. In general, the low sequence coverage in charge state 5+ hinders a clear identification of every CE-binding site.

5.3.4 Ion Mobility-Mass Spectrometry Experiments

The major focus of the study was to test whether the effect of microsolvation can still be observed with one or several lysine residues being replaced by arginine and, thus, to reveal if all lysines contribute equally to this effect or if specific residues play a major role. To clarify this question, the protein-CE complexes were investigated further using ion mobility-mass spectrometry.

The measured arrival time distributions (ATDs) for the proteins and their CE complexes were converted to contour plots, in which the colour intensity reflects the peak intensity of the raw data. The corresponding contour plot for wt ubiquitin is shown in Figure 5.6A and reveals a similar trend for a microsolvation-induced compaction as discussed above. Experiments under the same conditions were performed with the mutant-CE complexes and the resulting ATDs are presented as contour plots in Figure 5.6B. Raw IM-MS data of all possible lysine-to-arginine mutants (charge state 5+) are shown in Figure B.5. Generally, a different behavior is observed for all mutants compared to the wt. The ATD of the bare noK-mutant appears much narrower than that of the wt and reveals a compact conformation. Further compaction does not occur upon the addition of crown ethers. Similar behavior is observed for the K11- and K27-CE complexes (and K29-, K33, K48-, and K63-CE complexes, Figure B.5). In all these cases, the drift time increases steadily with growing mass of the complex, and shows a rather narrow and homogenous ATD, corresponding to a compact conformation. The K6-mutant, on the other hand, reveals the same conformational heterogeneity as the wt in the absence of CE. More importantly, however, K6 shows a similar effect as the wt sequence when CE molecules are bound, and requires fewer ligands to undergo this effect. The second, more compact conformational family evolves with the first attached CE and dominates after addition of the second CE molecule. As a result, K6 requires at least two (compared to three for the wt) coordinated CEs to undergo compaction. Therefore it is likely that both the N-terminus and K6 are involved in binding and that K6 plays a key role for the extending of the structure in ubiquitin. This assumption is supported by the ETD experiments, as both N-terminal and K6 CE-binding are observed with the first two CEs.

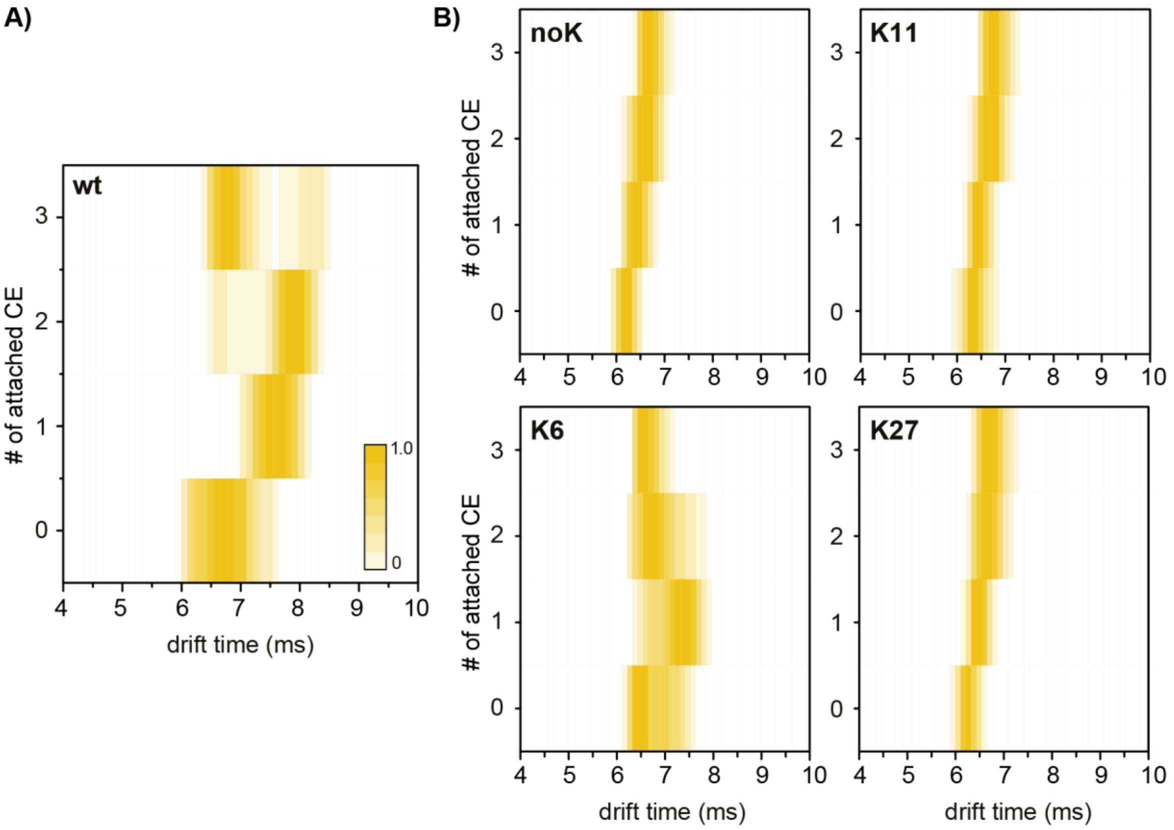


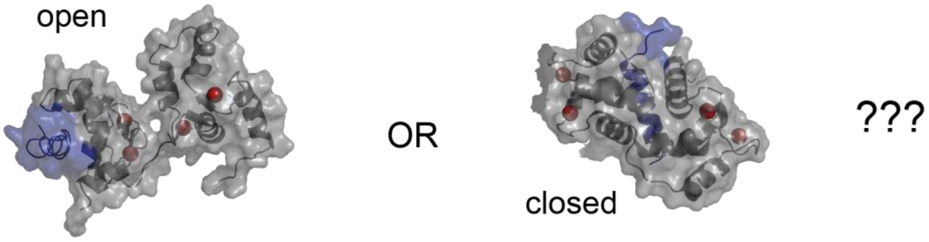
Figure 5.6 A) Arrival time distributions (ATDs) for wt ubiquitin with 0 to 3 CEs (5+) presented as contour plots. Upon attachment of CE molecules, the gas-phase structure increases at first, but a second more compact species evolves as well. This conformation dominates at higher number of crown ethers. **B)** Contour plots for the ATDs of the mutant complexes. Only the K6-mutant shows the same conformational heterogeneity and the compaction upon CE addition as the wt. The other mutant variants exhibit homogeneous and compact structures already without CE attachment.

Comparison of the microenvironment of the particular lysine residues in the crystal structure of the protein (Figure 5.2) reveals that K6 is the only one of the investigated residues that is not involved in an intramolecular salt bridge. As the next potential binding partner for K6 is approximately 5.4 Å away,[225] the charged residue is more likely to self-solvate onto the protein backbone after transfer into the gas phase. This, in turn, leads to a disruption of structure-stabilizing intramolecular hydrogen bonds and, eventually, to the unfolding of the protein's compact conformation. Salt bridges, on the other hand, are likely to support the compact and native-like structure of the 5+ ubiquitin ions. Previous density function theory calculations on the binding of 18-crown-6 to different possible conformations of protonated lysine revealed that in structures where lysine is involved in a salt bridge, the crown ether was likely to prefer binding to the corresponding N-terminus rather than side chain in order to retain the stabilization by the salt bridge.[223] Therefore, it is likely that lysine residues that take part in such structure-stabilizing interactions do not show a strong effect upon microsolvation with a CE. This agrees well with another study where CE complexes with ubiquitin lysine-to-asparagine mutants were analyzed *via* selective non-covalent adduct protein probing mass spectrometry (SNAPP-MS).[235-237] Here, the authors investigate the changing adduct distribution in mass spectra for mutants with only one lysine replaced in the sequence and reveal that intramolecular interactions, such as salt bridges and hydrogen bonds, can hamper the coordination of 18C6 to protonated lysine residues.[225].

5.4 Conclusions

In this chapter the influence of a protein's microenvironment on its gas-phase structure was investigated by non-covalent attachment of crown ether molecules to ubiquitin and ubiquitin lysine-to-arginine mutants. It was demonstrated that the structure of partially folded conformations of ubiquitin becomes more compact upon crown-ether attachment, which is in good agreement with a previous study on cytochrome c.[127] The results strongly indicate that crown ether molecules not only coordinate to protonated lysine and arginine residues, but also the N-terminus. Applying harsher conditions in the MS and CID experiments showed that the single crown-ether complexes differ in relative dissociation energies, which enables the differentiation between a coordination of the crown ether to lysine or arginine. Furthermore, differences in the dissociation energies of the respective complexes are not significant between mutants of the same type (K6, K11 and K27-mutant). ETD-MS experiments were performed to locate the crown-ether binding sites. Although a clear identification of all binding sites is not possible, the data reveal that the N-terminus is involved in the CE-microsolvation, and in some cases even by binding the first CE. When comparing the arrival time distributions of the different mutants, major differences in the overall gas-phase structure are observed upon attachment of CEs: Only with the K6-mutant a structural rearrangement similar to wild type ubiquitin is detected. All other mutants exhibit compact and homogeneous ATDs that increase in drift time without a noticeable structural change. In other words, the lysine residue on position 6 induces a structural extension and shows a compaction upon CE addition, whereas salt-bridged residues adopt a compact structure already without CEs attached. It is likely that the formation and preservation of salt bridges to spatially adjacent residues is more favorable than backbone solvation. Taken together, the results presented in this chapter strongly indicate that the effect of crown-ether microsolvation is dependent on the microenvironment of the specific residue.

6 Structural Investigation of CaM/Munc13-Peptide Complexes*



6.1 Introduction

Calcium ions (Ca^{2+}) are central cellular messengers in an extraordinarily high number of cellular processes, including gene transcription, cell proliferation, apoptosis, muscle contraction, and mitochondrial function. Multiple proteins can directly bind Ca^{2+} *via* Ca^{2+}-binding domains in their primary sequence (*e.g.* EF hands, C2 domains). Alternatively, an array of proteins have evolved that bind Ca^{2+} across a broad concentration range and additionally multiple target proteins, thereby regulating their activity in a Ca^{2+}-dependent manner. A large group of such proteins, typically grouped as the EF hand superfamily, is expressed in neurons.[238-240] The prototype of this group is calmodulin (CaM), a small acidic protein (148 amino acids, ~17 kDa) with exceptionally high sequence conservation among eukaryotes,[241] and sequence identity across vertebrates. CaM is expressed ubiquitously in all eukaryotic cell types. The high concentrations of CaM in the brain (up to 100 μM)[242] may reflect its important role for nervous system function. Among the myriad of cellular functions regulated by CaM,[243] the focus for this work lies on synaptic transmission.

CaM consists of a N- and a C-terminal globular domain, each containing two Ca^{2+}-binding helix-loop-helix motifs of the EF-hand type. These domains (here referred to as N- and C-lobes) have a considerable backbone flexibility that is key to their ability to bind a wide range of targets[244] and are connected by a flexible linker. Ca^{2+} binds to CaM in a cooperative manner ($K_d = 5{\cdot}10^{-7}$ to $5{\cdot}10^{-6}$ M), with the C-lobe EF hands having a three- to fivefold higher affinity for Ca^{2+} than the N-lobe EF hands.[245] Upon Ca^{2+}-binding CaM undergoes dramatic structural rearrangements, including conformational changes within the EF hands and the repositioning of the N- and C-lobes.[246] These conformational changes expose hydrophobic

---

* This chapter is in parts based on the work "*Presynaptic Calmodulin Targets: Lessons from Structural Proteomics*", published in N. Lipstein, M.Göth, C. Piotrowski, K. Pagel, A. Sinz, O. Jahn, *Expert Rev. Proteomics* **2017**, *14*, 223-242, http://dx.doi.org/10.1080/14789450.2017.1275966. Figures and content adapted with permission. Copyright 2017 Taylor & Francis.

surfaces where protein-protein interaction can occur. It is therefore postulated that the various Ca^{2+}-bound conformational states of CaM allow for unique interactions with target proteins.[246-248]

A recent analysis of the >80 unique CaM complex structures deposited in the Protein Data Bank (PDB) as of 2013[246] impressively highlights that a defined consensus sequence for CaM binding does not exist, and that target sequences often show very low homology. Nonetheless, CaM-binding sites do share several characteristic features, including a high helical propensity, a net positive charge, and the presence of hydrophobic anchor residues. The spacing of the hydrophobic anchor residues historically serves as criterion to the classification of CaM-binding motifs, such as in the prototypic 1-14 and 1-5-10 motifs of myosin light chain kinase (MLCK) and CaM-dependent protein kinase II (CaMKII), respectively. The recognition of such features has been implemented into bioinformatics tools for the prediction of CaM-binding sites,[249, 250] and novel concepts in computational biology (*e.g.* conversion of biological sequences into feature vectors to be combined with machine-learning algorithms[251, 252]) may lead to refined, ideally web server-based prediction tools.

CaM can bind its target proteins either in a Ca^{2+}-free form (apo-CaM) or in one of its Ca^{2+}-bound forms. The interactions are tightly dependent on the surrounding Ca^{2+}-concentration and can range from transient to irreversible. Thus, a variety of regulation modes by CaM have been described.[245] Further structural rearrangements in CaM occur upon binding to the target protein. For example, in the well-studied cases of CaM-dependent kinases like MLCK (1-14 motif) and CAMKII (1-5-10 motif), binding leads to a rather compact CaM structure in the complexes.[253-255] In contrast, when the flexible linker region between the lobes is fully extended, CaM can adopt an open conformation, and the anchoring residues in the target protein can theoretically be as far as 70 Å apart.[246]

An important neuronal CaM target is the Munc13 (mammalian homolog of C.elegans unc13 protein) protein family, which consists of four isoforms. These are Munc13-1, bMunc13-2, ubMunc13-2, and Munc13-3, described in more detail in section 6.3.3. Munc13s bind to CaM in a Ca^{2+}-dependent manner and form a sensor/effector complex that controls short term synaptic plasticity.[256] The presence of Munc13s is essential for the completion of the synaptic vesicle cycle: in synapses lacking Munc13s synaptic vesicles fail to attach to the plasma membrane, which results in a complete blocking of synaptic transmission and the immediate death of Munc13-1/2 double-knock-out mice after birth.[257] The central and essential role of Munc13 in synaptic transmission emerged recently also in form of its involvement in human neurological and neuropsychiatric disease. Previous studies on Munc13-derived peptides already gave valuable insights on the structure and function of these complexes.[258, 259] However, additional information is required also with respect to solve CaM-binding sites for all Munc13 variants and to better understand the function and role of Munc13-CaM complexes.

Ion mobility-mass spectrometry (IM-MS), which was introduced in Chapter 2.3, presents a promising method to get additional insights into the overall protein-peptide complex structures. In this study, IM-MS is applied as a tool for conformational screening for CaM/Munc13 peptide complexes to reveal if they adopt either compact (closed) or extended (open) gas-phase structures. In the first part of this chapter data on the conformational analysis on CaM and its Ca^{2+} complexes are briefly discussed. Then the screening approach is explained and test experiments with peptides derived from skeletal MLCK (skMLCK) are presented. Subsequently, a variety of different Munc13-derived peptides in complex with CaM are screened with IM-MS to analyze their gas-phase structure. In the last part of the chapter, the unfolding and dissociation of these CaM/Munc13 complexes are studied by collisional activation.

6.2 Experimental Details

Samples and Sample Preparation

Calmodulin (CaM) wild type recombinant (human) and all Munc13 peptides were provided by Dr. Olaf Jahn and coworkers (Max Planck Institute for Experimental Medicine, Göttingen, Proteomics Group). All samples were obtained as powder and dissolved to stock solutions as annotated. CaM was dissolved in 10 mM aqueous ammonium acetate solution (Sigma Aldrich, 7.5 M, pH ~ 7.5) and salt-exchanged using ultracentrifugation units (Amicon, Thermo Fisher, cut-off 3 kDa). Therefore, the sample was flushed 3x with a 2 mM ethylene diamine tetra acid (EDTA) solution and 1x with an ammonium acetate solution to wash out most of the chelator and therefore avoid masking the protein signals in the mass spectra. However, EDTA-CaM adducts are still observed in case that calcium was added. The protein concentration was determined *via* UV absorption at 280 nm using and extinction coefficient of 2980 $cm^{-1}M^{-1}$ (DeNovix DS-11, Biozym) and CaM was subsequently diluted to 10 μM in the final sample solution. Calcium acetate was purchased by Sigma Aldrich and dissolved in water (HPLC-grade, VWR) to a concentration of 1.5 mM. The peptides were dissolved in water to a stock solution of 1 mM, which was further diluted to prepare the final samples. The CaM/peptide/Ca^{2+} samples were prepared freshly from the stock solutions for every measurement, with a concentration of 10 µM CaM, 20 µM of peptide, and 150 µM of Ca^{2+} in the final sample. Lower calcium concentrations had also been tested, but the complex with four calcium ions attached was formed only at a concentration of > 100 µM as the main species in the mass spectrum.

Ion Mobility-Mass Spectrometry

Measurements were performed in negative ion mode on a Synapt G2-S quadrupole-ion mobility time of flight (Q-IMS-ToF) mass spectrometer (Waters), equipped with a Z-spray nano electrospray ionization (nESI) source. nESI tips were produced in-house from borosilicate capillaries using a needle puller (Flaming/Brown Micropipette P-1000, Sutter Instrument Company) followed by Pt/Pd (80/20) coating (Sputter Coater HR 208, Cressington). External *m/z* calibration was performed using a CsI (20 mg/mL) solution (1:1 isopropanol/H_2O). The original traveling wave ion mobility cell had been replaced by a linear RF confined drift cell as reported in detail elsewhere[260] and described in Chapter 2.5. Due to the modification, collision cross section (CCS) values can be determined directly without the use of an external calibrant by applying the Mason-Shamp equation (see Chapter 2.3 for details). All measurements were performed in helium as drift gas with a pressure of 2.2 Torr (~2.9 mbar), which was controlled by a multi gas controller (647C, MKS Instruments Deutschland GmbH). Wave velocity and wave height parameters were inactive and the parameters within the drift cell were altered by changing IMS bias or helium DC voltages. To minimize errors, drift times were recorded for eight different voltages and CCS values were averaged. The relative errors of determined CCSs were in most cases below 1 %, however, for all values only accepted if below 3 %. Typical other instrument parameters were as follows: *m/z* range 100-5000; capillary voltage 0.7-1.2 kV; sample cone voltage 30 V; source offset 30V; nanoflow gas pressure 0-0.3 bar; He cell gas flow switched off (He cell inactive), IMS gas flow (He) 20 mL min^{-1}; trap collision voltage 2 V; transfer collision energy 2 V; trap DC bias 10 V; IMS DC entrance -20 V; Helium Cell DC 50 V; IMS bias 60-95 V, transfer DC entrance 5 V; transfer DC exit 15 V.

Collision-Induced Unfolding and Collision-Induced Dissociation Experiments

Unfolding and dissociation of protein complexes were performed by tandem IM-MS. Collisional activation was carried out in the ion trap cell prior to the ion mobility drift cell.

Therefore ions of interest were *m/z*-selected in the quadrupole mass filter, accelerated in the trap region and subsequently separated in the ion mobility cell based on their conformation, followed by time-of-flight analysis. The trap collision voltage was gradually increased in 2-5 V steps starting at non-activating conditions (2 V) until the complex was completely dissociated (~ 80 to 100 V). IM data and mass spectra were recorded for each distinct trap voltage and all measurements were performed in triplicate.

Data Analysis

Mass spectra and arrival time distributions (ATDs) were processed using MassLynx 4.1 (Waters) and subsequent data analysis was performed with OriginLab 8.6 (OriginLab Corporation). To illustrate the unfolding of the complexes, the recorded trap voltages were plotted against the relative unfolding. This relative unfolding corresponds to the ratio of integrated initial complex conformer divided by the integral of the total complex ATD at one specific collision voltage. The 50 % unfolding voltage was subsequently determined by sigmoidal fit using a Boltzmann function. The value for a 50 % complex dissociation is determined in a similar manner. Therefore peak intensities were read-out for every tandem mass spectrum. The intensity of the parent ion (CaM/peptide/Ca_4^{2+}-complex) was divided by the intensity of all detected species (CaM/peptide/Ca_4^{2+}complex and dissociation products) in the corresponding spectrum, which results in the relative fraction of the complex. This fraction was plotted against the applied collision voltage to give a relative dissociation curve, which subsequently was fitted with a Boltzmann function.

6.3 Results and Discussion

6.3.1 Conformational Analysis of CaM and CaM/Ca^{2+} Complexes

The protein CaM has been widely studied and a large variety of methods have been applied for the experimental discovery of CaM binding partners. These involve for example probing of expression libraries,[261] mRNA display techniques,[262] affinity chromatography followed by mass spectrometric protein identification[263] and protein arrays.[264] In addition, native MS - alone and in combination with IMS - was applied to CaM and some of its non-covalent peptide complexes,[126, 236, 265-271] providing valuable information on the metal and target binding properties of CaM, as well as the stoichiometry and gas-phase structure of CaM/peptide complexes.

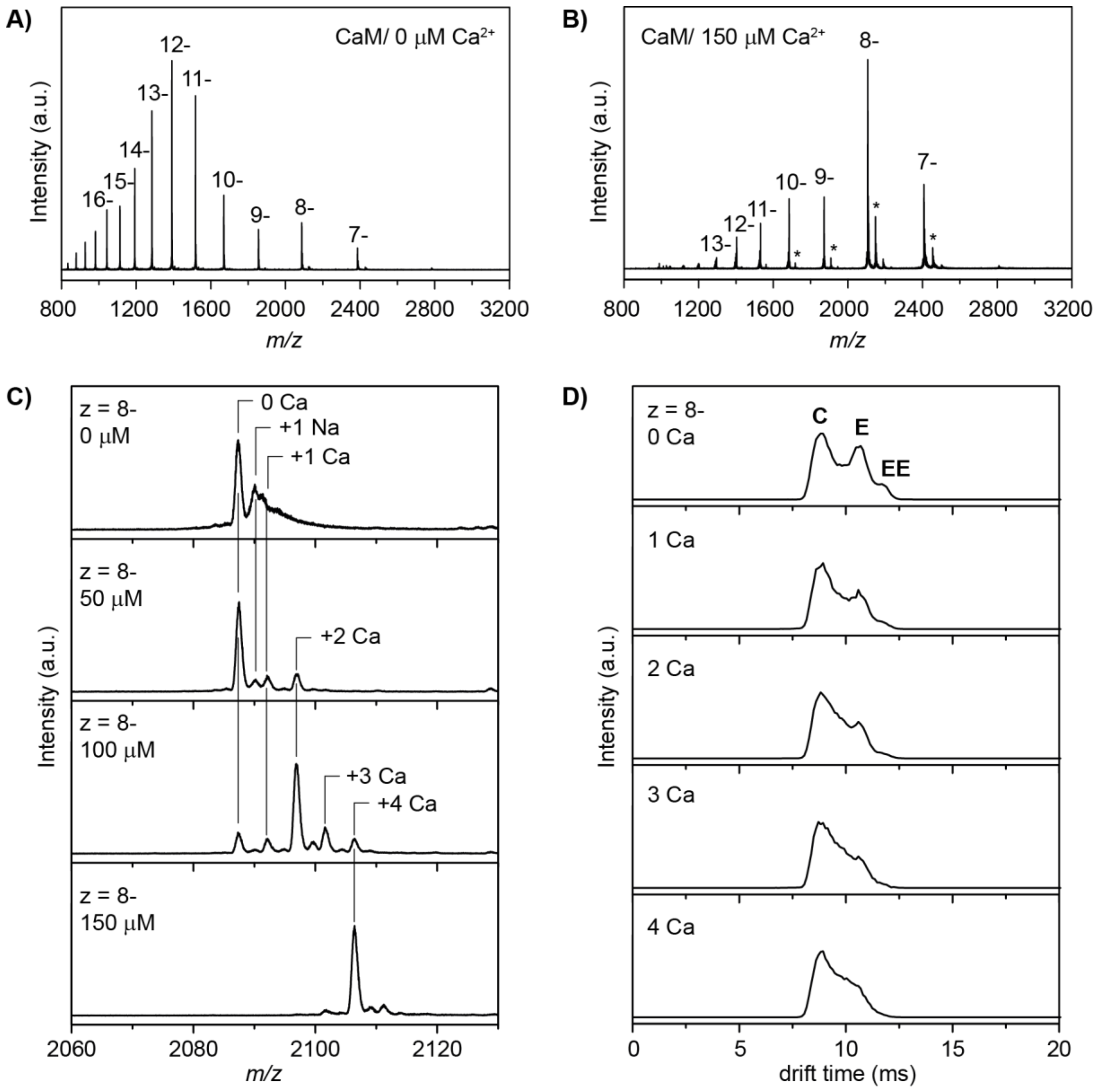


Figure 6.1: Negative ion mode nESI mass spectra of CaM in aqueous ammonium acetate (10 mM) **A)** without and **B)** with additional Ca^{2+} in the spray solution. Upon Ca^{2+}-binding, the intensities in the mass spectrum shift towards lower charge states. Asterisks indicate binding of the chelator EDTA to CaM. EDTA- adducts appear only when Ca^{2+} was added. **C)** Blow-up of z = 8- for different amounts of calcium acetate added to the spray solution. With 150 µM Ca^{2+} added to the solution, a CaM complex with four Ca^{2+} forms exclusively. **D)** Arrival time distributions (ATDs) of z = 8- for Ca^{2+} complexes with n = 0 to 4. The intensity of extended (E) and highly extended (EE) conformations decreases upon Ca^{2+}-binding.

Figure 6.1 shows two mass spectra of CaM, measured in negative ion mode, where either no Ca^{2+} (A) or 150 µM Ca^{2+} (B) were added to the spray solution. Without any additional Ca^{2+} the mass spectrum shows a rather broad charge-state distribution with a maximum intensity at z = 12-. Although mainly apo-CaM is detected from this solution, complex species with one Na^{+} or one Ca^{2+} are also formed, as it is indicated for z = 8- in Figure 6.1C. Upon addition of calcium acetate to the solution the formation of higher Ca^{2+} complexes is observed and with 150 µM of calcium acetate the CaM/Ca_4^{2+} complex is the main species (Figure 6.1C, bottom). The corresponding mass spectrum (Figure 6.1B) shows a shift in intensities towards lower charge states with a maximum at charge state 8- and a slightly narrower charge-state distribution.

The conformational analysis of apo-CaM and CaM/Ca_4^{2+} complexes with IM-MS shows two main populations for most charge states. These two conformations can be assigned to globule (compact) and dumbbell-like (extended) structures, which, using X-ray crystallography, have previously been shown to coexist under certain conditions.[126, 272, 273] Figure 6.1D exemplarily shows arrival time distributions of CaM at charge state 8- with increasing amounts of bound Ca^{2+} (n = 0 - 4). Apo-CaM (0 Ca) adopts three major populations: a compact conformer (**C**), an extended, dumbbell-like conformer (**E**) and a highly extended conformer (**EE**), which was previously referred to as an annealed dumbbell gas-phase structure.[126] Upon Ca^{2+}-binding the intensity of extended conformations decreases significantly. It is assumed that Ca^{2+}-binding stabilizes the globular structure of CaM, whereas the dumbbell-like structure collapses. The data shown here and previous studies on CaM and its Ca^{2+} complexes support this assumption and strongly indicate that structural motifs from solution are memorized in the gas phase. Furthermore, in comparison with X-ray and NMR structures, the obtained gas-phase structure of CaM/Ca_4^{2+} complexes from a previous study slightly agrees better with X-ray data than with structures from NMR.[126]

6.3.2 IM-MS as a Tool for Conformational Screening

The gas-phase structures of CaM/peptide complexes have been studied for a variety of targets with IM-MS.[126, 270, 271, 274] In most cases the complexes adopt compact structures, which are in analogy to the prototypic canonical geometry of NO synthase peptide complex (PDB entry: 2O60). An extended complex structure has been observed only in one case with IM-MS.[270]

A recent cross-linking mass spectrometry study on the CaM binding site of 20- to 24-amino acid Munc13 peptides revealed a common binding mode for all Munc13 proteins, which indicates that the Ca^{2+}/CaM-dependent regulation of priming activity is structurally conserved throughout the entire Munc13 protein family.[258, 275] The obtained model structures were best compatible with the compact structures based on the CaM/NO synthase complex, in which CaM wraps around the target peptide by binding through an 1-5-8 motif in an antiparallel orientation. While binding of the N-terminal part of the Munc13 peptides to the C-lobe of CaM was clearly evident in the structural models, the same level of confidence was never achieved between the C-terminus of the peptides and the N-lobe of CaM. Therefore, selected Munc13 peptides were C-terminally elongated and further experiments revealed an additional binding site between the peptide and the N-lobe of CaM leading to the discovery of a 1-5-8-26 binding motif in Munc13-1 and ubMunc13-2. Furthermore, NMR experiments of these complexes reveal a largely extended structure.[259] This result led to the question if the CaM-binding motif in Munc13s is unique. The application of IM-MS to elongated CaM-binding peptides derived from Munc13 proteins and other CaM targets is expected to shed light on the question whether extended CaM conformations and sequential CaM binding are unique to Munc13s or a rather general feature of (presynaptic) CaM targets.

Therefore, CCS values shall be experimentally determined for CaM and CaM/peptide complexes, and be subsequently compared as a function of molecular mass. Figure 6.2 shows such a theoretical plot with helium drift tube CCSs ($^{DT}CCS_{He}$) of the compact and dumbbell conformation of apo-CaM. The estimated mass-to-CCS correlation for compact conformations is assumed to behave like an idealized spherical particle (CCS ~ $M^{2/3}$),[276] whereas extended conformations are estimated to grow in cylindrical-like (dumbbell) shape (CCS ~ M). Thus, individual CaM/peptide-complexes can be distinguished on the basis of the trend line they follow.

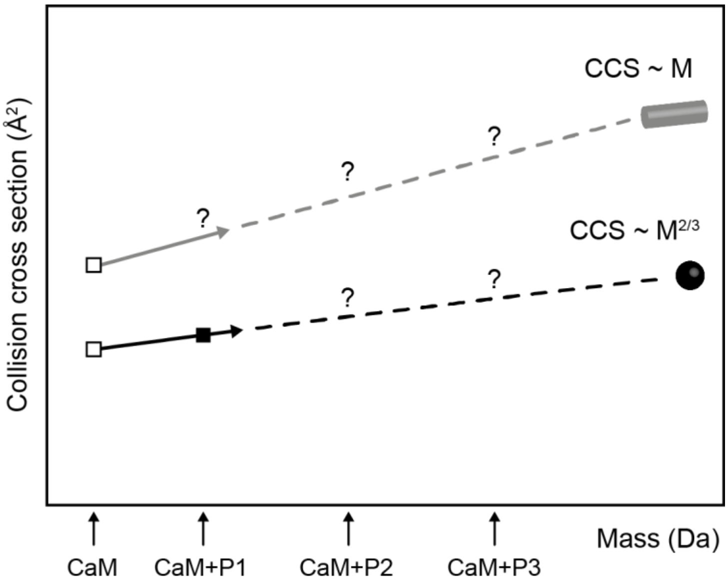


Figure 6.2: Collision cross sections (CCSs) of CaM and CaM/peptide complexes as a function of their protein mass. CCSs of the compact and dumbbell conformations of CaM are shown as open squares. The estimated mass-to-CCS correlation of idealized spherical particles and idealized cylindrical particles are shown as black and grey lines, respectively. The trend lines provide an indication whether the investigated complexes formed by CaM and different peptides (P1, P2, P3) are of the compact or extended type.

In a pilot experiment, a complex between CaM and two different peptides derived from the CaM-binding region of skMLCK were analyzed. The shorter peptide skMLCK$_{577\text{-}602}$ contains a prototypic 1-5-8-14 binding motif and leads to a compact complex (Figure 6.3A).[253] The longer, C-terminally elongated peptide skMLCK$_{575\text{-}607}$ contains an additional hydrophobic amino acid in motif position 26 (Leu 605), similar to the 1-5-8-26 motif of Munc13-1 and ubMunc13-2. It should be evaluated whether the peptide elongation leads to a CaM conformation that is compact (as indicated by previous cross-linking experiments[277]) or extended. Mass spectra of both complexes are shown in Figure 6.3B and C and reveal a clear preference for a 1:1:4 CaM/peptide/Ca_4^{2+} stoichiometry and a narrow charge state distribution with the most abundant complex intensities in charge state 8- and 9-. Figure 6.3D shows ATDs of these two charge states for CaM/Ca_4^{2+} (black) and for the corresponding complexes with skMLCK$_{577\text{-}602}$ (red) and skMLCK$_{575\text{-}607}$ (blue), respectively. In general, ATDs for both complexes appear narrower and more defined than those for CaM/Ca_4^{2+}.

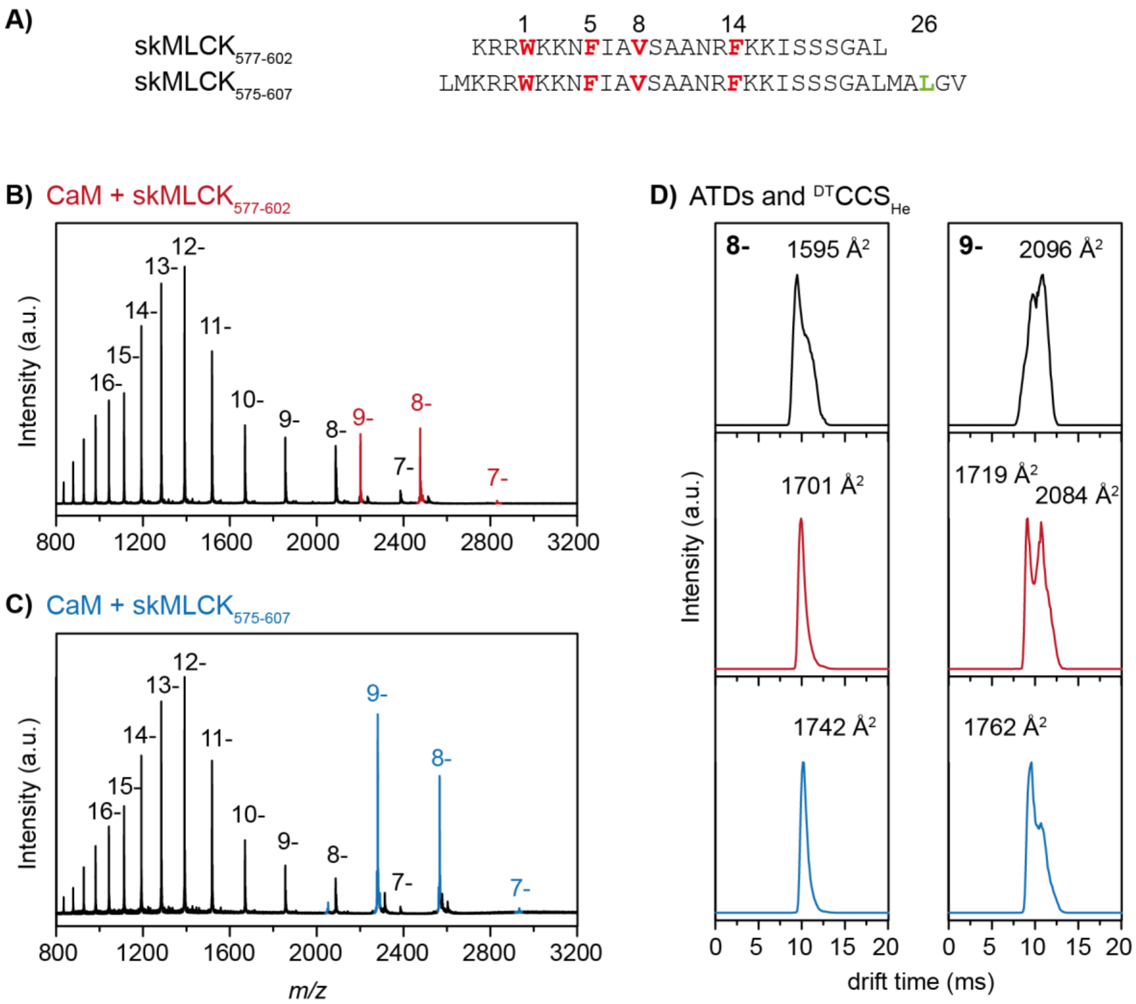


Figure 6.3: A) Amino acid sequence of the investigated peptides skMLCK$_{577\text{-}602}$ and skMLCK$_{575\text{-}607}$. Hydrophobic anchor residues of the 1-5-8-14 CaM-binding motif position are annotated in red, while a potential hydrophobic contact site in motif position 26 of skMLCK$_{575\text{-}607}$ is annotated in green. **B)**, **C)** Negative ion mass spectra of CaM/peptide complexes. Spectra were derived from 10 µM CaM without external addition of Ca^{2+} and in presence of B) 20 µM skMLCK$_{577\text{-}602}$ or C) 20 µM skMLCK$_{575\text{-}607}$, respectively. Charge states of bare CaM are shown in black, whereas the charge states of CaM/skMLCK$_{577\text{-}602}$ and the CaM/skMLCK$_{575\text{-}607}$ complex are highlighted in red and blue. In all cases the ratio of bound complex is Ca^{2+}/CaM/peptide 4:1:1. **D)** Arrival time distributions (ATDs) and experimental helium CCSs (^{DT}CCS$_{He}$) of CaM (black) and corresponding CaM/peptide complexes with four Ca^{2+} in charge state 8- and 9-. Notably, in both charge states the experimental ^{DT}CCS$_{He}$ of the CaM/skMLCK$_{575\text{-}607}$ complex does not differ significantly from that of the CaM/skMLCK$_{577\text{-}602}$ complex, identifying both CaM/peptide complexes as compact type.

The experimentally determined CCSs are similar and identify both complexes as compact type, despite the elongated amino acid sequence and the resulting ~20 % gain in mass with skMLCK$_{575\text{-}607}$ in comparison to skMLCK$_{577\text{-}602}$. Interestingly, for charge state 9-, two distinct conformations are clearly observed for CaM and both CaM/peptide complexes. In particular, the ATD of the $\text{skMLCK}_{577\text{-}602}$ complex exhibits two distinct features, which are roughly similar in intensity, whereas in the ATD of the $\text{skMLCK}_{575\text{-}607}$ complex, the extended feature is much less abundant. This result is in good agreement with previous cross-linking experiments[277] and indicates that CaM/ $\text{skMLCK}_{575\text{-}607}$ is even more compact than the complex with the short skMLCK peptide.

Taken together, the results of the pilot experiment indicate that IM-MS can provide both, the throughput and the resolving power for straightforward conformational screening of CaM/peptide complexes, and application to CaM/Munc13 peptide complexes is outlined in the next section.

6.3.3 Conformational Screening of CaM/Munc13 Complexes

As mentioned briefly in the introduction of this chapter, the Munc13 family comprises four neuronal isoforms: Munc13-1, ubMunc13-2, bMunc13-2, and Munc13-3 (Figure 6.4). Munc13-1 is the major isoform and is expressed essentially in all neurons. Munc13s are regulated in an activity-dependent manner *via* a diacylglycerol- and phorbol ester-binding C1 domain, a Ca^{2+}/CaM-binding domain, and a C2 domain that binds to Ca^{2+} and phospholipids. While all Munc13 proteins share a highly homologous C-terminus, only Munc13-1 and ubMunc13-2 are also conserved N-terminally of the C1 domain (Figure 6.4). Three peptides of the major isoform Munc13-1 were investigated and the corresponding amino acid sequences are shown in Figure 6.5A. In addition to the minimal binding motif with 21 amino acids (13-1), the C-terminally elongated peptide (13-1C, 34 residues), and a CaM binding-deficient mutant sequence, where one of the hydrophobic tryptophan anchor residues was replaced by arginine (13-1M),[256, 278] were analyzed. When spraying the CaM/Munc13-1 complexes from a 10 mM ammonium acetate solution with 150 μM calcium acetate, the mass spectra show a similar charge-state distribution for all three CaM/peptide complexes (Figure 6.5B, C, and Figure C.1A) with a clear preference for a 1:1:4 CaM/peptide/Ca_4^{2+} complex (inset Figure 6.5B). The most dominant charge states are $z = 8$- and $z = 9$-. Furthermore, under the conditions used (10 μM CaM and 20 μM peptide), apo-CaM is not observed in the mass spectra. Figure C.1A shows that the mutated variant of 13-1 can still bind to CaM and that the appearance in the mass spectrum strongly resembles the complex with the wild type peptide.

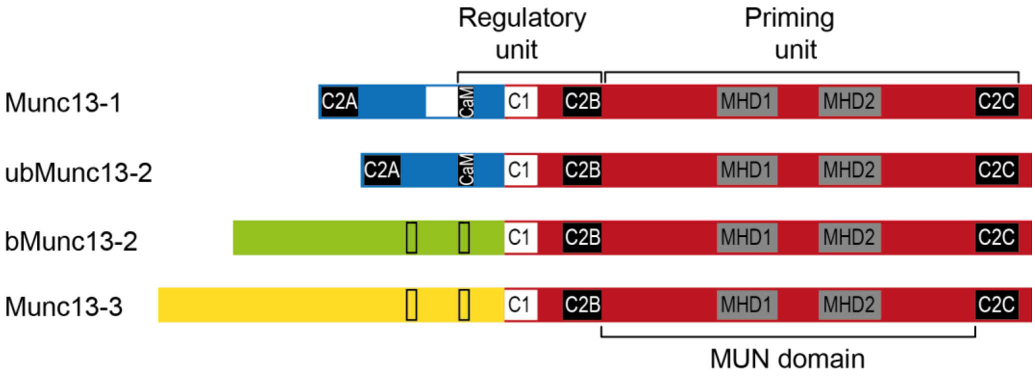


Figure 6.4: Domain structure of Munc13 proteins. All Munc13 proteins share a highly homologous C-terminus (red region) with two Munc13 homology domains (MHDs), one diacylglycerol/phorbolester binding site (C1) and two C2 domains. Munc13-1 and ubMunc13-2 are also conserved N-terminally of the C1 domain (L-region, blue), whereas the N-termini of bMunc13-2 (green) and Munc13-3 (yellow) are unrelated. CaM-binding sites are represented by black bars in the N-terminal region of the proteins. Whereas binding sites in Munc13-1 and ubMunc13-2 are established and conserved, corresponding binding sites in bMunc13-2 and Munc13-3 have been only recently identified and are not conserved. Figure adapted with permission from [279]. Copyright 2012 American Society for Microbiology

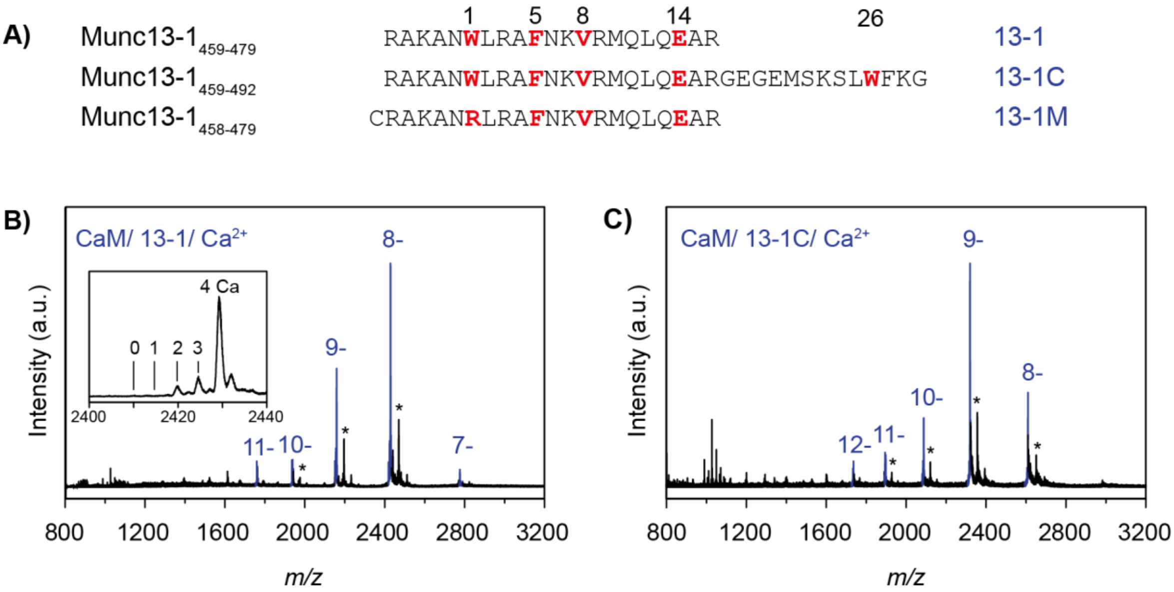


Figure 6.5: A) Amino acid sequences of the investigated Munc13-1 peptides. The minimal binding motif Munc13-1$_{459\text{-}479}$, 13-1), and a C-terminally elongated version (Munc13-1$_{459\text{-}492}$, 13-1C) as well as the mutated sequence Munc13-1$_{458\text{-}479}$, 13-1M) were analyzed with CaM. Hydrophobic anchor residues are indicated in red. **B) C)** Mass spectra of CaM/peptide/Ca^{2+} complexes with B) the peptide 13-1 and C) the C-terminally elongated 13-1C. Concentrations of 10 µM:20 µM:150 µM CaM/peptide/Ca^{2+} were applied. The complex stoichiometry in the spectrum is with both peptides 1:1:4. Asterisks indicate species where the chelator EDTA binds to the complex.

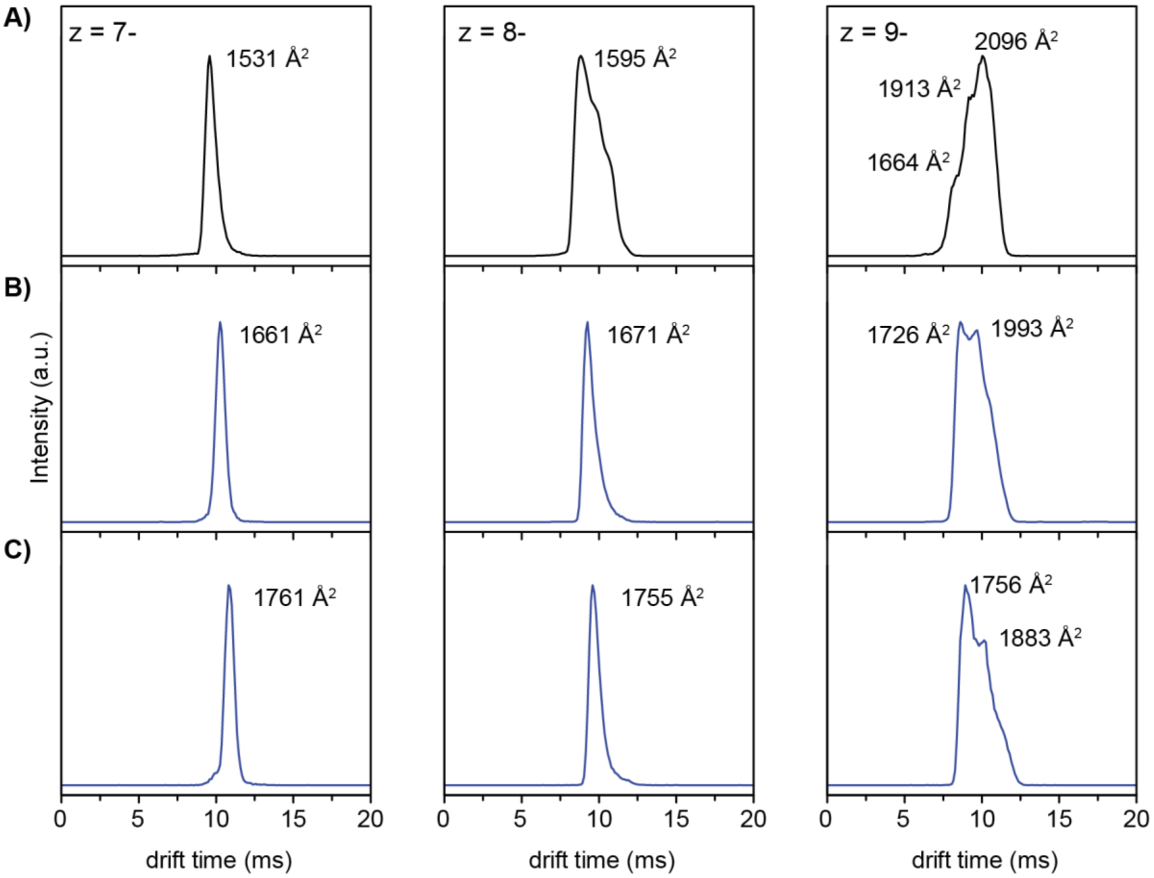


Figure 6.6: Arrival time distributions (ATDs) of CaM/Ca_4^{2+} (black) complexes **(A)** and CaM/peptide/Ca_4^{2+} complexes (blue) with **B)** peptide 13-1 and **C)** peptide 13-1C in charge state 7-, 8- and 9- respectively. $^{DT}CCS_{He}$ values were determined for the most dominant conformations.

Inspection of the ATDs of CaM and CaM/Munc13-1 complexes gives indications on their gas-phase structures. Figure 6.6 presents an overview of ATDs from A) CaM/Ca_4^{2+}, B) CaM/13-1/Ca_4^{2+}, and C) CaM/13-1C/Ca_4^{2+} in charge states 7-, 8-, and 9-. The determined collision cross section ($^{DT}CCS_{He}$) values are given in Å². ATDs of the CaM/13-1M complex and corresponding CCS values are shown in Figure C.1B.

In case of all complexes, charge state 7- adopts only one conformation with a very narrow peak shape (Figure 6.6, left column and Figure C.1B). While extended conformations are observed as shoulders in a broad ATD for CaM/Ca_4^{2+} complex in charge state 8-, the corresponding ATDs appear much narrower for all three CaM/Munc13-1 complexes with only one major population (Figure 6.6, middle column and Figure C.1B). In charge state 9-, the ATDs are generally broader with multiple features and in case of CaM (A), the assignment of clear conformations is somewhat challenging. Complexes with peptide 13-1 and 13-1C in charge state 9- show two major conformations, whereby the compact conformation (1756 Å²) forms with higher intensity in the complex with the C-terminally elongated 13-1C peptide. The ATDs of the complex with 13-1M resembles the complex with 13-1, however the features are slightly better separated. A comparison of the determined CCS values shows that they are in general quite similar for charge states 7- and 8- (~ 1600 Å^2 for CaM, ~1670 Å^2 for 13-1, ~1760 Å^2 for 13-1C complexes, and ~ 1650 Å^2 for 13-1M, respectively), and in case of the two CaM/Munc13-1 complexes shown in Figure 6.6 similar values are as well determined for the compact conformations of charge state 9- (1726 Å^2 and 1756 Å^2, respectively). Furthermore, CCS values increase when a target peptide is bound to CaM (*e.g.* z = 7-: 1531 Å^2 to 1661 Å^2). However, in a relative comparison this increase in CCS is lower compared to the gain in mass. For example, the mass increase is 15 % when 13-1 binds to CaM and 24 % when the 13-1C peptide binds to the protein, with an increase in CCS from 4-9 % and 6-15 %, respectively (calculated for z = 7-, 8-, and compact conformation of 9-). This indicates that the complexes with Munc13-1 derived peptides adopt rather compact gas-phase structures. Furthermore, comparing the complexes with 13-1 and 13-1C to each other, the gain in mass from a complex with the short peptide to a complex with the C-terminally elongated peptide is 7 %, whereas the increase in CCS is only 2-6 %. This in turn leads to the assumption that CaM does not adopt an extended structure with the elongated 13-1C in the gas phase, but even forms a more compact complex. IM-MS data of the CaM/mutant peptide complex indicate that the substitution of one amino acid in the peptide does not alter the complex structure significantly. Before drawing any further conclusions from these data, CaM/peptide complexes with the three other isoforms ubMunc13-2, bMunc13-2, and Munc13-3 are analyzed with IM-MS.

Figure 6.7A presents the amino acid sequences of the three isoforms and a color code for the abbreviation of the different peptides. Indicated in red are the mutated amino acids, which were replaced to generate CaM binding-deficient control peptides.[279] In case of ubMunc13-2 and Munc13-3 only one amino acid is replaced (ub13-2M, blue; 13-3M, yellow), whereas in the sequence of bMunc13-2 four amino acids were substituted (b13-2M, green). Figure 6.7B shows the corresponding mass spectra of CaM/peptide complexes with short and C-terminally elongated peptides and 150 µM Ca^{2+}. Mass spectra of the mutant complexes are shown in Figure C.2. In general, the expected CaM/peptide complexes are formed and the mass spectra show a charge-state distribution from z = ~7- to ~11-, with z = 8- and 9- being the most dominant charge states.

A)

ubMunc13-2$_{382-402}$	CQARAHWFRAVTKVRLQLQEIS	ub13-2
ubMunc13-2$_{383-415}$	QARAHWFRAVTKVRLQLQEISDDGDPSLPQWLPE	ub13-2C
ubMunc13-2$_{382-402}$	CQARAHRFRAVTKVRLQLQEIS	ub13-2M
bMunc13-2$_{720-742}$	INNFKNVLREKRLRQKKLLQELV	b13-2
bMunc13-2$_{720-752}$	INNFKNVLREKRLRQKKLLQELVQTASHLSVED	b13-2C
bMunc13-2$_{719-742}$	CINNRENVLEEKELRQKKLLQELV	b13-2M
Munc13-3$_{962-980}$	SFKEAALRAYKKQMAELEEK	13-3
Munc13-3$_{962-991}$	SFKEAALRAYKKQMAELEEKILAGDSSSVDE	13-3C
Munc13-3$_{962-980}$	SRKEAALRAYKKQMAELEEK	13-3M

B)

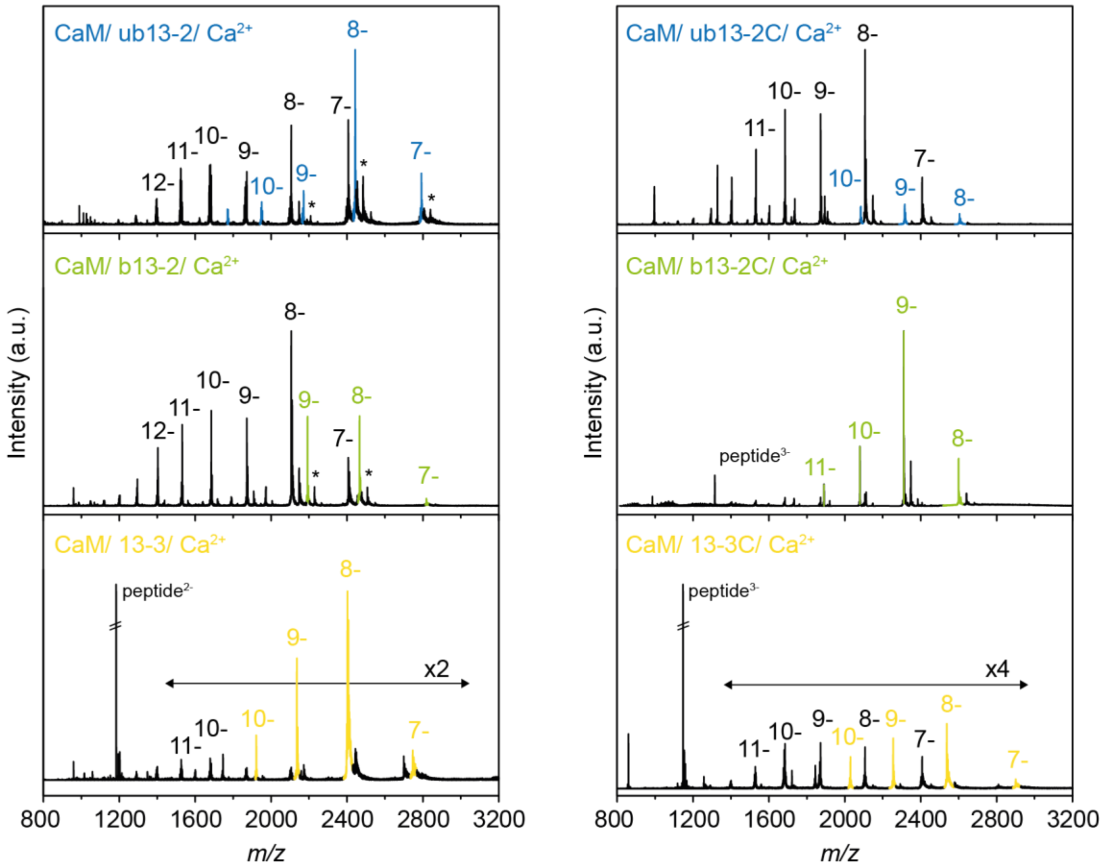


Figure 6.7: A) Amino acid sequences and abbreviation codes of the following three isoforms of Munc13 derived peptides: ubMunc13-2 (blue), bMunc13-2 (green), and Munc13-3 (yellow). Three different peptides (short, elongated, mutant) are investigated for each isoform. Mutated amino acids are indicated in red. **B)** nESI mass spectra of CaM/ peptide/ Ca^{2+} (10 µM: 20 µM: 150 µM) complexes with ubMunc13-2 (blue), bMunc13-2 (green), and Munc13-3 (yellow). Spectra of complexes with the short peptides (left column) and C-terminally elongated peptides (right column) are shown.

There are likely differences in the binding affinity of the three peptide isoforms, but free CaM (with four Ca^{2+} bound) is still detected in most spectra in contrast to previously analyzed CaM/Munc13-1 complexes. In all cases a 1:1:4 CaM/peptide/Ca_4^{2+} stoichiometry is formed predominantly. The mutant peptides also bind to CaM, even in the case where four amino acids are substituted in the peptide (b13-2M), although apparently with low affinity. Quantitative binding in case of 13-1M and ub13-2M was however a rather unexpected result, as in a previous publication a decreased affinity to CaM was revealed for all four mutant peptides by photo affinity labeling experiments,[278] and CaM binding was even completely abolished with 13-1 and ub13-2 mutants on the protein level.[256] As discussed in Chapter 2.4, affinities measured in the gas phase do in some cases not reflect affinities measured in solution due to altering interaction strengths in vacuum. This may apply for the CaM/Munc13 system, resulting in slightly differing affinities and peak intensities. Obviously, this system presents an example where care has to be taken when comparing affinities from solution to such determined in the gas-phase.

CCS values in helium were determined as well for these complex species in several charge states and are listed in Table C.2. The shape of the corresponding ATDs resembles those of the CaM/Munc13-1 complexes in Figure 6.6, with rather narrow ATDs for $z = 7$- and 8- showing mainly only one conformational family and $z = 9$- being broader with several features. To give a qualitative overview and to decide whether the analyzed complexes belong to the compact or extended type, the determined CCS values are plotted against the mass of the respective CaM/peptide/Ca_4^{2+} complex (Figure 6.8). In this plot the CaM/mutant peptide complexes are not considered. However, as the determined CCSs values (Table C.1 and Table C2) are similar to those of complexes with the short peptides, the complexes with mutant peptides are assumed to be located also in a similar range.

The CCSs of the charge states 7-, 8- and the compact conformation of 9- were averaged. These charge states were selected, as the corresponding CCSs lie all in a very similar range and therefore it can be assumed that they correspond to native-like structures. Extended conformations of $z = 9$- were not considered in this diagram, as the ATDs in this charge state often show multiple, not well-resolved features and it is uncertain whether the corresponding conformations are comparable. The estimated mass-to-CCS correlation of idealized spherical particles is shown as black dotted line for the CaM/Ca_4^{2+} complex using an effective density of $d_{eff} = 0.67$ g cm^{-3} (Figure 6.8, Table C.3).[64, 79]

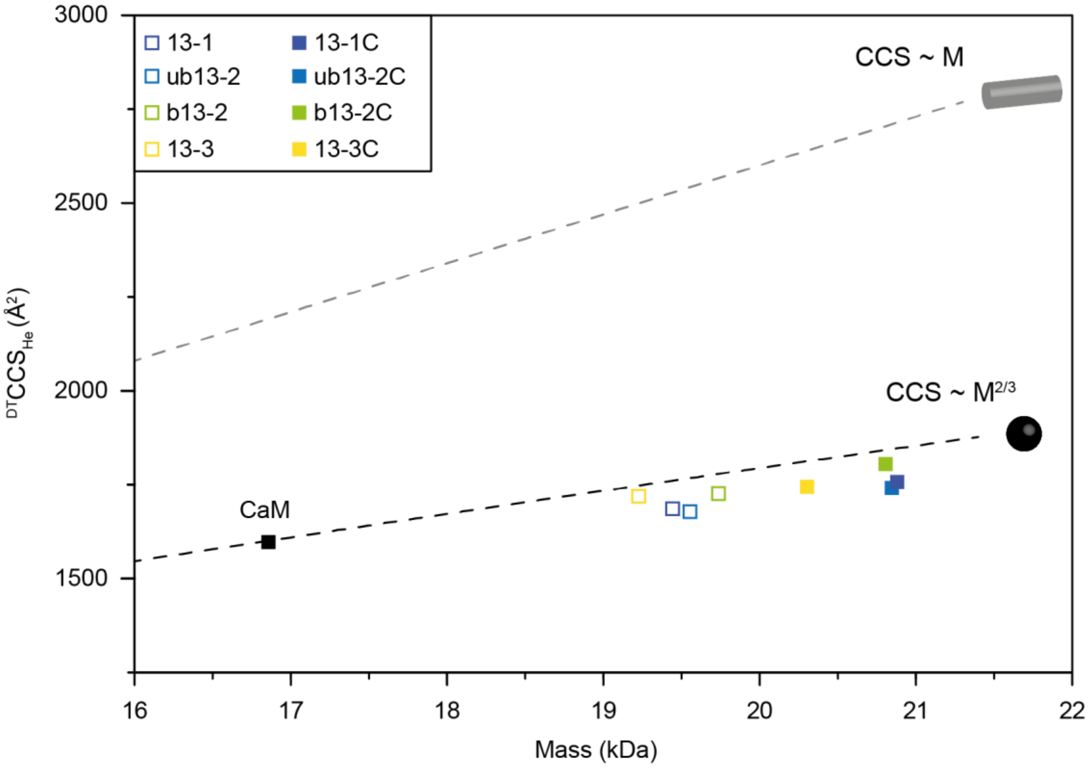


Figure 6.8: Collision cross sections ($^{DT}CCS_{He}$) of CaM/Ca_4^{2+} (black square) and CaM/peptide/Ca_4^{2+} complexes as a function of mass. The estimated mass-to-CCS correlation of idealized spherical particles and idealized cylindrical particles are shown as black and grey dotted trend lines, respectively. Complexes with short peptides are indicated as open squares, whereas complexes with C-terminally elongated peptides are shown as filled squares. The color code is as follows: CaM, black; Munc13-1, dark blue; ubMunc13-2, blue; bMunc13-2, green; Munc13-3, yellow. All complexes are located below the black trend line, thus being of compact type.

A CCS of 1601 Å^2 was calculated from this density for CaM/Ca_4^{2+} and agrees well with the experimentally determined CCS of 1597 Å^2 (average of 7-, 8-, and 9-). The obtained data are also in agreement with previously calculated values for the native-like structure of this complex by exact hard sphere scattering (EHSS)[73] of 1680 Å^2 and 2060 Å^2,[126] when considering that EHSS typically overestimates protein CCSs in the gas phase.[280] When inspecting Figure 6.8 in more detail, it is striking that CaM complexes with short peptides (open squares) as well as complexes with C-terminally elongated peptides (filled squares) are all located below the black trend line. These results imply that not only the complexes with short peptides are even more compact than assumed for an idealized spherical particle, but also the complexes with C-terminally elongated peptides adopt globular structures. Although there are some minor differences between the CCSs of different complexes, in total they adopt roughly similar gas-phase structures.

Previous NMR experiments of CaM/13-1C revealed a characteristic extended structure in complex with this elongated peptide.[259] A similar pattern of interaction was obtained for the complex with ub13-2C by photoaffinity labeling and NMR titration within the same study, which indicates a similar structure of the complex with the highly homologous ub13-2C. Unfortunately, complexes with both peptides do not adopt extended structures in the gas phase, which implies a huge discrepancy between solution and gas-phase structures. How can the structural collapse in the gas phase be explained? In a previous study on CaM/peptide complexes[126] it was proposed that a CaM dumbbell-like structure is only stable with a weakly bound peptide. This could for example be the case with very small peptides, which bind only with one lobe of CaM. In case of larger peptides (~20 residues), on the other hand, the globular CaM structure provides a much better binding site involving both lobes. The energy cost due to the conversion would be a downside, however, prior formation of a compact CaM complex is possible when Ca^{2+} is present and this lowers the energy barrier of a compact structure. In other words, this means that dumbbell-like structures are preferentially formed with peptides that bind only one lobe of CaM, while longer peptides rather adopt compact structures in complex with CaM. Although this assumption sounds quite plausible, it stands in contradiction to the results on the peptide complexes with minimal binding motif (*e.g.* 13-1, ub13-2): They only bind with one lobe of CaM and nonetheless adopt compact complexes in the gas phase.

The general structural flexibility of the helical linker could provide another explanation. It is assumed that extended structures are formed in solution, but the linker, which connects N-and C-lobe of CaM is too flexible and closes up upon transfer into the gas phase, which results in a collapsed compact structure. Such a gas phase collapse has been reported before with the GroEL complex.[281] In this case, it would also explain why CaM-intermediate structures from NMR experiments have not been detected by IM-MS.

6.3.4 Following the Unfolding and Dissociation of CaM/Munc13 Complexes

Following the unfolding and dissociation pathway of a complex with IM-MS can give further structural information and reveal potential differences between the investigated complexes. This is accomplished by collision-induced unfolding (CIU) and collision-induced dissociation (CID) experiments. Therefore, one charge state of interest is *m/z*-selected as precursor and subsequently activated by increasing the collision voltage. This induces unfolding and finally dissociation of the complex. The corresponding ATD is recorded for several different voltages and subsequently plotted for every applied collision voltage to give a contour plot, which illustrates the unfolding of the complex at a specific charge state (see Chapter 2.4.4 for details). A comparison of unfolding plots of different complexes can shed light on differences in gas-phase (intermediate) structures and binding events.[282]

Figure 6.9 shows contour plots exemplarily for the CaM/13-1/Ca_4^{2+} complex in charge states 7-, 8-, and 9- (from left to right). The experiment was also performed for a precursor ion in charge state 10-, however, for this ion the signal was too unstable and thus reliable data could not be generated. A comparison of the three charge states presented in Figure 6.9 illustrates that z = 7- is the most stable charge state, whereas z = 9- is the most unstable charge state. This behavior seems quite reasonable, as higher charge states have a higher charge density and therefore a larger Coulomb repulsion,[208] which likely supports unfolding at an earlier stage. In case of z = 7- the contour plot shows a quite narrow ATD for low and high collision voltages with only one unfolding event at ~15 V. The general structural rearrangement can be considered as relatively small (CCS not determined). The contour plot of charge state 8- also shows a narrow ATD at low collision voltages, however, starting from around 10-12 V, a gradual but significant unfolding is induced until the fully unfolded structure (2081 Å^2) is obtained at ~20-25 V. A similar unfolding pathway is observed for the precursor complex ion in charge state 9-. In this case, however, the ATD is already broader from the beginning and the fully extended (annealed) gas-phase structure dominates already at ~15-17 V. As charge state 7- does not undergo large structural rearrangements and charge state 9- is rather unstable, charge state 8- was selected as precursor for all further measurements.

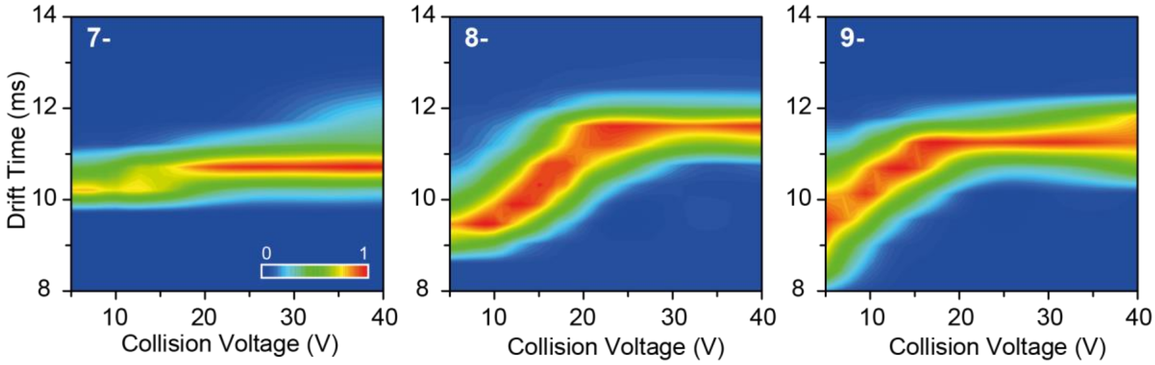


Figure 6.9: Contour plots of CaM/13-1/Ca_4^{2+} complexes in charge states 7-, 8-, and 9-, respectively. The complex species were activated to induce unfolding and dissociation.

Figure 6.10 presents contour plots for precursor complex ions in charge state 8- with short peptides (left column) and C-terminally elongated peptides (right column) for all four Munc13 isoforms, Munc13-1 (A), ubMunc13-2 (B), bMunc13-2 (C), Munc13-3 (D). The unfolding pattern is quite similar for complexes with short peptides: Starting from a narrow ATD, a gradual structural rearrangement occurs, which leads to a highly extended complex species with relatively similar CCS values for all four isoforms (2081-2130 Å^2, Table C.4). Minor differences are detected in the relative unfolding energies, which will be discussed later in this chapter. Comparing unfolding plots of precursor ions with short peptides (*e.g.* CaM/13-1) to precursor ions with the respective elongated peptides (*e.g.* CaM/13-1C), the pattern is quite similar, however the structural rearrangement occurs at a later stage, meaning that these complexes are more stable. This slightly higher stability likely arises from the gain in mass due to the elongated peptide. A comparison of contour plots of complex ions with C-terminally elongated peptides (right column) reveals isoform-dependent differences: While CaM complexes with peptides 13-1C, ub13-2C, and 13-3C (A, B, and D) unfold rather gradually, the complex with b13-2C (C) stands out particularly: Starting from ~12 V one distinct intermediate structure is formed (1951 Å^2), which is highly stable until the fully unfolded species (2155 Å^2) emerges at ~ 30 V. As observed for complexes with short peptides, also the C-terminally elongated peptide complexes adopt similar structure sizes in the fully unfolded state (2116 - 2155 Å^2). Thus, the major structural difference between these complexes is observed in the intermediate structures. How can these differences be explained? Model structures indicate that the N-terminal region of Munc13 peptides interacts with the C-lobe of CaM. In case of the peptides 13-1, ub13-2, and 13-3 this interaction is based on hydrophobic residues in positions 1, 5, and 8, whereas for b13-2 it is based on positions 1, 5, and 10.[275] Unfortunately, within the complexes with short peptides, b13-2 does not show a particular different unfolding pathway than the other isoforms. The fact that this particular stable intermediate structure is only formed in the b13-2C-complex, indicates that elongation of the peptide in case of bMunc13-2 leads to an additional, very strong, interaction and therefore likely to an additional binding contact with CaM. Although, sequence alignments do not support that bMunc13-2, nor Munc13-3 contain a 1-5-8-26 CaM-recognition motif[259] or any other hydrophobic cluster with a characteristic spacing, binding contacts could have been missed and further experiments could help for clarification.

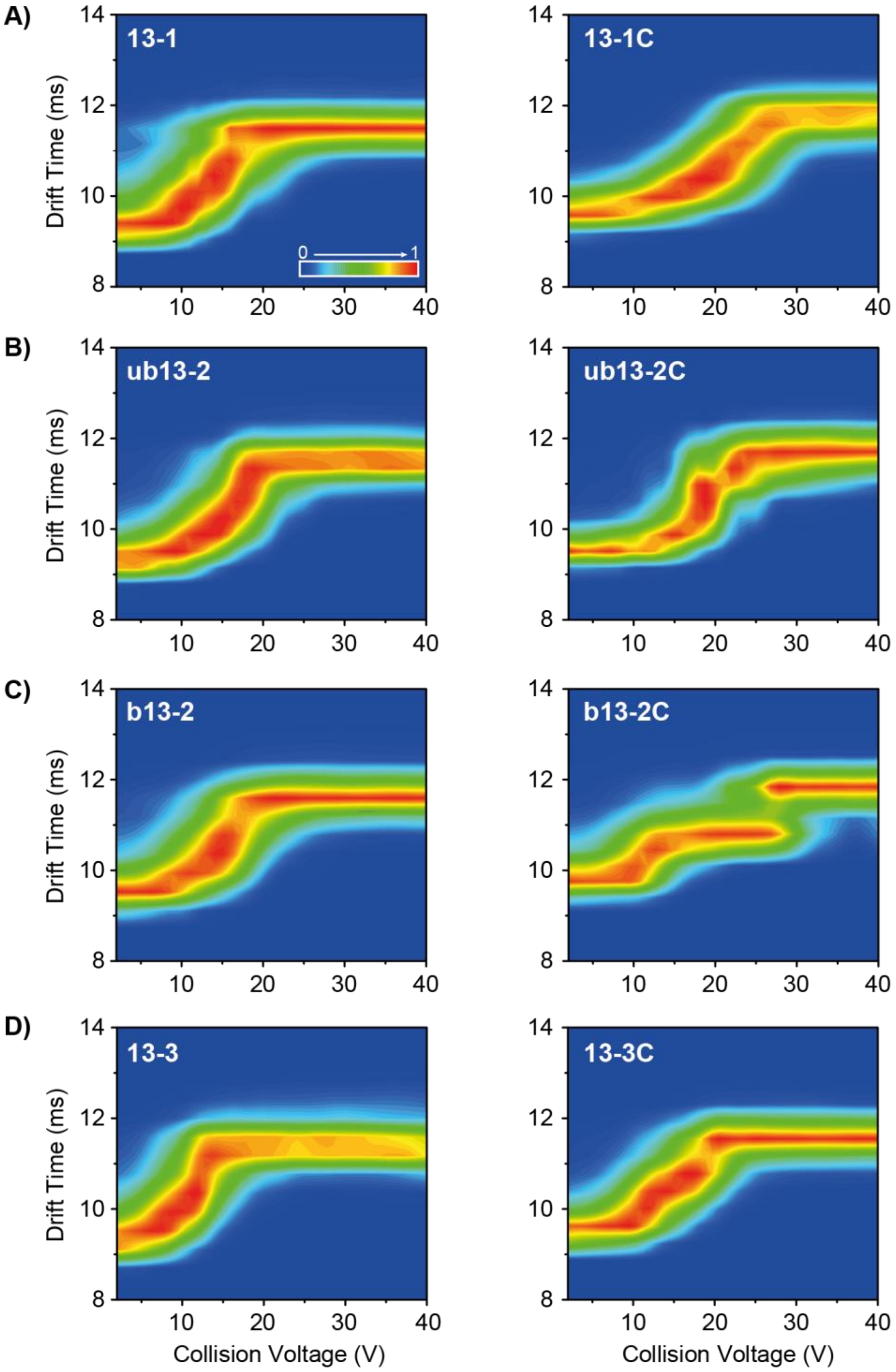


Figure 6.10: Contour plots of CaM/Munc13/Ca_4^{2+} complexes in charge state 8- with short peptides (left column) and C-terminally elongated peptides (right column) for **A)** Munc13-1, **B)** ubMunc13-2, **C)** bMunc13-2, and **D)** Munc13-3.

The fully unfolded structure, as shown in Figure 6.10, remains for all complexes until complete dissociation has occurred or the signal-to-noise ratio is too low at higher collision voltages.

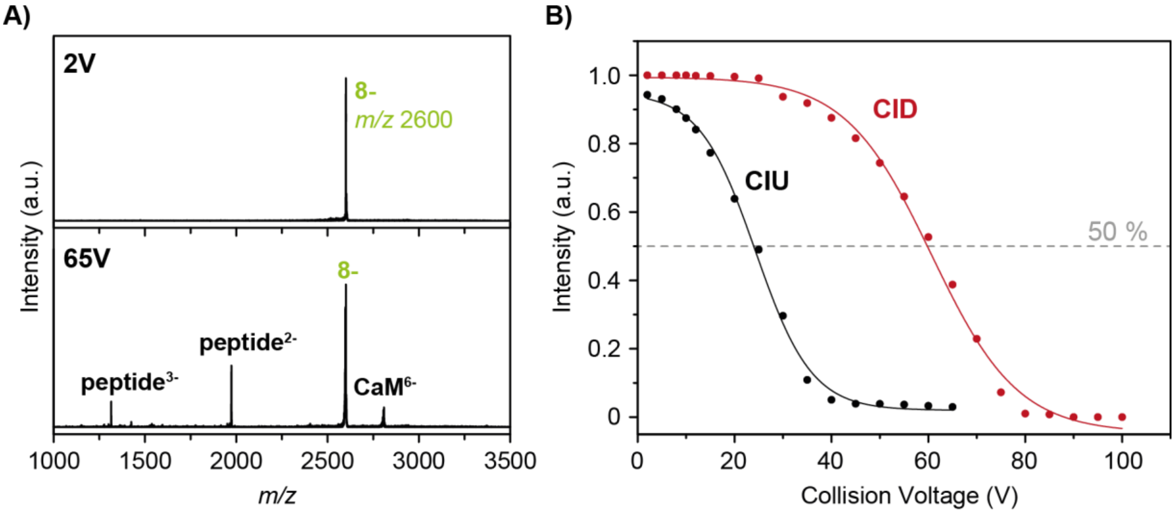


Figure 6.11: A) MSMS spectra of a CaM/b13-2C/Ca_4^{2+} complex precursor ion in charge state 8-. At soft conditions (2 V) the precursor ion is still completely intact, whereas it dissociates at activated conditions (65 V) and peptidic species in different charge states as well as CaM/Ca_4^{2+} complex in charge state 6- are formed as products. **B)** Depletion curves for CIU and CID experiments exemplarily for the CaM/b13-2C/Ca_4^{2+} complex. A Boltzmann fit is applied and a value for 50 % unfolding and 50 % dissociation is determined.

Exemplary MSMS spectra are presented in Figure 6.11A for the CaM/b13-2C/Ca_4^{2+} complex at soft conditions (2V) and at activated conditions (65 V). The precursor ion at *m/z* 2600 is still completely intact at 2 V, but dissociates into peptide species and CaM/Ca_4^{2+} species at higher voltages (65 V). In the exemplary spectra peptides in charge state 2- and 3- as well as CaM/Ca_4^{2+} species in charge state 6- are formed. A similar dissociation pathway was obtained for most of the other investigated complexes (spectra not shown).

For a better comparison of unfolding and dissociation between the CaM/Munc13 complexes, the data can be quantified and a value for 50 % unfolding or 50 % dissociation can be determined. Therefore, the relative unfolding is calculated by using the ratio of the integrated intensity of the initial conformer divided by the integral of the total complex ATD (Figure C.3). The decreasing signal of the initial conformation is subsequently plotted against the applied voltages and a value for 50 % unfolding can be determined using a Boltzmann fit (Figure 6.11B, black data and fit curve). In an analogous manner values for 50 % complex dissociation were determined. Therefore, the intensities of the precursor ions and respective dissociation products are extracted of every recorded tandem mass spectrum.

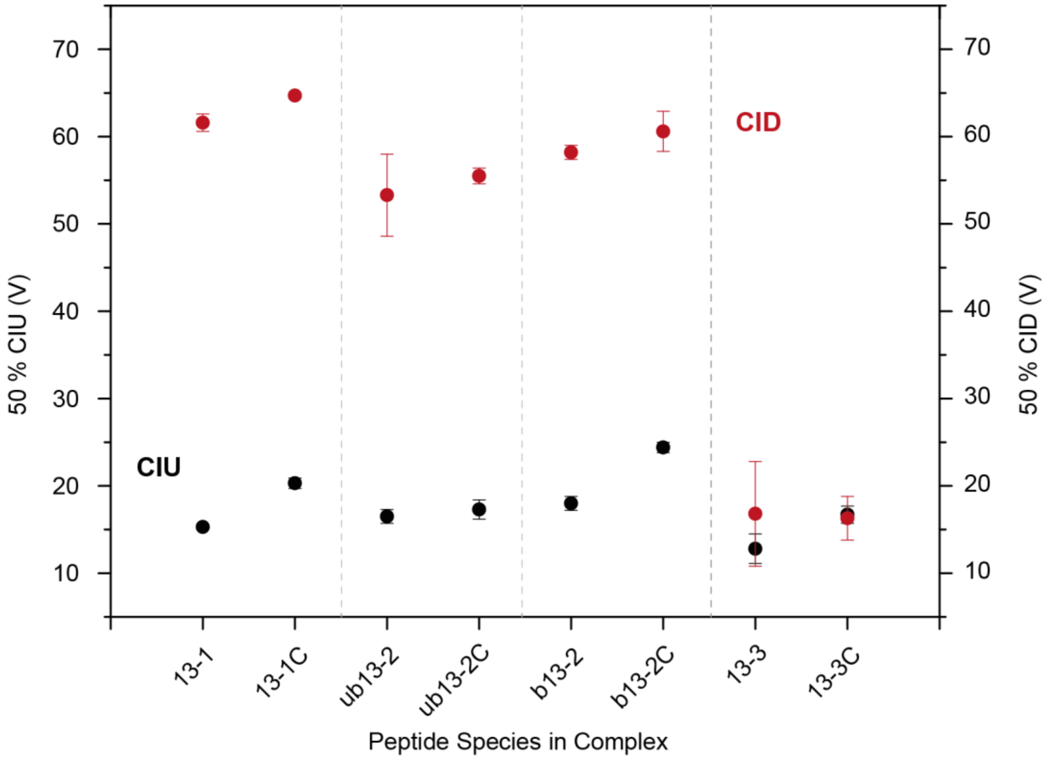


Figure 6.12: 50 % CIU and 50 % CID values (black and red data, respectively) for all investigated CaM/Munc13/Ca_4^{2+} complexes. Values correspond to an average of three measurements and the error corresponds to the standard deviation.

The depletion of the precursor ion with increasing collision voltage is plotted and fitted (Figure 6.11B, red data and fit curve). 50 % values for unfolding and dissociation were calculated for all investigated CaM/Munc13/Ca_4^{2+} complexes and are plotted in Figure 6.12 (exact values are listed in Table C.5). As the measurements were performed in triplicate and are well reproducible (small standard deviations for most calculated values, Table C.5), differences in CIU and CID values can be considered as meaningful. In most cases unfolding and dissociation occurs at higher collision voltages with the C-terminally elongated peptide compared to the complex with the corresponding short peptide. The 50 % dissociation occurs mostly ~30-40 V higher than the 50 % unfolding. One exception is the Munc13-3 isoform, where the corresponding complex dissociates already at very low collision voltages and the values for 50 % unfolding and 50 % dissociation are almost similar for complexes with peptides of this isoform. 13-1C has the highest apparent affinity of all investigated peptides (50 % CIU 64.7 ± 0.3 V), whereas the b13-2-complex has the most stable intermediate structure and therefore the highest relative 50 % CIU value (24.4 ± 0.6 V). Although Munc13-1 and ubMunc13-2 are highly conserved, the peptide sequence of the CaM recognition motif is different, which may also lead to a different affinity and values for relative dissociation identify ubMunc13-2 as the weaker binding isoform. Intriguingly, the energy gap between 50 % unfolding and 50 % dissociation is quite similar within one isoform (*i.e.* 13-1 and 13-1C or ub13-2 and ub13-2C). It is likely that there is a defined energy barrier between unfolding and dissociation that stays constant within the isoform.

6.4 Conclusions

In the experiments described in this chapter, IM-MS has been used as a tool for conformational screening of CaM/Munc13-peptide complexes. The structure of CaM consists of two binding lobes (N- and C-terminal), which are connected by a flexible helical linker. Globular and extended, dumbbell-like structure have been reported for CaM and CaM/Ca^{2+} complexes. The question whether the complexes adopt extended (dumbbell-like) or compact structures, where CaM is wrapped around the peptide in an antiparallel orientation, was addressed. The screening experiment could be successfully performed with an adequate throughput concerning the amount of samples and a straightforward data acquisition and data analysis. This shows that IM-MS in general has the potential for being employed in screening scaffolds. The data reveal that all investigated CaM/Munc13-peptide complexes are of compact type. However, two of the peptides, 13-1C and ub13-2C, were previously confirmed to adopt an extended structure in complex with CaM (and Ca^{2+}) by NMR spectroscopy and photo affinity labeling.[259] This discrepancy leads to the conclusion that all complexes undergo a structural collapse when being transferred into the gas phase. The collapse can probably be attributed to the high flexibility of the helical linker, which energetically favors compact CaM conformations.

The determined CCS values do not differ substantially between the complexes and therefore it can be assumed that they adopt similar gas-phase structures. Unfolding and dissociation studies on the complexes in charge state 8- show an interesting unfolding pathway for a CaM/peptide/Ca$_4^{2+}$-complex with one specific Munc13 isoform, which differs from the gradual unfolding observed in complex with the other peptides. A stable intermediate structure is formed, which likely hints at an additional, very strong, binding contact with CaM, which the other elongated peptides do not share. Unfolding and dissociation were quantified (50 % values) for better comparability. The 13-1C peptide has the highest apparent affinity towards CaM/Ca^{2+}, whereas the homologous isoform ubMunc13-2 binds generally weaker in the gas phase. Munc13-3 is the weakest binding isoform, as the values for 50 % unfolding and dissociation are roughly equal at relatively low voltages.

In summary, it can be concluded that the CaM/Munc13 system is unfortunately not suitable for conformational screening by IM-MS, as the complexes collapse and adopt in all cases a compact structure. Collisional-induced unfolding is a useful technique to visualize structural differences between complex species at activated conditions. Here, it has shown differences in the unfolding pathway, which could hint to additional binding contacts. However, further experiments, ideally with even longer peptides or Munc13 protein fragments, have to be performed to verify this assumption.

7 Summary and Future Perspectives

In this work, protein-ligand complexes were analyzed with native mass spectrometry (MS) and ion mobility-mass spectrometry (IM-MS). Native MS was successfully applied to investigate the influence of the solvent dimethylsulfoxide (DMSO) on a protein's gas-phase structure and protein-ligand affinity. Based on two bromodomain proteins it was shown that low DMSO concentrations (~1-2 %) lead to a shift in the charge-state distribution to higher *m/z* values and therefore likely to a native-like gas-phase structure. In turn, higher DMSO concentrations (starting from ~5 %) most likely cause protein unfolding. In addition to influencing the protein structure, DMSO also affects the binding of ligands to a protein. Titration experiments with a high-affinity binder and varying DMSO concentrations revealed that the binding affinity decreases significantly when increasing the concentration of DMSO from 1 % to 5 %.

Furthermore, a screening experiment against four proteins was performed to evaluate the potential of native MS for fragment-based drug discovery. The hits from MS were compared with results from the well-established thermal shift assay technique and for two of the proteins a good overlap was shown. The results revealed that native MS is an excellent method to analyze protein-ligand complexes in more detail, as it can display direct, indirect and multiple binding. The measureable number of samples is presently limited to a few hundred per day, which identifies native MS as a medium- to low-throughput method.

IM-MS provides information on mass, charge, size, and shape and was used to study the influence of the microenvironment on the protein gas-phase structure. Previously, it had been reported that non-covalent attachment of small crown-ether molecules to positively charged amino acid side chains prevents these side chains from collapsing onto the protein backbone.[127] In this work, crown-ether binding sites on the protein ubiquitin were identified using tandem MS and IM-MS. The results revealed that for residues involved in salt bridges crown-ether binding does not have a crucial effect as these amino acid side chains already stabilize the gas-phase structure. In turn, without salt bridges or other types of contact, the residues contribute to a structural collapse and therefore crown-ether binding influences the gas-phase structure to a greater extent.

In the last part of this thesis, the ability of IM-MS to display conformational changes was exploited to test it as a tool for conformational screening. Protein-peptide complexes with peptides of different length were analyzed and it was evaluated whether they form globular (compact) or dumbbell-like (extended) structures. The screening identified all of the investigated complexes to be of the compact type. A comparison to a previous study with NMR spectroscopy and photo affinity labeling, in which two of the complexes have been revealed to adopt extended conformations,[259] strongly indicates that the investigated complexes undergo a structural collapse when transferred into the gas phase. Differences in complex stability and unfolding behavior were revealed, by using collision-induced unfolding and collision-induced dissociation.

The last two decades have shown immense developments in native MS and related hybrid methods like IM-MS, not only regarding instrumental improvement, but also with respect to the accessibility of challenging targets such as membrane proteins. These improvements make the methods attractive for further applications and for a deeper integration into the drug discovery process. In general, three key factors should be considered for successful future applications: These are a simple and fast sample handling, a certain throughput power as well as the automation of data accumulation and analysis. As it was shown in Chapter 3 the solvent DMSO, in which compounds for screening assays are usually dissolved, can already have an influence on the protein structure and on the protein-ligand affinity. This is especially important to consider when comparing affinities of experiments with differing DMSO levels. Unfortunately, there is probably no equally appropriate substituent in near future, implying that the only solution is to reduce the DMSO concentration in the samples as much as possible. As far as the sample handling of the target protein is concerned, an additional buffer exchange step has to be carried out before electrospray ionization and the instrument conditions currently have to be optimized for every system individually. A new approach that uses submicrometer emitter tips for ESI, which tolerate common buffer salts, could present a way to reduce the amount of time from protein production to characterization.[283] However, proteins can also be taken directly from crude cell lysates to perform native MS experiments.[284] Thus, native MS could also serve as a direct quality control for protein overexpression. In addition to recombinant proteins, the investigation of endogenous protein assemblies with native MS is likely an interesting new area in near future, as it can shed light for example on the proper oligomeric state of protein assemblies or on posttranslational modifications that are rarely found in recombinant proteins.[285, 286]

The throughput power of native MS and IM-MS is currently not good enough to embed them as primary screening methods in high-throughput screening. However, allowing the measurement of several hundreds of samples per day, they can certainly be implemented for hit identification of a low number of samples or for verification of promising compounds from primary screens with other methods. Future technical developments will most probably improve this situation further. A huge step in this regard had already been made by introducing an automatic chip-based electrospray ionization source that allows for automated sample ionization of hundreds of samples.[30]

Automation of data analysis and interpretation is still a bottleneck, as it is challenging to consider all different appearances in a mass spectrum within a single program. Integrated software programs of commercial mass spectrometers allow automated measurements and the before-mentioned chip-based ionization source can be programmed as well, which enables automated data accumulation. Likewise, a series of software packages, such as Amphitrite,[287] UniDec,[288] or CIUsuite,[289] have been developed and can be readily applied for automated analysis of large data sets. However, there are still a couple of challenges arising, *e.g.* from non-specific ligand binding or signal suppression during the ESI process, and it can therefore be cumbersome to identify real binders. A few approaches are already available to identify non-specific binding, such as the reference method[99] or algorithms to calculate and subtract a non-specifically bound fraction.[102] These approaches could be incorporated in software tools for data analysis and combined with routine protocols in which standardized control samples are used. This would make the identification of non-specific binders more straightforward and enable the automation of native MS and IM-MS screening experiments.

Although MS and IM-MS offer a great extent of information on the overall structure, they do not yield high-resolution details. However, structural models can be produced as support using molecular density calculations, even though they are challenging for large protein assemblies. During the last years, the number of approaches for such calculations have tremendously improved so that more accurate studies could be possible in near future. Not only with respect to applications in pharmaceutical industry, but also in a more general sense, the best way to answer biological questions is and will be the combination of different techniques. In this regard, native MS and IM-MS can provide complementary information to other biological methods and can contribute considerably to answer remaining questions in biological systems.

8 Bibliography

[1] J. M. Berg, J. L. Tymoczko, L. Stryer, in *Biochemistry*, 5th edition ed., W.H. Freeman and Company, New York, **2002**.

[2] R. Kiss, Á. Fizil, C. Szántay, What NMR can do in the biopharmaceutical industry, *J. Pharm. Biomed. Anal.* **2017**, doi.org./10.1016/j.jpba.2017.1007.1004

[3] A. L. Carvalho, J. Trincão, M. J. Romão, in *Ligand-Macromolecular Interactions in Drug Discovery: Methods and Protocols* (Ed.: A. C. A. Roque), Humana Press, Totowa, NJ, **2010**, pp. 31-56.

[4] G. Klebe, Applying thermodynamic profiling in lead finding and optimization, *Nat. Rev. Drug Discov.* **2015**, *14*, 95-110.

[5] C. Boozer, G. Kim, S. Cong, H. Guan, T. Londergan, Looking towards label-free biomolecular interaction analysis in a high-throughput format: a review of new surface plasmon resonance technologies, *Curr. Opin. Biotechnol.* **2006**, *17*, 400-405.

[6] M. Wilhelm, J. Schlegl, H. Hahne, A. M. Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx*, et al.*, Mass-spectrometry-based draft of the human proteome, *Nature* **2014**, *509*, 582-587.

[7] D. M. Ramanathan, R. M. Lelacheur, in *Mass Spectrometry in Drug Metabolism and Pharmacokinetics*, John Wiley & Sons, Inc., **2008**, pp. 1-85.

[8] R. H. H. v. d. Heuvel, A. J. R. Heck, Native protein mass spectrometry: from intact oligomers to functional machineries, *Curr. Opin. Chem. Biol.* **2004**, *8*, 519-526.

[9] J. S. Richardson, The Anatomy and Taxonomy of Protein Structure, *Adv. Protein Chem.* **1981**, *34*, 167-339.

[10] C. M. Dobson, Protein folding and misfolding, *Nature* **2003**, *426*, 884-890.

[11] B. Ganem, Y. T. Li, J. D. Henion, Detection of noncovalent receptor-ligand complexes by mass spectrometry, *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.

[12] V. Katta, B. T. Chait, Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry, *J. Am. Chem. Soc.* **1991**, *113*, 8534-8535.

[13] C. V. Robinson, S. E. Radford, Weighing the evidence for structure: electrospray ionization mass spectrometry of proteins, *Structure* **1995**, *3*, 861-865.

[14] J. A. Loo, Studying noncovalent protein complexes by electrospray ionization mass spectrometry, *Mass Spectrom. Rev.* **1997**, *16*, 1-23.

[15] A. C. Leney, A. J. R. Heck, Native Mass Spectrometry: What is in the Name?, *J. Am. Soc. Mass. Spectrom.* **2017**, *28*, 5-13.

[16] A. J. R. Heck, Native mass spectrometry: a bridge between interactomics and structural biology, *Nat. Methods* **2008**, *5*, 927-933.

[17] J. Fenn, M. Mann, C. Meng, S. Wong, C. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, *Science* **1989**, *246*, 64-71.

[18] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.

[19] M. O. Glocker, S. H. J. Bauer, J. Kast, J. Volz, M. Przybylski, Characterization of Specific Noncovalent Protein Complexes by UV Matrix-assisted Laser Desorption Ionization Mass Spectrometry, *J. Mass Spectrom.* **1996**, *31*, 1221-1227.

[20] F. Chen, S. Gerber, K. Heuser, V. M. Korkhov, C. Lizak, S. Mireku, K. P. Locher, R. Zenobi, High-Mass Matrix-Assisted Laser Desorption Ionization-Mass Spectrometry of Integral Membrane Proteins and Their Complexes, *Anal. Chem.* **2013**, *85*, 3483-3488.

[21] F. Chen, B. Gülbakan, S. Weidmann, S. R. Fagerer, A. J. Ibáñez, R. Zenobi, Applying mass spectrometry to study non-covalent biomolecule complexes, *Mass Spectrom. Rev.* **2016**, *35*, 48-70.

[22] P. Kebarle, U. H. Verkerk, Electrospray: From ions in solution to ions in the gas phase, what we know now, *Mass Spectrom. Rev.* **2009**, *28*, 898-917.

[23] L. Konermann, E. Ahadi, A. D. Rodriguez, S. Vahidi, Unraveling the Mechanism of Electrospray Ionization, *Anal. Chem.* **2013**, *85*, 2-9.

[24] L. Rayleigh, XX. On the equilibrium of liquid conducting masses charged with electricity, *Philos. Mag.* **1882**, *14*, 184-186.

[25] J. V. Iribarne, B. A. Thomson, On the evaporation of small ions from charged droplets, *J. Chem. Phys.* **1976**, *64*, 2287-2294.

[26] C. D. Daub, N. M. Cann, How Are Completely Desolvated Ions Produced in Electrospray Ionization: Insights from Molecular Dynamics Simulations, *Anal. Chem.* **2011**, *83*, 8372-8376.

[27] L. Konermann, A. D. Rodriguez, J. Liu, On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization, *Anal. Chem.* **2012**, *84*, 6798-6804.

[28] M. Wilm, M. Mann, Analytical Properties of the Nanoelectrospray Ion Source, *Anal. Chem.* **1996**, *68*, 1-8.

[29] K. Lorenzen, E. v. Duijn, in *Current Protocols in Protein Science*, John Wiley & Sons, Inc., **2001**.

[30] C. A. Keetch, H. Hernández, A. Sterling, M. Baumert, M. H. Allen, C. V. Robinson, Use of a Microchip Device Coupled with Mass Spectrometry for Ligand Screening of a Multi-Protein Target, *Anal. Chem.* **2003**, *75*, 4937-4941.

[31] A. Konijnenberg, A. Butterer, F. Sobott, Native ion mobility-mass spectrometry and related methods in structural biology, *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 1239-1256.

[32] A. N. Verentchikov, W. Ens, K. G. Standing, Reflecting time-of-flight mass spectrometer with an electrospray ion source and orthogonal extraction, *Anal. Chem.* **1994**, *66*, 126-133.

[33] B. E. Winger, K. J. Light-Wahl, R. R. Ogorzalek Loo, H. R. Udseth, R. D. Smith, Observation and implications of high mass-to-charge ratio ions from electrospray ionization mass spectrometry, *J. Am. Soc. Mass. Spectrom.* **1993**, *4*, 536-545.

[34] B. A. Collings, D. J. Douglas, An extended mass range quadrupole for electrospray mass spectrometry, *Int. J. Mass Spectrom. Ion Process* **1997**, *162*, 121-127.

[35] N. Tahallah, M. Pinkse, C. S. Maier, A. J. R. Heck, The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument, *Rapid Commun. Mass Spectrom.* **2001**, *15*, 596-601.

[36] F. Sobott, H. Hernández, M. G. McCammon, M. A. Tito, C. V. Robinson, A Tandem Mass Spectrometer for Improved Transmission and Analysis of Large Macromolecular Assemblies, *Anal. Chem.* **2002**, *74*, 1402-1407.

[37] R. H. H. van den Heuvel, E. van Duijn, H. Mazon, S. A. Synowsky, K. Lorenzen, C. Versluis, S. J. J. Brouns, D. Langridge, J. van der Oost, J. Hoyes*, et al.*, Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry, *Anal. Chem.* **2006**, *78*, 7473-7483.

[38] J. Snijder, R. J. Rose, D. Veesler, J. E. Johnson, A. J. R. Heck, Studying 18 MDa Virus Assemblies with Native Mass Spectrometry, *Angew. Chem. Int. Ed.* **2013**, *52*, 4020-4023.

[39] A. G. Marshall, C. L. Hendrickson, G. S. Jackson, Fourier transform ion cyclotron resonance mass spectrometry: a primer, *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

[40] M. Wigger, J. R. Eyler, S. A. Benner, W. Li, A. G. Marshall, Fourier transform-ion cyclotron resonance mass spectrometric resolution, identification, and screening of non-covalent complexes of Hck Src homology 2 domain receptor and ligands from a 324-member peptide combinatorial library, *J. Am. Soc. Mass. Spectrom.* **2002**, *13*, 1162-1169.

[41] R. H. Perry, R. G. Cooks, R. J. Noll, Orbitrap mass spectrometry: Instrumentation, ion motion and applications, *Mass Spectrom. Rev.* **2008**, *27*, 661-699.

[42] R. J. Rose, E. Damoc, E. Denisov, A. Makarov, A. J. R. Heck, High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies, *Nat. Methods* **2012**, *9*, 1084-1086.

[43] Y. Yang, F. Liu, V. Franc, L. A. Halim, H. Schellekens, A. J. R. Heck, Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity, *Nat. Commun.* **2016**, *7*, 13397.

[44] K. Breuker, F. W. McLafferty, Stepwise evolution of protein native structure with electrospray into the gas phase, 10−12 to 102 s, *Proc. Natl. Acad. Sci.* **2008**, *105*, 18145-18152.

[45] D. E. Clemmer, R. R. Hudgins, M. F. Jarrold, Naked Protein Conformations: Cytochrome c in the Gas Phase, *J. Am. Chem. Soc.* **1995**, *117*, 10141-10142.

[46] M. Peschke, U. H. Verkerk, P. Kebarle, Prediction of the Charge States of Folded Proteins in Electrospray Ionization, *Eur. J. Mass Spectrom.* **2004**, *10*, 993-1002.

[47] S. K. Chowdhury, V. Katta, B. T. Chait, Probing conformational changes in proteins by mass spectrometry, *J. Am. Chem. Soc.* **1990**, *112*, 9012-9013.

[48] L. Konermann, D. J. Douglas, Equilibrium unfolding of proteins monitored by electrospray ionization mass spectrometry: distinguishing two-state from multi-state transitions, *Rapid Commun. Mass Spectrom.* **1998**, *12*, 435-442.

[49] J. A. Loo, R. R. Loo, H. R. Udseth, C. G. Edmonds, R. D. Smith, Solvent-induced conformational changes of polypeptides probed by electrospray-ionization mass spectrometry, *Rapid Commun. Mass Spectrom.* **1991**, *5*, 101-105.

[50] A. Dobo, I. A. Kaltashov, Detection of Multiple Protein Conformational Ensembles in Solution via Deconvolution of Charge-State Distributions in ESI MS, *Anal. Chem.* **2001**, *73*, 4763-4773.

[51] G. Siuzdak, B. Bothner, M. Yeager, C. Brugidou, C. M. Fauquet, K. Hoey, C. M. Chang, Mass spectrometry and viral analysis, *Chem. Biol.* **1996**, *3*, 45-48.

[52] V. Frankevich, K. Barylyuk, K. Chingin, R. Nieckarz, R. Zenobi, Native biomolecules in the gas phase? The case of green fluorescent protein, *ChemPhysChem* **2013**, *14*, 929-935.

[53] Z. Hall, C. V. Robinson, Do Charge State Signatures Guarantee Protein Conformations?, *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 1161-1168.

[54] B. Gülbakan, K. Barylyuk, R. Zenobi, Determination of thermodynamic and kinetic properties of biomolecules by mass spectrometry, *Curr. Opin. Biotechnol.* **2015**, *31*, 65-72.

[55] K. Breuker, S. Brüschweiler, M. Tollinger, Electrostatic Stabilization of a Native Protein Structure in the Gas Phase, *Angew. Chem. Int. Ed.* **2011**, *50*, 873-877.

[56] J. A. Loo, B. Berhane, C. S. Kaddis, K. M. Wooding, Y. Xie, S. L. Kaufman, I. V. Chernushevich, Electrospray Ionization Mass Spectrometry and Ion Mobility Analysis of the 20S Proteasome Complex, *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 998-1008.

[57] M. Sharon, S. Witt, E. Glasmacher, W. Baumeister, C. V. Robinson, Mass Spectrometry Reveals the Missing Links in the Assembly Pathway of the Bacterial 20 S Proteasome, *J. Biol. Chem.* **2007**, *282*, 18448-18457.

[58] N. P. Barrera, N. Di Bartolo, P. J. Booth, C. V. Robinson, Micelles Protect Membrane Complexes from Solution to Vacuum, *Science* **2008**, *321*, 243-246.

[59] M. Zhou, N. Morgner, N. P. Barrera, A. Politis, S. C. Isaacson, D. Matak-Vinković, T. Murata, R. A. Bernal, D. Stock, C. V. Robinson, Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding, *Science (New York, N.Y.)* **2011**, *334*, 380-385.

[60] R. G. McAllister, H. Metwally, Y. Sun, L. Konermann, Release of Native-like Gaseous Proteins from Electrospray Droplets via the Charged Residue Mechanism: Insights from Molecular Dynamics Simulations, *J. Am. Chem. Soc.* **2015**, *137*, 12667-12676.

[61] Ulrik H. Mistarz, Jeffery M. Brown, Kim F. Haselmann, Kasper D. Rand, Probing the Binding Interfaces of Protein Complexes Using Gas-Phase H/D Exchange Mass Spectrometry, *Structure* **2016**, *24*, 310-318.

[62] C. Arlt, V. Flegler, C. H. Ihling, M. Schäfer, I. Thondorf, A. Sinz, An Integrated Mass Spectrometry Based Approach to Probe the Structure of the Full-Length Wild-Type Tetrameric p53 Tumor Suppressor, *Angew. Chem. Int. Ed.* **2017**, *56*, 275-279.

[63] 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.

[64] K. Pagel, E. Natan, Z. Hall, A. R. Fersht, C. V. Robinson, Intrinsically Disordered p53 and Its Complexes Populate Compact Conformations in the Gas Phase, *Angew. Chem. Int. Ed.* **2013**, *52*, 361-365.

[65] P. Langevin, *Ann. Chim. Phys.* **1903**, *28*, 289-384.

[66] C. Uetrecht, R. J. Rose, E. van Duijn, K. Lorenzen, A. J. R. Heck, Ion mobility mass spectrometry of proteins and protein assemblies, *Chem. Soc. Rev.* **2010**, *39*, 1633-1655.

[67] M. A. Mäkinen, O. A. Anttalainen, M. E. T. Sillanpää, Ion Mobility Spectrometry and Its Applications in Detection of Chemical Warfare Agents, *Anal. Chem.* **2010**, *82*, 9594-9600.

[68] M. Mäkinen, M. Nousiainen, M. Sillanpää, Ion spectrometric detection technologies for ultra-traces of explosives: A review, *Mass Spectrom. Rev.* **2011**, *30*, 940-973.

[69] J. I. Baumbach, Process analysis using ion mobility spectrometry, *Anal. Bioanal. Chem.* **2006**, *384*, 1059-1070.

[70] A. B. Kanu, P. Dwivedi, M. Tam, L. Matz, H. H. Hill, Ion mobility–mass spectrometry, *J. Mass Spectrom.* **2008**, *43*, 1-22.

[71] M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, M. F. Jarrold, Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential, *J. Phys. Chem.* **1996**, *100*, 16082-16086.

[72] E. A. Mason, H. W. Schamp, Mobility of gaseous lons in weak electric fields, *Ann. Phys.* **1958**, *4*, 233-270.

[73] A. A. Shvartsburg, M. F. Jarrold, An exact hard-spheres scattering model for the mobilities of polyatomic ions, *Chem. Phys. Lett.* **1996**, *261*, 86-91.

[74] A. A. Shvartsburg, R. R. Hudgins, P. Dugourd, M. F. Jarrold, Structural Elucidation of Fullerene Dimers by High-Resolution Ion Mobility Measurements and Trajectory Calculation Simulations, *J. Phys. Chem. A* **1997**, *101*, 1684-1688.

[75] C. S. Creaser, J. R. Griffiths, C. J. Bramwell, S. Noreen, C. A. Hill, C. L. P. Thomas, Ion mobility spectrometry: a review. Part 1. Structural analysis by mobility measurement, *Analyst* **2004**, *129*, 984-994.

[76] B. C. Bohrer, S. I. Merenbloom, S. L. Koeniger, A. E. Hilderbrand, D. E. Clemmer, Biomolecule Analysis by Ion Mobility Spectrometry, *Annu. Rev. Anal. Chem.* **2008**, *1*, 293-327.

[77] S. I. Merenbloom, R. S. Glaskin, Z. B. Henson, D. E. Clemmer, High-Resolution Ion Cyclotron Mobility Spectrometry, *Anal. Chem.* **2009**, *81*, 1482-1487.

[78] K. Giles, S. D. Pringle, K. R. Worthington, D. Little, J. L. Wildgoose, R. H. Bateman, Application of a travelling wave-based radio-frequency-only stacked ring ion guide, *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2401-2414.

[79] M. F. Bush, Z. Hall, K. Giles, J. Hoyes, C. V. Robinson, B. T. Ruotolo, Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology, *Anal. Chem.* **2010**, *82*, 9557-9565.

[80] S. D. Pringle, K. Giles, J. L. Wildgoose, J. P. Williams, S. E. Slade, K. Thalassinos, R. H. Bateman, M. T. Bowers, J. H. Scrivens, An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument, *Int. J. Mass Spectrom.* **2007**, *261*, 1-12.

[81] R. W. Purves, R. Guevremont, Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry−Mass Spectrometry, *Anal. Chem.* **1999**, *71*, 2346-2357.

[82] F. Lanucara, S. W. Holman, C. J. Gray, C. E. Eyers, The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics, *Nat. Chem.* **2014**, *6*, 281-294.

[83] F. Fernandez-Lima, D. A. Kaplan, J. Suetering, M. A. Park, Gas-phase separation using a trapped ion mobility spectrometer, *Int. J. Ion Mobil. Spectrom.* **2011**, *14*, 93-98.

[84] E. W. McDaniel, D. W. Martin, W. S. Barnes, Drift Tube‐Mass Spectrometer for Studies of Low‐Energy Ion‐Molecule Reactions, *Rev. Sci. Instrum.* **1962**, *33*, 2-7.

[85] D. E. Clemmer, M. F. Jarrold, Ion Mobility Measurements and their Applications to Clusters and Biomolecules, *J. Mass Spectrom.* **1997**, *32*, 577-592.

[86] B. T. Ruotolo, K. Giles, I. Campuzano, A. M. Sandercock, R. H. Bateman, C. V. Robinson, Evidence for Macromolecular Protein Rings in the Absence of Bulk Water, *Science* **2005**, *310*, 1658-1661.

[87] S. L. Koeniger, S. I. Merenbloom, S. Sevugarajan, D. E. Clemmer, Transfer of Structural Elements from Compact to Extended States in Unsolvated Ubiquitin, *J. Am. Chem. Soc.* **2006**, *128*, 11713-11719.

[88] T. Wyttenbach, M. T. Bowers, Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment, *J. Phys. Chem. B* **2011**, *115*, 12266-12275.

[89] X. Du, Y. Li, Y.-L. Xia, S.-M. Ai, J. Liang, P. Sang, X.-L. Ji, S.-Q. Liu, Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods, *Int. J. Mol. Sci.* **2016**, *17*, 144.

[90] A. Tjernberg, S. Carnö, F. Oliv, K. Benkestock, P.-O. Edlund, W. J. Griffiths, D. Hallén, Determination of Dissociation Constants for Protein−Ligand Complexes by Electrospray Ionization Mass Spectrometry, *Anal. Chem.* **2004**, *76*, 4325-4331.

[91] J. Daniel, G. McCombie, S. Wendt, R. Zenobi, Mass spectrometric determination of association constants of adenylate kinase with two noncovalent inhibitors, *J. Am. Soc. Mass. Spectrom.* **2003**, *14*, 442-448.

[92] S. Zhang, C. K. Van Pelt, D. B. Wilson, Quantitative Determination of Noncovalent Binding Interactions Using Automated Nanoelectrospray Mass Spectrometry, *Anal. Chem.* **2003**, *75*, 3010-3018.

[93] M. C. Jecklin, S. Schauer, C. E. Dumelin, R. Zenobi, Label-free determination of protein–ligand binding constants using mass spectrometry and validation using surface plasmon resonance and isothermal titration calorimetry, *J. Mol. Recogn.* **2009**, *22*, 319-329.

[94] E. Kitova, A. El-Hawiet, P. Schnier, J. Klassen, Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet?, *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431-441.

[95] A. El-Hawiet, E. N. Kitova, J. S. Klassen, Quantifying Protein Interactions with Isomeric Carbohydrate Ligands Using a Catch and Release Electrospray Ionization-Mass Spectrometry Assay, *Anal. Chem.* **2013**, *85*, 7637-7644.

[96] A. El-Hawiet, E. N. Kitova, J. S. Klassen, Quantifying Carbohydrate–Protein Interactions by Electrospray Ionization Mass Spectrometry Analysis, *Biochemistry* **2012**, *51*, 4244-4253.

[97] V. Gabelica, N. Galic, F. Rosu, C. Houssier, E. De Pauw, Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry, *J. Mass Spectrom.* **2003**, *38*, 491-501.

[98] N. Sun, J. Sun, E. N. Kitova, J. S. Klassen, Identifying Nonspecific Ligand Binding in Electrospray Ionization Mass Spectrometry Using the Reporter Molecule Method, *J. Am. Soc. Mass. Spectrom.* **2009**, *20*, 1242-1250.

[99] J. Sun, E. N. Kitova, W. Wang, J. S. Klassen, Method for Distinguishing Specific from Nonspecific Protein−Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry, *Anal. Chem.* **2006**, *78*, 3010-3018.

[100] W. Wang, E. N. Kitova, J. S. Klassen, Nonspecific Protein−Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Factors Influencing Their Formation and Stability, *Anal. Chem.* **2005**, *77*, 3060-3071.

[101] T. Daubenfeld, A.-P. Bouin, G. van der Rest, A Deconvolution Method for the Separation of Specific Versus Nonspecific Interactions in Noncovalent Protein-Ligand Complexes Analyzed by ESI-FT-ICR Mass Spectrometry, *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1239-1248.

[102] L. Shimon, M. Sharon, A. Horovitz, A Method for Removing Effects of Nonspecific Binding on the Distribution of Binding Stoichiometries: Application to Mass Spectroscopy Data, *Biophys. J.* **2010**, *99*, 1645-1649.

[103] S. Guan, M. J. Trnka, D. A. Bushnell, P. J. J. Robinson, J. E. Gestwicki, A. L. Burlingame, Deconvolution Method for Specific and Nonspecific Binding of Ligand to Multiprotein Complex by Native Mass Spectrometry, *Anal. Chem.* **2015**, *87*, 8541-8546.

[104] W. Wang, E. N. Kitova, J. S. Klassen, Influence of Solution and Gas Phase Processes on Protein−Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, *Anal. Chem.* **2003**, *75*, 4945-4955.

[105] M. C. Jecklin, D. Touboul, C. Bovet, A. Wortmann, R. Zenobi, Which Electrospray-Based Ionization Method Best Reflects Protein-Ligand Interactions Found in Solution? A Comparison of ESI, nanoESI, and ESSI for the Determination of Dissociation Constants with Mass Spectrometry, *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 332-343.

[106] K. Barylyuk, B. Gülbakan, X. Xie, R. Zenobi, DNA Oligonucleotides: A Model System with Tunable Binding Strength to Study Monomer–Dimer Equilibria with Electrospray Ionization-Mass Spectrometry, *Anal. Chem.* **2013**, *85*, 11902-11912.

[107] K. A. Sannes-Lowery, R. H. Griffey, S. A. Hofstadler, Measuring Dissociation Constants of RNA and Aminoglycoside Antibiotics by Electrospray Ionization Mass Spectrometry, *Anal. Biochem.* **2000**, *280*, 264-271.

[108] C. V. Robinson, E. W. Chung, B. B. Kragelund, J. Knudsen, R. T. Aplin, F. M. Poulsen, C. M. Dobson, Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein−CoA Ligand Binding and Assembly, *J. Am. Chem. Soc.* **1996**, *118*, 8646-8653.

[109] L. Liu, D. Bagal, E. N. Kitova, P. D. Schnier, J. S. Klassen, Hydrophobic Protein−Ligand Interactions Preserved in the Gas Phase, *J. Am. Chem. Soc.* **2009**, *131*, 15980-15981.

[110] T. D. Veenstra, Electrospray Ionization Mass Spectrometry: A Promising New Technique in the Study of Protein/DNA Noncovalent Complexes, *Biochem. Biophys. Res. Commun.* **1999**, *257*, 1-5.

[111] C. Schmidt, K. Kramer, H. Urlaub, Investigation of protein–RNA interactions by mass spectrometry—Techniques and applications, *J. Proteomics* **2012**, *75*, 3478-3494.

[112] L. Han, E. Kitova, M. Tan, X. Jiang, J. Klassen, Identifying Carbohydrate Ligands of a Norovirus P Particle using a Catch and Release Electrospray Ionization Mass Spectrometry Assay, *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 111-119.

[113] Y. Zhang, L. Liu, R. Daneshfar, E. N. Kitova, C. Li, F. Jia, C. W. Cairo, J. S. Klassen, Protein–Glycosphingolipid Interactions Revealed Using Catch-and-Release Mass Spectrometry, *Anal. Chem.* **2012**, *84*, 7618-7621.

[114] L. Han, E. N. Kitova, J. S. Klassen, Detecting Protein–Glycolipid Interactions Using Glycomicelles and CaR-ESI-MS, *J. Am. Soc. Mass. Spectrom.* **2016**, *27*, 1878-1886.

[115] A. J. R. Heck, R. H. H. van den Heuvel, Investigation of intact protein complexes by mass spectrometry, *Mass Spectrom. Rev.* **2004**, *23*, 368-389.

[116] S. Mehmood, T. M. Allison, C. V. Robinson, Mass Spectrometry of Protein Complexes: From Origins to Applications, *Annu. Rev. Phys. Chem.* **2015**, *66*, 453-474.

[117] L. Pedro, R. Quinn, Native Mass Spectrometry in Fragment-Based Drug Discovery, *Molecules* **2016**, *21*, 984-999.

[118] J. S. Brodbelt, Ion Activation Methods for Peptides and Proteins, *Anal. Chem.* **2016**, *88*, 30-51.

[119] S. A. McLuckey, G. J. Van Berker, G. L. Glish, Tandem mass spectrometry of small, multiply charged oligonucleotides, *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 60-70.

[120] H. Hernandez, C. V. Robinson, Determining the stoichometry and interactions of macromolecular assemblies from mass spectrometry, *Nat. Prot.* **2007**, *2*.

[121] X. Ma, L. B. Lai, S. M. Lai, A. Tanimoto, M. P. Foster, V. H. Wysocki, V. Gopalan, Uncovering the Stoichiometry of Pyrococcus furiosus RNase P, a Multi-Subunit Catalytic Ribonucleoprotein Complex, by Surface-Induced Dissociation and Ion Mobility Mass Spectrometry, *Angew. Chem. Int. Ed.* **2014**, *53*, 11483-11487.

[122] K. Breuker, F. W. McLafferty, Native Electron Capture Dissociation for the Structural Characterization of Noncovalent Interactions in Native Cytochrome c, *Angew. Chem. Int. Ed.* **2003**, *42*, 4900-4904.

[123] F. Lermyte, F. Sobott, Electron transfer dissociation provides higher-order structural information of native and partially unfolded protein complexes, *Proteomics* **2015**, *15*, 2813-2822.

[124] J. S. Brodbelt, J. J. Wilson, Infrared multiphoton dissociation in quadrupole ion traps, *Mass Spectrom. Rev.* **2009**, *28*, 390-424.

[125] J. P. O'Brien, W. Li, Y. Zhang, J. S. Brodbelt, Characterization of Native Protein Complexes Using Ultraviolet Photodissociation Mass Spectrometry, *J. Am. Chem. Soc.* **2014**, *136*, 12920-12928.

[126] T. Wyttenbach, M. Grabenauer, K. Thalassinos, J. H. Scrivens, M. T. Bowers, The Effect of Calcium Ions and Peptide Ligands on the Relative Stabilities of the Calmodulin Dumbbell and Compact Structures, *J. Phys. Chem. B* **2010**, *114*, 437-447.

[127] S. Warnke, G. von Helden, K. Pagel, Protein Structure in the Gas Phase: The Influence of Side-Chain Microsolvation, *J. Am. Chem. Soc.* **2013**, *135*, 1177-1180.

[128] E. van Duijn, I. M. Barbu, A. Barendregt, M. M. Jore, B. Wiedenheft, M. Lundgren, E. R. Westra, S. J. J. Brouns, J. A. Doudna, J. van der Oost*, et al.*, Native Tandem and Ion Mobility Mass Spectrometry Highlight Structural and Modular Similarities in Clustered-Regularly-Interspaced Shot-Palindromic-Repeats (CRISPR)-associated Protein Complexes From Escherichia coli and Pseudomonas aeruginosa, *Mol. Cell. Proteomics* **2012**, *11*, 1430-1441.

[129] A. Laganowsky, E. Reading, T. M. Allison, M. B. Ulmschneider, M. T. Degiacomi, A. J. Baldwin, C. V. Robinson, Membrane proteins bind lipids selectively to modulate their structure and function, *Nature* **2014**, *510*, 172-175.

[130] E. Reading, I. Liko, T. M. Allison, J. L. P. Benesch, A. Laganowsky, C. V. Robinson, The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase, *Angew. Chem. Int. Ed.* **2015**, *54*, 4577-4581.

[131] L. M. Young, J. C. Saunders, R. A. Mahood, C. H. Revill, R. J. Foster, L.-H. Tu, D. P. Raleigh, S. E. Radford, A. E. Ashcroft, Screening and classifying small-molecule inhibitors of amyloid formation using ion mobility spectrometry–mass spectrometry, *Nat. Chem.* **2015**, *7*, 73-81.

[132] X. Cheng, R. Chen, J. E. Bruce, B. L. Schwartz, G. A. Anderson, S. A. Hofstadler, D. C. Gale, R. D. Smith, J. Gao, G. B. Sigal*, et al.*, Using Electrospray Ionization FTICR Mass Spectrometry To Study Competitive Binding of Inhibitors to Carbonic Anhydrase, *J. Am. Chem. Soc.* **1995**, *117*, 8859-8860.

[133] J. Gao, X. Cheng, R. Chen, G. B. Sigal, J. E. Bruce, B. L. Schwartz, S. A. Hofstadler, G. A. Anderson, R. D. Smith, G. M. Whitesides, Screening Derivatized Peptide Libraries for Tight Binding Inhibitors to Carbonic Anhydrase II by Electrospray Ionization-Mass Spectrometry, *J. Med. Chem.* **1996**, *39*, 1949-1955.

[134] F. H. Cederkvist, A. D. Zamfir, S. Bahrke, V. G. H. Eijsink, M. Sørlie, J. Peter-Katalinić, M. G. Peter, Identification of a High-Affinity-Binding Oligosaccharide by (+) Nanoelectrospray Quadrupole Time-of-Flight Tandem Mass Spectrometry of a Noncovalent Enzyme–Ligand Complex, *Angew. Chem. Int. Ed.* **2006**, *45*, 2429-2434.

[135] A. El-Hawiet, G. K. Shoemaker, R. Daneshfar, E. N. Kitova, J. S. Klassen, Applications of a Catch and Release Electrospray Ionization Mass Spectrometry Assay for Carbohydrate Library Screening, *Anal. Chem.* **2012**, *84*, 50-58.

[136] Y. Zhong, L. Han, B. T. Ruotolo, Collisional and Coulombic Unfolding of Gas-Phase Proteins: High Correlation to Their Domain Structures in Solution, *Angew. Chem. Int. Ed.* **2014**, *53*, 9209-9212.

[137] H. Zhang, H. Liu, Y. Lu, N. R. Wolf, M. L. Gross, R. E. Blankenship, Native mass spectrometry and ion mobility characterize the orange carotenoid protein functional domains, *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2016**, *1857*, 734-739.

[138] J. T. S. Hopper, N. J. Oldham, Collision Induced Unfolding of Protein Ions in the Gas Phase Studied by Ion Mobility-Mass Spectrometry: The Effect of Ligand Binding on Conformational Stability, *J. Am. Soc. Mass. Spectrom.* **2009**, *20*, 1851-1858.

[139] S.-J. Hyung, C. V. Robinson, B. T. Ruotolo, Gas-Phase Unfolding and Disassembly Reveals Stability Differences in Ligand-Bound Multiprotein Complexes, *Chem. Biol.* **2009**, *16*, 382-390.

[140] T. M. Allison, E. Reading, I. Liko, A. J. Baldwin, A. Laganowsky, C. V. Robinson, Quantifying the stabilizing effects of protein-ligand interactions in the gas phase, *Nat. Commun.* **2015**, *6*, 8551-8551.

[141] J. N. Rabuck, S.-J. Hyung, K. S. Ko, C. C. Fox, M. B. Soellner, B. T. Ruotolo, Activation State-Selective Kinase Inhibitor Assay Based on Ion Mobility-Mass Spectrometry, *Anal. Chem.* **2013**, *85*, 6995-7002.

[142] J. D. Eschweiler, R. M. Martini, B. T. Ruotolo, Chemical Probes and Engineered Constructs Reveal a Detailed Unfolding Mechanism for a Solvent-Free Multidomain Protein, *J. Am. Chem. Soc.* **2017**, *139*, 534-540.

[143] K. Giles, J. P. Williams, I. Campuzano, Enhancements in travelling wave ion mobility resolution, *Rapid Commun. Mass Spectrom.* **2011**, *25*, 1559-1566.

[144] S. J. Allen, M. F. Bush, Radio-Frequency (rf) Confinement in Ion Mobility Spectrometry: Apparent Mobilities and Effective Temperatures, *J. Am. Soc. Mass. Spectrom.* **2016**, 1-10.

[145] A. Tjernberg, N. Markova, W. J. Griffiths, D. Hallén, DMSO-Related Effects in Protein Characterization, *J. Biomol. Screen.* **2006**, *11*, 131-137.

[146] D. Cubrilovic, R. Zenobi, Influence of Dimehylsulfoxide on Protein–Ligand Binding Affinities, *Anal. Chem.* **2013**, *85*, 2724-2730.

[147] H. J. Sterling, J. S. Prell, C. A. Cassou, E. R. Williams, Protein Conformation and Supercharging with DMSO from Aqueous Solution, *J. Am. Soc. Mass. Spectrom.* **2011**, *22*, 1178-1186.

[148] H. Sterling, A. Kintzer, G. Feld, C. Cassou, B. Krantz, E. Williams, Supercharging Protein Complexes from Aqueous Solution Disrupts their Native Conformations, *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 191-200.

[149] M. Landreh, G. Alvelius, J. Johansson, H. Jörnvall, Protective Effects of Dimethyl Sulfoxide on Labile Protein Interactions during Electrospray Ionization, *Anal. Chem.* **2014**, *86*, 4135-4139.

[150] P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar*, et al.*, Selective inhibition of BET bromodomains, *Nature* **2010**, *468*, 1067-1073.

[151] G. Lolli, R. Battistutta, Different orientations of low-molecular-weight fragments in the binding pocket of a BRD4 bromodomain, *Acta Crytallogr. D* **2013**, *69*, 2161-2164.

[152] C. Dhalluin, J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal, M.-M. Zhou, M.-M. Zhou, Structure and ligand of a histone acetyltransferase bromodomain, *Nature* **1999**, *399*, 491-496.

[153] C. A. French, I. Miyoshi, J. C. Aster, I. Kubonishi, T. G. Kroll, P. Dal Cin, S. O. Vargas, A. R. Perez-Atayde, J. A. Fletcher, BRD4 Bromodomain Gene Rearrangement in Aggressive Carcinoma with Translocation t(15;19), *Am. J. Pathol.* **2001**, *159*, 1987-1992.

[154] C. A. French, I. Miyoshi, I. Kubonishi, H. E. Grier, A. R. Perez-Atayde, J. A. Fletcher, BRD4-NUT Fusion Oncogene: A Novel Mechanism in Aggressive Carcinoma, *Cancer Res.* **2003**, *63*, 304-307.

[155] J. Lovén, Heather A. Hoke, Charles Y. Lin, A. Lau, David A. Orlando, Christopher R. Vakoc, James E. Bradner, Tong I. Lee, Richard A. Young, Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers, *Cell* **2013**, *153*, 320-334.

[156] H. J. Maple, R. A. Garlish, L. Rigau-Roca, J. Porter, I. Whitcombe, C. E. Prosser, J. Kennedy, A. J. Henry, R. J. Taylor, M. P. Crump, *et al.*, Automated Protein–Ligand Interaction Screening by Mass Spectrometry, *J. Med. Chem.* **2011**, *55*, 837-851.

[157] H. Sterling, M. Daly, G. Feld, K. Thoren, A. Kintzer, B. Krantz, E. Williams, Effects of supercharging reagents on noncovalent complex structure in electrospray ionization from aqueous solutions, *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 1762-1774.

[158] Y. Yao, M. R. Richards, E. N. Kitova, J. S. Klassen, Influence of Sulfolane on ESI-MS Measurements of Protein–Ligand Affinities, *J. Am. Soc. Mass. Spectrom.* **2016**, *27*, 498-506.

[159] D. S.-H. Chan, D. Matak-Vinković, A. G. Coyne, C. Abell, Insight into Protein Conformation and Subcharging by DMSO from Native Ion Mobility Mass Spectrometry, *ChemistrySelect* **2016**, *1*, 5686-5690.

[160] A. W. Adamson, A. P. Gast, *Physical chemistry of surfaces*, 6th ed., Wiley, **1997**.

[161] D. R. Lide, *CRC Handbook of Chemistry and Physics*, 89th ed., CRC Press, Taylor & Francis Group, **2008 - 2009**.

[162] H. Metwally, R. G. McAllister, V. Popa, L. Konermann, Mechanism of Protein Supercharging by Sulfolane and m-Nitrobenzyl Alcohol: Molecular Dynamics Simulations of the Electrospray Process, *Anal. Chem.* **2016**, *88*, 5345-5354.

[163] M. Whittaker, R. J. Law, O. Ichihara, T. Hesterkamp, D. Hallett, Fragments: past, present and future, *Drug Discov. Today: Technol.* **2010**, *7*, e163-e171.

[164] D. E. Scott, A. G. Coyne, S. A. Hudson, C. Abell, Fragment-Based Approaches in Drug Discovery and Chemical Biology, *Biochemistry* **2012**, *51*, 4990-5003.

[165] C. W. Murray, D. C. Rees, The rise of fragment-based drug discovery, *Nat. Chem.* **2009**, *1*, 187-192.

[166] M. M. Hann, A. R. Leach, G. Harper, Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856-864.

[167] A. W. Hung, A. Ramek, Y. Wang, T. Kaya, J. A. Wilson, P. A. Clemons, D. W. Young, Route to three-dimensional fragments using diversity-oriented synthesis, *Proc. Natl. Acad. Sci.* **2011**, *108*, 6799-6804.

[168] H. J. Maple, R. A. Garlish, L. Rigau-Roca, J. Porter, I. Whitcombe, C. E. Prosser, J. Kennedy, A. J. Henry, R. J. Taylor, M. P. Crump, *et al.*, Automated Protein–Ligand Interaction Screening by Mass Spectrometry, *J. Med. Chem.* **2012**, *55*, 837-851.

[169] C. Dalvit, NMR methods in fragment screening: theory and a comparison with other biophysical techniques, *Drug Discov. Today* **2009**, *14*, 1051-1057.

[170] H. Jhoti, A. Cleasby, M. Verdonk, G. Williams, Fragment-based screening using X-ray crystallography and NMR spectroscopy, *Curr. Opin. Chem. Biol.* **2007**, *11*, 485-493.

[171] M. Geitmann, M. Elinder, C. Seeger, P. Brandt, I. J. P. De Esch, U. H. Danielson, Identification of a Novel Scaffold for Allosteric Inhibition of Wild Type and Drug Resistant HIV-1 Reverse Transkriptase by Fragment Library Screening, *J. Med. Chem.* **2011**, *54*, 699-708.

[172] M. W. Pantoliano, E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B. A. Springer, P. Lane, *et al.*, High-Density Miniaturized Thermal Shift Assays as a General Strategy for Drug Discovery, *J. Biomol. Screen.* **2001**, *6*, 429-440.

[173] G. M. Keserű, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray, S. D. Pickett, Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia, *J. Med. Chem.* **2016**, *59*, 8189-8206.

[174] A. Ahmed-Belkacem, L. Colliandre, N. Ahnou, Q. Nevers, M. Gelin, Y. Bessin, R. Brillet, O. Cala, D. Douguet, W. Bourguet, *et al.*, Fragment-based discovery of a new family of non-peptidic small-molecule cyclophilin inhibitors with potent antiviral activities, *Nat. Commun.* **2016**, *7*, 1-11.

[175] P. M. Collins, J. T. Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N. Wright, N. M. Pearce, F. von Delft, Gentle, fast and effective crystal soaking by acoustic dispensing, *Acta Cryst. D. Biol. Struct.* **2017**, *73*, 246-255.

[176] V. Borsi, V. Calderone, M. Fragai, C. Luchinat, N. Sarti, Entropic Contribution to the Linking Coefficient in Fragment Based Drug Design: A Case Study, *J. Med. Chem.* **2010**, *53*, 4285-4289.

[177] A. Friberg, D. Vigil, B. Zhao, R. N. Daniels, J. P. Burke, P. M. Garcia-Barrantes, D. Camper, B. A. Chauder, T. Lee, E. T. Olejniczak, *et al.*, Discovery of potent myeloid cell leukemia 1 (Mcl 1) inhibitors using fragment based methods and structure based design, *J. Med. Chem.* **2013**, *56*, 15-30.

[178] G. Deng, G. Sanyal, Applications of mass spectrometry in early stages of target based drug discovery, *J. Pharm. Biomed. Anal.* **2006**, *40*, 528-538.

[179] S. A. Hofstadler, K. A. Sannes-Lowery, Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes, *Nat. Rev. Drug Discov.* **2006**, *5*, 585-595.

[180] V. Vivat Hannah, C. Atmanene, D. Zeyer, A. Van Dorsselaer, S. Sanglier-Cianférani, Native MS: an 'ESI, way to support structure- and fragment-based drug discovery, *Fut. Med. Chem.* **2009**, *2*, 35-50.

[181] N. Drinkwater, H. Vu, K. M. Lovell, K. R. Criscione, B. M. Collins, T. E. Prisinzano, S. A. Poulsen, M. J. McLeish, G. L. Grunewald, J. L. Martin, Fragment-based screening by X-ray crystallography, MS and isothermal titration calorimetry to identify PNMT (phenylethanolamine N-methyltransferase) inhibitors, *Biochem. J* **2010**, *431*, 51-61.

[182] S.-A. Poulsen, Fragment Screening by Native State Mass Spectrometry, *Austr. J. Chem.* **2013**, *66*, 1495-1501.

[183] F. Riccardi Sirtori, D. Caronni, M. Colombo, C. Dalvit, M. Paolucci, L. Regazzoni, C. Visco, G. Fogliatto, Establish an automated flow injection ESI-MS method for the screening of fragment based libraries: Application to Hsp90, *Eur. J. Pharm. Sci.* **2015**, *76*, 83-94.

[184] X. Chen, S. Qin, S. Chen, J. Li, L. Li, Z. Wang, Q. Wang, J. Lin, C. Yang, W. Shui, A Ligand-observed Mass Spectrometry Approach Integrated into the Fragment Based Lead Discovery Pipeline, *Sci. Rep.* **2015**, *5*, 1-8.

[185] L. A. Woods, O. Dolezal, B. Ren, J. H. Ryan, T. S. Peat, S.-A. Poulsen, Native State Mass Spectrometry, Surface Plasmon Resonance, and X-ray Crystallography Correlate Strongly as a Fragment Screening Combination, *J. Med. Chem.* **2016**, *59*, 2192-2204.

[186] J. Schiebel, N. Radeva, H. Köster, A. Metz, T. Krotzky, M. Kuhnert, W. E. Diederich, A. Heine, L. Neumann, C. Atmanene*, et al.*, One Question, Multiple Answers: Biochemical and Biophysical Screening Methods Retrieve Deviating Fragment Hit Lists, *ChemMedChem* **2015**, *10*, 1511-1521.

[187] M.-C. Lo, A. Aulabaugh, G. Jin, R. Cowling, J. Bard, M. Malamas, G. Ellestad, Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery, *Anal. Biochem.* **2004**, *332*, 153-159.

[188] J. K. Kranz, C. Schalk-Hihi, in *Methods Enzymol., Vol. Volume 493* (Ed.: C. K. Lawrence), Academic Press, **2011**, pp. 277-298.

[189] D. Matulis, J. K. Kranz, F. R. Salemme, M. J. Todd, Thermodynamic Stability of Carbonic Anhydrase: Measurements of Binding Affinity and Stoichiometry Using ThermoFluor, *Biochemistry* **2005**, *44*, 5258-5266.

[190] G. A. Senisterra, E. Markin, K. Yamazaki, R. Hui, M. Vedadi, D. E. Awrey, Screening for Ligands Using a Generic and High-Throughput Light-Scattering-Based Assay, *J. Biomol. Screen.* **2006**, *11*, 940-948.

[191] M. Vedadi, F. H. Niesen, A. Allali-Hassani, O. Y. Fedorov, P. J. Finerty, G. A. Wasney, R. Yeung, C. Arrowsmith, L. J. Ball, H. Berglund*, et al.*, Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15835-15840.

[192] Z. Wang, M. Cui, F. Song, L. Lu, Z. Liu, S. Liu, Evaluation of Flavonoids Binding to DNA Duplexes by Electrospray Ionization Mass Spectrometry, *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 914-922.

[193] F. Rosu, V. Gabelica, C. Houssier, E. De Pauw, Determination of affinity, stoichiometry and sequence selectivity of minor groove binder complexes with double-stranded oligodeoxynucleotides by electrospray ionization mass spectrometry, *Nucleic Acids Res.* **2002**, *30*, e82-e82.

[194] H. Gad, T. Koolmeister, A.-S. Jemth, S. Eshtad, S. A. Jacques, C. E. Strom, L. M. Svensson, N. Schultz, T. Lundback, B. O. Einarsdottir*, et al.*, MTH1 inhibition eradicates cancer by preventing sanitation of the dNTP pool, *Nature* **2014**, *508*, 215-221.

[195] A. Patriksson, E. Marklund, D. van der Spoel, Protein Structures under Electrospray Conditions, *Biochemistry* **2007**, *46*, 933-945.

[196] C. Van Pelt, S. Zhang, J. Henion, Characterization of a fully automated nanoelectrospray system with mass spectrometric detection for proteomic analyses, *J. Biomol. Tech.* **2002**, *13*, 72-84.

[197] E. N. Kitova, A. El-Hawiet, J. S. Klassen, Screening Carbohydrate Libraries for Protein Interactions Using the Direct ESI-MS Assay. Applications to Libraries of Unknown Concentration, *J. Am. Soc. Mass. Spectrom.* **2014**, *25*, 1908-1916.

[198] E. A. Blackburn, J. K. F. Maclean, B. S. Sherborne, M. D. Walkinshaw, Estimating the affinity of protein–ligand complex from changes to the charge-state distribution of a protein in electrospray ionization mass spectrometry, *Biochem. Biophys. Res. Commun.* **2010**, *403*, 190-193.

[199] L. H. Kristensen, A. L. Nielsen, C. Helgstrand, M. Lees, P. Cloos, J. S. Kastrup, K. Helin, L. Olsen, M. Gajhede, Studies of H3K4me3 demethylation by KDM5B/Jarid1B/PLU1 reveals strong substrate recognition in vitro and identifies 2,4-pyridine-dicarboxylic acid as an in vitro and in cell inhibitor, *FEBS Journal* **2012**, *279*, 1905-1914.

[200] N. Avvakumov, J. Cote, The MYST family of histone acetyltransferases and their intimate links to cancer, *Oncogene* **2000**, *26*, 5395-5407.

[201] S. B. Rothbart, B. M. Dickson, M. S. Ong, K. Krajewski, S. Houliston, D. B. Kireev, C. H. Arrowsmith, B. D. Strahl, Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation, *Genes Dev.* **2013**, *27*, 1288-1298.

[202] L. Cui, J. Chen, Q. Zhang, X. Wang, J. Qu, J. Zhang, S. Dang, Up-regulation of UHRF1 by oncogenic Ras promoted the growth, migration, and metastasis of pancreatic cancer cells, *Mol. Cell. Biochem.* **2015**, *400*, 223-232.

[203] C. Johansson, S. Velupillai, A. Tumber, A. Szykowska, E. S. Hookway, R. P. Nowak, C. Strain-Damerell, C. Gileadi, M. Philpott, N. Burgess-Brown, *et al.*, Structural analysis of human KDM5B guides histone demethylase inhibitor development, *Nat. Chem. Biol.* **2016**, *12*, 539-545.

[204] S. Yin, Y. Xie, J. A. Loo, Mass Spectrometry of Protein-Ligand Complexes: Enhanced Gas Phase Stability of Ribonuclease-Nucleotide Complexes, *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1199-1208.

[205] J. T. S. Hopper, C. V. Robinson, Mass Spectrometry Quantifies Protein Interactions—From Molecular Chaperones to Membrane Porins, *Angew. Chem. Int. Ed.* **2014**, *53*, 14002-14015.

[206] S. Myung, E. R. Badman, Y. J. Lee, D. E. Clemmer, Structural Transitions of Electrosprayed Ubiquitin Ions Stored in an Ion Trap over ~10 ms to 30 s, *J. Phys. Chem. A* **2002**, *106*, 9976-9982.

[207] E. Segev, T. Wyttenbach, M. T. Bowers, R. B. Gerber, Conformational evolution of ubiquitin ions in electrospray mass spectrometry: molecular dynamics simulations at gradually increasing temperatures, *Phys. Chem. Chem. Phys.* **2008**, *10*, 3077-3082.

[208] Y. Mao, M. A. Ratner, M. F. Jarrold, Molecular Dynamics Simulations of the Charge-Induced Unfolding and Refolding of Unsolvated Cytochrome c, *J. Phys. Chem. B* **1999**, *103*, 10017-10021.

[209] H. Shi, N. A. Pierson, S. J. Valentine, D. E. Clemmer, Conformation Types of Ubiquitin [M+8H]8+ Ions from Water:Methanol Solutions: Evidence for the N and A States in Aqueous Solution, *J. Phys. Chem. B* **2012**, *116*, 3344-3352.

[210] A. I. González Flórez, E. Mucha, D.-S. Ahn, S. Gewinner, W. Schöllkopf, K. Pagel, G. von Helden, Charge-Induced Unzipping of Isolated Proteins to a Defined Secondary Structure, *Angew. Chem. Int. Ed.* **2016**, *55*, 3295-3299.

[211] T. Wyttenbach, M. T. Bowers, in *Modern Mass Spectrometry* (Ed.: C. A. Schalley), Springer Berlin Heidelberg, Berlin, Heidelberg, **2003**, pp. 207-232.

[212] C. Bleiholder, N. F. Dupuis, T. Wyttenbach, M. T. Bowers, Ion mobility–mass spectrometry reveals a conformational conversion from random assembly to β-sheet in amyloid fibril formation, *Nat. Chem.* **2011**, *3*, 172-177.

[213] K. Breuker, H. Oh, D. M. Horn, B. A. Cerda, F. W. McLafferty, Detailed Unfolding and Folding of Gaseous Ubiquitin Ions Characterized by Electron Capture Dissociation, *J. Am. Chem. Soc.* **2002**, *124*, 6407-6420.

[214] S. L. Koeniger, S. I. Merenbloom, D. E. Clemmer, Evidence for Many Resolvable Structures within Conformation Types of Electrosprayed Ubiquitin Ions, *J. Phys. Chem. B* **2006**, *110*, 7017-7021.

[215] S. L. Koeniger, D. E. Clemmer, Resolution and structural transitions of elongated states of ubiquitin, *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 322-331.

[216] O. S. Skinner, F. W. McLafferty, K. Breuker, How Ubiquitin Unfolds after Transfer into the Gas Phase, *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 1011-1014.

[217] J. W. Lee, S. W. Heo, S. J. C. Lee, J. Y. Ko, H. Kim, H. I. Kim, Probing Conformational Changes of Ubiquitin by Host–Guest Chemistry Using Electrospray Ionization Mass Spectrometry, *J. Am. Soc. Mass. Spectrom.* **2013**, *24*, 21-29.

[218] S. Warnke, C. Baldauf, M. T. Bowers, K. Pagel, G. von Helden, Photodissociation of Conformer-Selected Ubiquitin Ions Reveals Site-Specific Cis/Trans Isomerization of Proline Peptide Bonds, *J. Am. Chem. Soc.* **2014**, *136*, 10308-10314.

[219] K. A. Servage, J. A. Silveira, K. L. Fort, D. E. Clemmer, D. H. Russell, Water-Mediated Dimerization of Ubiquitin Ions Captured by Cryogenic Ion Mobility-Mass Spectrometry, *J. Phys. Chem. Lett.* **2015**, *6*, 4947-4951.

[220] A. Hershko, A. Ciechanover, The Ubiquitin System, *Annu. Rev. Biochem.* **1998**, *67*, 425-479.

[221] S. Maleknia, J. Brodbelt, Cavity-size-dependent dissociation of crown ether/ammonium ion complexes in the gas phase, *J. Am. Chem. Soc.* **1993**, *115*, 2837-2843.

[222] C. C. Liou, H. F. Wu, J. S. Brodbelt, Hydrogen-bonding interactions in gas-phase polyether/ammonium ion complexes, *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 260-273.

[223] R. R. Julian, J. L. Beauchamp, Site specific sequestering and stabilization of charge in peptides by supramolecular adduct formation with 18-crown-6 ether by way of electrospray ionization1, *Int. J. Mass Spectrom.* **2001**, *210/211*, 613-623.

[224] P. R. Kemper, N. F. Dupuis, M. T. Bowers, A new, higher resolution, ion mobility mass spectrometer, *Int. J. Mass Spectrom.* **2009**, *287*, 46-57.

[225] Z. Liu, S. Cheng, D. R. Gallie, R. R. Julian, Exploring the Mechanism of Selective Noncovalent Adduct Protein Probing Mass Spectrometry Utilizing Site-Directed Mutagenesis To Examine Ubiquitin, *Anal. Chem.* **2008**, *80*, 3846-3852.

[226] D. P. Weimann, H. D. F. Winkler, J. A. Falenski, B. Koksch, C. A. Schalley, Highly dynamic motion of crown ethers along oligolysine peptide chains, *Nat. Chem.* **2009**, *1*, 573-577.

[227] H. D. F. Winkler, D. P. Weimann, A. Springer, C. A. Schalley, Dynamic Motion in Crown Ether Dendrimer Complexes: A "Spacewalk" on the Molecular Scale, *Angew. Chem. Int. Ed.* **2009**, *48*, 7246-7250.

[228] S. N. Jackson, S. Dutta, A. S. Woods, The use of ECD/ETD to identify the site of electrostatic interaction in noncovalent complexes, *J. Am. Soc. Mass. Spectrom.* **2011**, *20*, 176-179.

[229] L. Muller, S. N. Jackson, A. S. Woods, ETD and sequential ETD localize the residues involved in D2-A2A heteromerization, *RSC Advances* **2014**, *4*, 42272-42277.

[230] D. M. Good, M. Wirtala, G. C. McAlister, J. J. Coon, Performance Characteristics of Electron Transfer Dissociation Mass Spectrometry, *Mol. Cell Proteomics* **2007**, *6*, 1942-1951.

[231] S. A. McLuckey, J. L. Stephenson, Ion/ion chemistry of high-mass multiply charged ions, *Mass Spectrom. Rev.* **1998**, *17*, 369-407.

[232] S. J. Pitteri, P. A. Chrisman, J. M. Hogan, S. A. McLuckey, Electron Transfer Ion/Ion Reactions in a Three-Dimensional Quadrupole Ion Trap: Reactions of Doubly and Triply Protonated Peptides with SO2•-, *Anal. Chem.* **2005**, *77*, 1831-1839.

[233] S. J. Pitteri, S. A. McLuckey, Recent developments in the ion/ion chemistry of high-mass multiply charged ions, *Mass Spectrom. Rev.* **2005**, *24*, 931-958.

[234] F. Lermyte, T. Verschueren, J. M. Brown, J. P. Williams, D. Valkenborg, F. Sobott, Characterization of top-down ETD in a travelling-wave ion guide, *Methods* **2015**, *89*, 22-29.

[235] T. Ly, R. R. Julian, Using ESI-MS to Probe Protein Structure by Site-Specific Noncovalent Attachment of 18-Crown-6, *J. Am. Soc. Mass. Spectrom.* **2006**, *17*, 1209-1215.

[236] T. Ly, R. R. Julian, Protein–Metal Interactions of Calmodulin and α-Synuclein Monitored by Selective Noncovalent Adduct Protein Probing Mass Spectrometry, *J. Am. Soc. Mass. Spectrom.* **2008**, *19*, 1663-1672.

[237] B. N. Moore, O. Hamdy, R. R. Julian, Protein structure evolution in liquid DESI as revealed by selective noncovalent adduct protein probing, *Int. J. Mass Spectrom.* **2012**, *330–332* 220-225.

[238] R. D. Burgoyne, Neuronal calcium sensor proteins: generating diversity in neuronal Ca2+ signalling, *Nat. Rev. Neurosci.* **2007**, *8*, 182-193.

[239] M. R. Kreutz, J. R. Naranjo, K. W. Koch, B. Schwaller, The Neuronal Functions of EF-Hand Ca2+-Binding Proteins, *Front. Mol. Neurosci.* **2012**, *5*, 92.

[240] M. Mikhaylova, J. Hradsky, M. R. Kreutz, Between promiscuity and specificity: novel roles of EF-hand calcium sensors in neuronal Ca2+ signalling, *J. Neurochem.* **2011**, *118*, 695-713.

[241] T. N. Davis, J. Thorner, Vertebrate and yeast calmodulin, despite significant sequence divergence, are functionally interchangeable, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7909-7913.

[242] A. Biber, G. Schmid, K. Hempel, Calmodulin content in specific brain areas, *Exp. Brain Res.* **1984**, *56*, 323-326.

[243] M. W. Berchtold, A. Villalobo, The many faces of calmodulin in cell proliferation, programmed cell death, autophagy, and cancer, *Biochim. Biophys. Acta* **2014**, *1843*, 398-435.

[244] J. J. Chou, S. Li, C. B. Klee, A. Bax, Solution structure of Ca2+-calmodulin reveals flexible hand-like properties of its domains, *Nat. Struct. Mol. Biol.* **2001**, *8*, 990-997.

[245] D. Chin, A. R. Means, Calmodulin: a prototypical calcium sensor, *Trends Cell Biol.* **2000**, *10*, 322-328.

[246] H. Tidow, P. Nissen, Structural diversity of calmodulin binding to its target sites, *FEBS Journal* **2013**, *280*, 5551-5565.

[247] P. Kursula, The many structural faces of calmodulin: a multitasking molecular jackknife, *Amino acids* **2014**, *46*, 2295-2304.

[248] A. Villarroel, M. Taglialatela, G. Bernardo-Seisdedos, A. Alaimo, J. Agirre, A. Alberdi, C. Gomis-Perez, M. V. Soldovieri, P. Ambrosino, C. Malo*, et al.*, The ever changing moods of calmodulin: how structural plasticity entails transductional adaptability, *J. Mol. Biol.* **2014**, *426*, 2717-2735.

[249] P. Radivojac, S. Vucetic, T. R. O'Connor, V. N. Uversky, Z. Obradovic, A. K. Dunker, Calmodulin signaling: analysis and prediction of a disorder-dependent molecular recognition, *Proteins* **2006**, *63*, 398-410.

[250] K. L. Yap, J. Kim, K. Truong, M. Sherman, T. Yuan, M. Ikura, Calmodulin target database, *J. Struct. Funct. Genomics* **2000**, *1*, 8-14.

[251] C. Kuo-Chen, Pseudo Amino Acid Composition and its Applications in Bioinformatics, Proteomics and System Biology, *Curr. Proteomics* **2009**, *6*, 262-274.

[252] B. Liu, F. Liu, X. Wang, J. Chen, L. Fang, K.-C. Chou, Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences, *Nucleic Acids Res.* **2015**, *43*, W65-W71.

[253] M. Ikura, G. Clore, A. Gronenborn, G. Zhu, C. Klee, A. Bax, Solution structure of a calmodulin-target peptide complex by multidimensional NMR, *Science* **1992**, *256*, 632-638.

[254] W. E. Meador, A. R. Means, F. A. Quiocho, Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex, *Science* **1992**, *257*, 1251-1255.

[255] W. E. Meador, A. R. Means, F. A. Quiocho, Modulation of calmodulin plasticity in molecular recognition on the basis of x-ray structures, *Science* **1993**, *262*, 1718-1721.

[256] H. J. Junge, J.-S. Rhee, O. Jahn, F. Varoqueaux, J. Spiess, M. N. Waxham, C. Rosenmund, N. Brose, Calmodulin and Munc13 Form a Ca2+ Sensor/Effector Complex that Controls Short-Term Synaptic Plasticity, *Cell* **2004**, *118*, 389-401.

[257] F. Varoqueaux, A. Sigler, J. S. Rhee, N. Brose, C. Enk, K. Reim, C. Rosenmund, Total arrest of spontaneous and evoked synaptic transmission but normal synaptogenesis in the absence of Munc13-mediated vesicle priming, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9037-9042.

[258] K. Dimova, S. Kalkhof, I. Pottratz, C. Ihling, F. Rodriguez-Castaneda, T. Liepold, C. Griesinger, N. Brose, A. Sinz, O. Jahn, Structural Insights into the Calmodulin−Munc13 Interaction Obtained by Cross-Linking and Mass Spectrometry, *Biochemistry* **2009**, *48*, 5908-5921.

[259] F. Rodriguez-Castaneda, M. Maestre-Martinez, N. Coudevylle, K. Dimova, H. Junge, N. Lipstein, D. Lee, S. Becker, N. Brose, O. Jahn*, et al.*, Modular architecture of Munc13/calmodulin complexes: dual regulation by Ca2+ and possible function in short-term synaptic plasticity, *EMBO J.* **2010**, *29*, 680-691.

[260] S. J. Allen, K. Giles, T. Gilbert, M. F. Bush, Ion mobility mass spectrometry of peptide, protein, and protein complex ions using a radio-frequency confining drift cell, *Analyst* **2016**, *141*, 884-891.

[261] X. Z. Xu, P. D. Wes, H. Chen, H. S. Li, M. Yu, S. Morgan, Y. Liu, C. Montell, Retinal targets for calmodulin include proteins implicated in synaptic transmission, *J. Biol. Chem.* **1998**, *273*, 31297-31307.

[262] X. Shen, C. A. Valencia, J. W. Szostak, B. Dong, R. Liu, Scanning the human proteome for calmodulin-binding proteins, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5969-5974.

[263] T. Berggard, G. Arrigoni, O. Olsson, M. Fex, S. Linse, P. James, 140 mouse brain proteins identified by Ca2+-calmodulin affinity chromatography and tandem mass spectrometry, *J. Proteome Res.* **2006**, *5*, 669-687.

[264] D. J. O'Connell, M. C. Bauer, J. O'Brien, W. M. Johnson, C. A. Divizio, S. L. O'Kane, T. Berggard, A. Merino, K. S. Akerfeldt, S. Linse*, et al.*, Integrated protein array screening and high throughput validation of 70 novel neural calmodulin-binding proteins, *Mol. Cell. Proteomics* **2010**, *9*, 1118-1132.

[265] O. V. Nemirovskiy, R. Ramanathan, M. L. Gross, Investigation of calcium-induced, noncovalent association of calmodulin with melittin by electrospray ionization mass spectrometry, *J. Am. Soc. Mass. Spectrom.* **1997**, *8*, 809-812.

[266] J. Pan, L. Konermann, Calcium-Induced Structural Transitions of the Calmodulin−Melittin System Studied by Electrospray Mass Spectrometry: Conformational Subpopulations and Metal-Unsaturated Intermediates, *Biochemistry* **2010**, *49*, 3477-3486.

[267] T. L. Pukala, T. Urathamakul, S. J. Watt, J. L. Beck, R. J. Jackway, J. H. Bowie, Binding studies of nNOS-active amphibian peptides and Ca2+ calmodulin, using negative ion electrospray ionisation mass spectrometry, *Rapid Commun. Mass Spectrom.* **2008**, *22*, 3501-3509.

[268] S. Shirran, P. Garnaud, S. Daff, D. McMillan, P. Barran, The formation of a complex between calmodulin and neuronal nitric oxide synthase is determined by ESI-MS, *J. R. Soc. Interface* **2005**, *2*, 465-476.

[269] S. J. Watt, A. Oakley, M. M. Sheil, J. L. Beck, Comparison of negative and positive ion electrospray ionization mass spectra of calmodulin and its complex with trifluoperazine, *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2123-2130.

[270] A. N. Calabrese, L. A. Speechley, T. L. Pukala, Characterisation of Calmodulin Structural Transitions by Ion Mobility Mass Spectrometry, *Austr. J. Chem.* **2012**, *65*, 504-511.

[271] A. N. Calabrese, J. H. Bowie, T. L. Pukala, Structural Analysis of Calmodulin Binding by nNOS Inhibitory Amphibian Peptides, *Biochemistry* **2015**, *54*, 567-576.

[272] Y. S. Babu, C. E. Bugg, W. J. Cook, Structure of calmodulin refined at 2.2 Å resolution, *J. Mol. Biol.* **1988**, *204*, 191-204.

[273] J. L. Fallon, F. A. Quiocho, A Closed Compact Structure of Native Ca2+-Calmodulin, *Structure* **2003**, *11*, 1303-1307.

[274] P. A. Faull, K. E. Korkeila, J. M. Kalapothakis, A. Gray, B. J. McCullough, P. E. Barran, Gas-phase metalloprotein complexes interrogated by ion mobility-mass spectrometry, *Int. J. Mass Spectrom.* **2009**, *283*, 140-148.

[275] N. Lipstein, S. Schaks, K. Dimova, S. Kalkhof, C. Ihling, K. Kolbel, U. Ashery, J. Rhee, N. Brose, A. Sinz*, et al.*, Nonconserved Ca(2+)/calmodulin binding sites in Munc13s differentially control synaptic short-term plasticity, *Mol. Cell. Biol.* **2012**, *32*, 4628-4641.

[276] B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S.-J. Hyung, C. V. Robinson., Ion mobility-mass spectrometry analysis of largen protein complexes, *Nat. Prot.* **2008**, *3*, 1139-1152.

[277] S. Herbst, D. Maucher, M. Schneider, C. H. Ihling, O. Jahn, A. Sinz, Munc13-Like skMLCK Variants Cannot Mimic the Unique Calmodulin Binding Mode of Munc13 as Evidenced by Chemical Cross-Linking and Mass Spectrometry, *PLoS ONE* **2013**, *8*, e75119.

[278] K. Dimova, H. Kawabe, A. Betz, N. Brose, O. Jahn, Characterization of the Munc13-calmodulin interaction by photoaffinity labeling, *Biochim. Biophys. Acta* **2006**, *1763*, 1256-1265.

[279] N. Lipstein, S. Schaks, K. Dimova, S. Kalkhof, C. Ihling, K. Kölbel, U. Ashery, J. Rhee, N. Brose, A. Sinz*, et al.*, Nonconserved Ca(2+)/Calmodulin Binding Sites in Munc13s Differentially Control Synaptic Short-Term Plasticity, *Mol. Cell. Biol.* **2012**, *32*, 4628-4641.

[280] E. Jurneczko, P. E. Barran, How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase, *Analyst* **2011**, *136*, 20-28.

[281] C. J. Hogan, B. T. Ruotolo, C. V. Robinson, J. Fernandez de la Mora, Tandem Differential Mobility Analysis-Mass Spectrometry Reveals Partial Gas-Phase Collapse of the GroEL Complex, *J. Phys. Chem. B* **2011**, *115*, 3614-3621.

[282] J. D. Eschweiler, R. Kerr, J. Rabuck-Gibbons, B. T. Ruotolo, Sizing Up Protein–Ligand Complexes: The Rise of Structural Mass Spectrometry Approaches in the Pharmaceutical Sciences, *Annu. Rev. Anal. Chem.* **2017**, *10*, 25-44.

[283] A. C. Susa, Z. Xia, E. R. Williams, Native Mass Spectrometry from Common Buffers with Salts That Mimic the Extracellular Environment, *Angew. Chem. Int. Ed.* **2017**, *56*, 7912-7915.

[284] J. Gan, G. Ben-Nissan, G. Arkind, M. Tarnavsky, D. Trudeau, L. Noda Garcia, D. S. Tawfik, M. Sharon, Native Mass Spectrometry of Recombinant Proteins from Crude Cell Lysates, *Anal. Chem.* **2017**, *89*, 4398-4404.

[285] P. D. B. Olinares, A. D. Dunn, J. C. Padovan, J. Fernandez-Martinez, M. P. Rout, B. T. Chait, A robust workflow for native mass spectrometric analysis of affinity-isolated endogenous protein assemblies, *Anal. Chem.* **2016**, *88*, 2799-2807.

[286] M. van de Waterbeemd, J. Snijder, I. B. Tsvetkova, B. G. Dragnea, J. J. Cornelissen, A. J. R. Heck, Examining the Heterogeneous Genome Content of Multipartite Viruses BMV and CCMV by Native Mass Spectrometry, *J. Am. Soc. Mass. Spectrom.* **2016**, *27*, 1000-1009.

[287] G. N. Sivalingam, J. Yan, H. Sahota, K. Thalassinos, Amphitrite: A program for processing travelling wave ion mobility mass spectrometry data, *Int. J. Mass Spectrom.* **2013**, *345–347*, 54-62.

[288] M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, J. L. P. Benesch, C. V. Robinson, Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles, *Anal. Chem.* **2015**, *87*, 4370-4376.

[289] J. D. Eschweiler, J. N. Rabuck-Gibbons, Y. Tian, B. T. Ruotolo, CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions, *Anal. Chem.* **2015**, *87*, 11516-11522.

Appendix A

Fragment Screening with Native Mass Spectrometry

Table A.1: Overview of the investigated fragments and their chemical properties. Fragments were studied with the proteins MTH1 (M1 - M33), with KDM5B (K01 – K16), with BRPF1 (B01 – B21), and UHRF1 (U01 – U21), respectively.

Fragment ID	MW (Da)	clog_{10}(D)@pH 7.5 [a]	TPSA [Å^2][b]	H-bond donors	H-bond acceptors	Rotatable bonds
M01	232.13	2.02	15.27	1	2	1
M02	254.78	2.29	26.02	2	2	1
M03	224.75	2.19	26.02	1	1	1
M04	301.25	2.39	6.48	0	2	2
M05	177.24	0.85	32.26	2	2	1
M06	199.25	2.19	33.12	1	2	0
M07	212.25	1.81	41.99	1	2	2
M08	186.25	2.83	24.92	1	2	2
M09	196.29	0.74	32.34	1	2	1
M10	240.34	1.73	55.56	1	3	2
M11	247.12	1.50	39.12	2	1	1
M12	199.25	1.35	40.81	2	1	0
M13	240.34	1.76	15.27	1	2	2
M14	235.28	2.46	30.71	0	2	1
M15	295.20	2.54	15.60	0	3	1
M16	160.18	1.05	54.46	1	3	1
M17	187.24	2.10	31.92	1	2	2
M18	170.21	2.16	38.90	1	2	1
M19	222.69	1.43	39.06	1	3	1
M20	163.22	1.86	35.25	1	2	1
M21	246.35	3.15	32.34	1	1	2
M22	207.23	0.78	49.77	1	4	2
M23	236.14	0.85	40.71	2	2	2
M24	244.33	2.52	33.62	1	3	2
M25	209.25	2.40	39.12	2	1	1
M26	249.31	2.43	27.63	1	3	1
M27	268.11	1.76	50.94	1	3	3
M28	215.25	2.07	39.08	1	3	1
M29	248.73	1.53	21.26	1	2	2
M30	221.71	1.93	38.37	1	2	0
M31	186.25	2.81	38.90	1	2	2
M32	249.18	2.37	12.03	1	2	2
M33	228.76	2.12	12.03	1	2	2
K01	174.20	0.71	56.73	1	3	2
K02	173.17	0.66	50.19	1	3	1
K03	188.18	1.77	52.08	0	4	2
K04	214.26	1.49	55.12	2	2	2
K05	204.29	0.51	49.41	1	3	1
K06	207.27	1.26	54.02	2	2	3
K07	203.21	1.55	33.12	1	2	2
K08	189.26	2.25	29.85	1	2	2
K09	165.19	0.39	54.02	2	2	2
K10	225.29	2.35	20.31	0	1	2
K11	189.21	1.93	50.95	1	3	3
K12	211.26	2.07	29.10	1	1	1
K13	197.66	2.82	29.10	1	1	2
K14	191.23	1.52	37.72	0	4	2

K15	179.22	1.00	50.19	1	3	2
K16	263.74	1.04	34.15	1	3	2
B01	200.24	2.04	32.67	0	2	0
B02	187.24	2.06	32.86	1	1	2
B03	186.21	1.41	41.99	1	2	1
B04	277.77	1.36	34.15	1	3	2
B05	190.20	2.08	44.12	0	3	2
B06	272.17	1.25	31.06	2	2	1
B07	209.25	1.47	30.71	0	2	1
B08	189.25	2.18	20.31	0	1	0
B09	233.26	1.61	55.4	1	3	2
B10	186.21	0.99	41.99	1	2	1
B11	205.25	0.91	40.54	1	3	3
B12	201.22	1.3	53.09	2	2	0
B13	207.27	1.93	54.02	2	3	3
B14	213.24	1.01	46.92	1	2	1
B15	215.25	1.58	38.33	1	2	2
B16	192.21	1.17	49.41	1	2	1
B17	222.69	1.43	39.06	1	3	1
B18	200.24	1.98	32.67	0	2	0
B19	233.26	1.00	46.61	0	3	2
B20	255.74	3.17	25.36	0	3	1
B21	201.22	1.70	54.88	1	3	1
U01	342.39	1.67	16.13	2	6	3
U02	254.78	1.83	23.47	1	2	2
U03	191.23	1.60	45.65	2	4	2
U04	195.30	1.96	23.47	1	2	1
U05	203.21	2.09	32.86	1	1	0
U06	192.21	1.42	58.14	2	3	2
U07	221.27	1.73	32.26	2	2	1
U08	210.23	1.43	43.08	0	3	1
U09	182.22	2.06	55.13	1	2	3
U10	271.81	2.01	32.34	1	2	2
U11	199.21	0.78	50.19	1	3	2
U12	294.18	2.06	43.84	1	2	2
U13	184.28	0.88	24.50	1	3	2
U14	202.25	1.06	33.20	0	2	1
U15	216.28	1.16	33.20	0	2	2
U16	189.26	1.75	28.49	0	3	2
U17	213.28	2.74	21.06	0	2	1
U18	226.68	1.86	52.32	1	3	2
U19	204.22	2.51	43.37	0	3	2
U20	227.26	1.87	57.78	2	2	2
U21	191.23	1.16	40.54	1	3	2

[a] $\mathrm{clog}_{10}(D)$ = calculated distribution coefficient – octanol water distribution; $\log_{10}(D)$ takes into account the compounds ionized and non-ionized forms and therefore the measurement is done at different pH.

[b] TPSA = topological polar surface area – metric for the optimization of a drug's ability to permeate cells

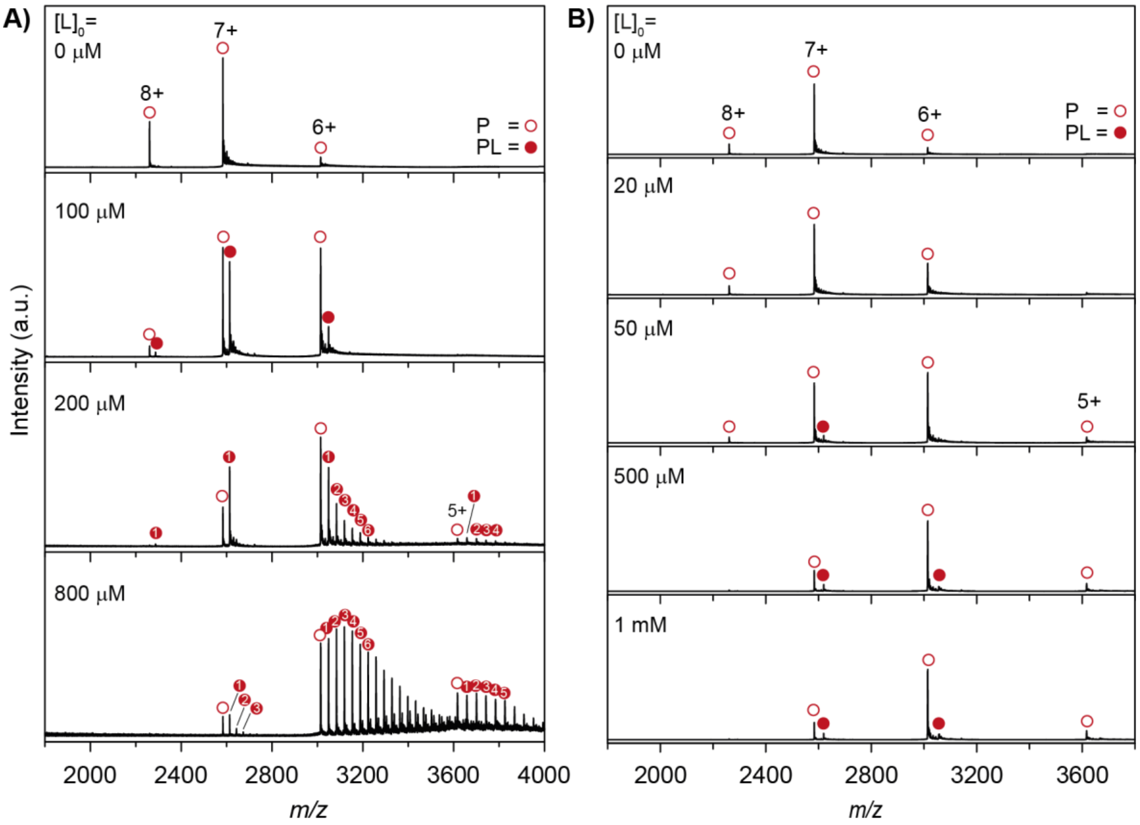


Figure A.1: A) Multiple, non-specific binding increases with ligand concentration. ESI-MS titration of target protein MTH1 with fragment **M25** in presence of 1 % DMSO. The protein concentration is kept constant at 10 µM, while the fragment concentration is varied from 0 to 800 µM. With increasing fragment concentration excessive multiple binding, especially in the lower charge states 6+ and 5+ occurs (indicated by filled red circles with numbers). In addition to multiple binding a charge-state shift is observed with increasing fragment concentrations. **B)** The effect of charge-state shift is concentration-dependent and therefore likely to reflect binding in solution. ESI-MS titration of MTH1 (10 µM) with increasing concentration of fragment **M26** (0 to 1000 µM) in presence of 1 % DMSO. With increasing fragment concentration a shift of charge-state intensities from 7+ to 6+ becomes apparent.

Table A.2: Comparison of manual and automated software-based data analysis using BiopharmaLynx (Waters) from an MS screen against MTH1 (10 µM protein with 100 µM fragment).

Fragment ID[a]	PL (%)[b] automated	PL (%)[b] manual	Fragment ID[a]	PL (%)[b] automated	PL (%)[b] manual
M01	29	30	**M18**	13	22
M02	22	24	**M19**	0	12
M03	0	9	**M20**	17	19
M04	0	7	**M21**	17	17
M05	34	33	**M22**	11	12
M06	0	7	**M23**	35	36
M07	7	8	**M24**	0	7
M08	10	12	**M25**	33	37
M09	14	15	**M26**	13	15
M10	58	46	**M27**	10	10
M11	36	37	**M28**	6	10
M12	53	53	**M29**	49	48
M13	52	52	**M30**	70	66
M14	15	20	**M31**	33	35
M15	19	20	**M32**	22	23
M16	32	27	**M33**	33	34
M17	11	12			

[a] Underlined fragments show multiple binding. [b] Value corresponds to the average of triplicate measurements.

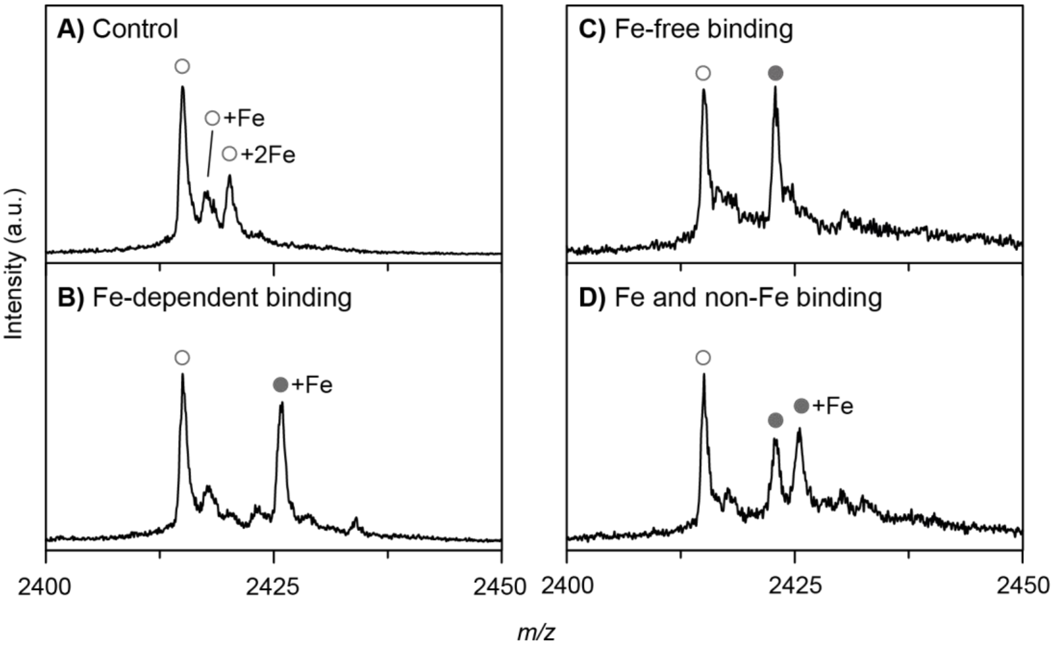


Figure A.2: Comparison of different binding modes observed with KDM5B exemplarily for charge state 23+. **A)** Control ESI mass spectrum of KDM5B that indicates coordination of 0, 1 or 2 irons. **B)** Spectrum with a ligand that binds only in the presence of iron (potential iron coordinator, *e.g.* **K03**). **C)** Spectrum with a ligand that displaces the iron upon binding (potential iron competitor, *e.g.* **K01**). **D)** Spectrum with a ligand that binds irrespectively of the presence of iron (*e.g.* **K09**).

Table A.3: Comparison of MS results with the orthogonal technique TSA for a screen of 21 fragments against the bromodomain of BRPF1 (10 μM protein with 500 μM and 100 μM fragment). Multiple binding is observed for all fragments at a concentration of 500 μM.

Fragment ID	MW (Da)	PL (%)[a] 500 μM/ 100 μM	css	ESI-MS hit[b] 500 μM / 100 μM	TSA hit[c]	TSA ΔT_m (°C)	TSA category	X-ray structure (unpublished)
B01	200.24	0/ 0	no	no	yes	2.7	stabilizer	yes
B02	187.24	2/ 0	no	no	yes	1.5	stabilizer	yes
B03	186.21	3/ 0	yes	no	yes	1.7	stabilizer	yes
B04	277.77	2/ 0	no	no	yes	2.1	stabilizer	yes
B05	190.20	0/ 0	no	no	no	0.3	neutral	
B06	272.17	6/ 1	no	no	no	0.2	neutral	
B07	209.25	2/ 0	no	no	yes	1.4	stabilizer	yes
B08	189.25	0/ 0	no	no	no	0.3	neutral	
B09	233.26	27/ 11	no	yes/ no	yes	1.6	stabilizer	yes
B10	186.21	10/ 2	no	no	yes	1.9	stabilizer	
B11	205.25	13/ 5	no	no	no	-1.0	destabili	
B12	201.22	9/ 4	no	no	yes	1.6	stabilizer	
B13	207.27	11/ 3	no	no	no	-4.8	destabili	
B14	213.24	15/ 4	no	no	yes	1.6	stabilizer	
B15	215.25	20/ 8	no	yes/ no	no	-0.9	destabili	
B16	192.21	17/ 4	no	no	no	-0.4	neutral	
B17	222.69	3/ 0	no	no	no	-7.3	destabili	
B18	200.24	1/ 0	no	no	yes	1.4	stabilizer	
B19	233.26	16/ 5	no	no	no	-0.4	neutral	
B20	255.74	9/ 5	no	no	no	-4.5	destabili	
B21	201.22	0/ 0	no	no	yes	1.5	stabilizer	

Table A.4: Comparison of MS results with the orthogonal technique TSA for a screen of 21 fragments against the SRA domain of UHRF1 (10 μM protein and 200 μM).

Fragment ID [d]	MW (Da)	PL (%)[a]	css	ESI-MS hit[b]	TSA hit[c]	TSA ΔT_m (°C)	TSA category
U01	342.39	22	no	yes	no	0.14	neutral
U02	254.78	0	no	no	yes	0.98	stabilizer
U03	191.23	0	no	no	yes	0.88	stabilizer
U04	195.30	0	yes	no	yes	0.85	stabilizer
U05	203.21	0	yes	no	no	-0.02	neutral
U06	192.21	9	yes	no	no	0.14	neutral
U07	221.27	10	yes	no	yes	0.87	stabilizer
U08	210.23	0	yes	no	no	-0.57	neutral
U09	182.22	0	yes	no	no	0.35	neutral
U10	271.81	3	no	no	yes	0.87	stabilizer
U11	199.21	19	no	no	no	-0.01	neutral
U12	294.18	3	yes	no	no	-0.57	neutral
U13	184.28	0	no	no	yes	0.94	stabilizer
U14	202.25	7	yes	no	no	-0.57	neutral
U15	216.28	8	yes	no	no	0.36	neutral
U16	189.26	0	no	no	yes	0.99	stabilizer
U17	213.28	0	yes	no	yes	0.90	stabilizer
U18	226.68	0	yes	no	yes	0.96	stabilizer
U19	204.22	0	yes	no	no	-0.58	neutral
U20	227.26	18	no	no	no	-0.57	neutral
U21	191.23	7	no	no	yes	1.01	stabilizer

[a] % PL determined by manual data analysis, where every charge state was considered. [b] A threshold of 20 % PL formation was set as hit criterion. [c] Stabilizing fragments were considered as hits, if ΔT_m exceeds the protein's melting temperature by six times the standard deviation. [d] Underlined fragments show multiple binding.

Appendix B

Ubiquitin Microsolvation

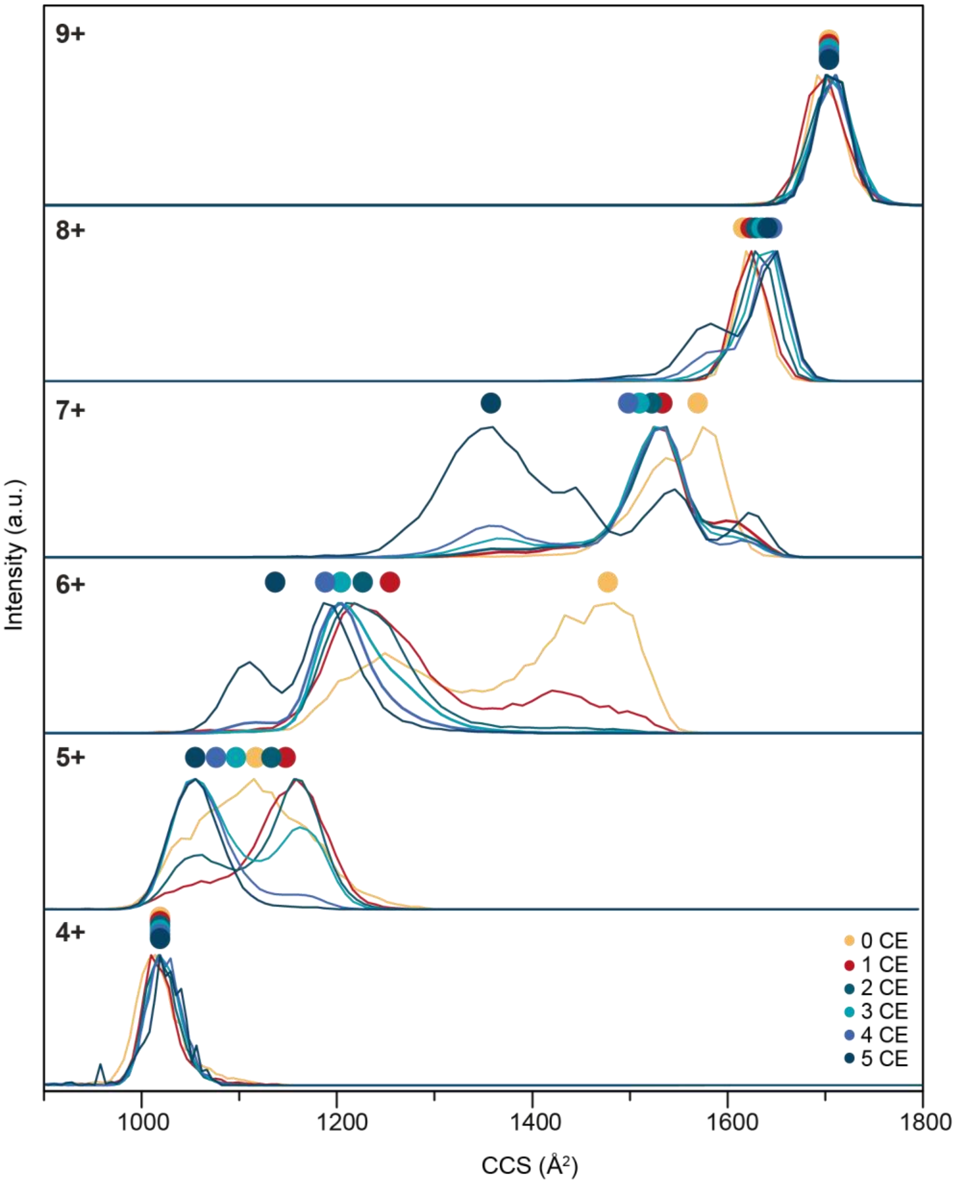


Figure B.1: Collision cross sections (CCSs) of ubiquitin wild type for different charge states and different numbers of attached crown ether molecules (CE). The average size increase of the protein per additional CE molecule (1.6 % to 1.7 %) is subtracted from the CCS of the complex. Intermediate charge states (5+ to 7+) undergo a compaction, whereas globular (4+) and unfolded (8+ and 9+) ions do not show this effect upon crown-ether attachment. The compaction of the intermediate charge states differs slightly depending on the charge state: 5+ ions adopt a compact form (~1050 Å^2) with 2CE to 3CE attached, whereas 6+ and 7+ both first adopt an intermediate form (~1250 Å^2 for 6+ and 1350 Å^2 for 7+) with 1CE to 4CE bound and only 6+ also shows the compact form with 5CEs attached (~1100 Å^2).

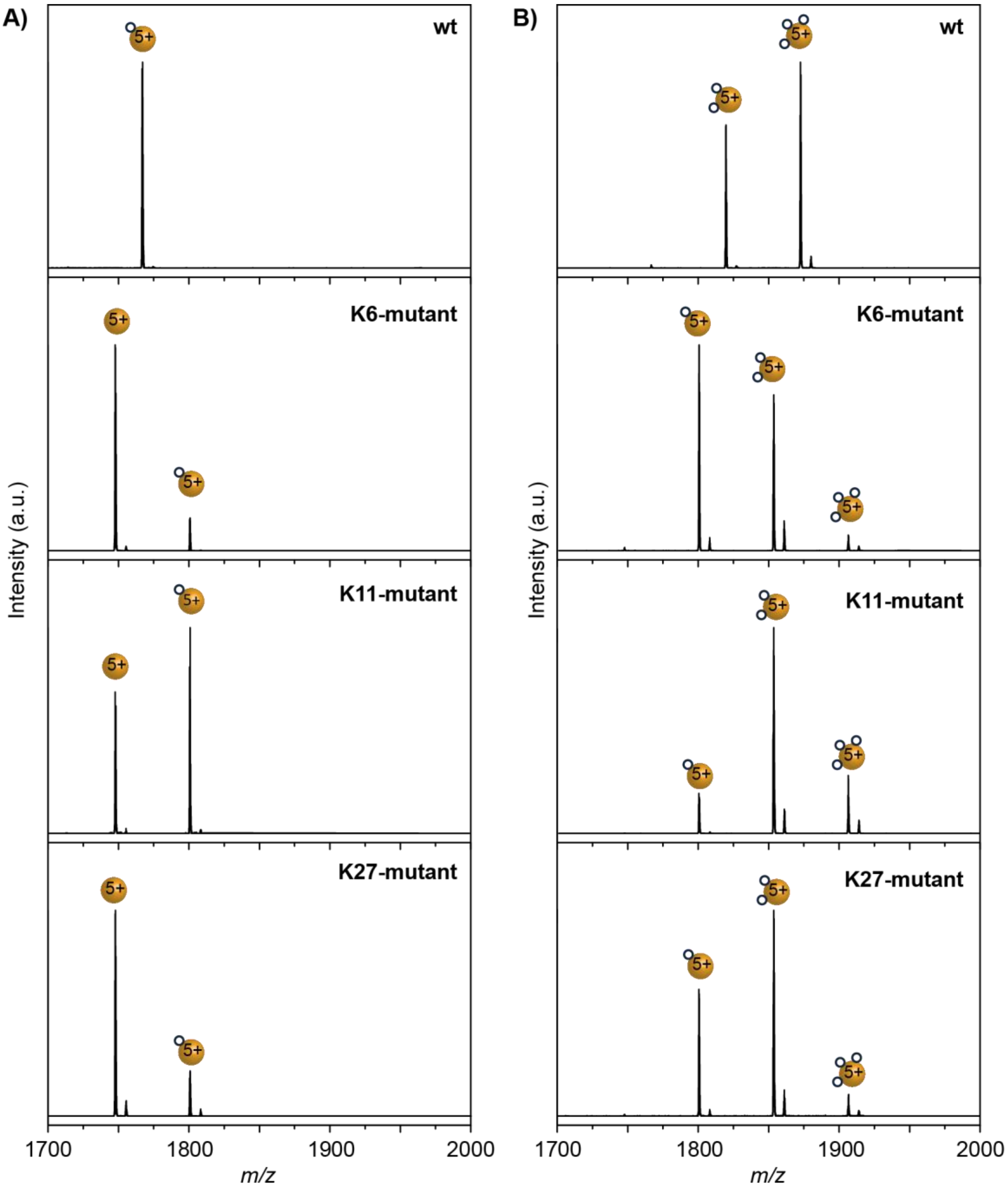


Figure B.2: MS/MS-experiments of both wild type and mutants. **A)** 1CE complexes at 16 V collision voltage and **B)** 3CE complexes at 10 V collision voltage. At 9 V collision voltage (Figure 5.4, noK) very low fragmentation is observed for wt and other variants.

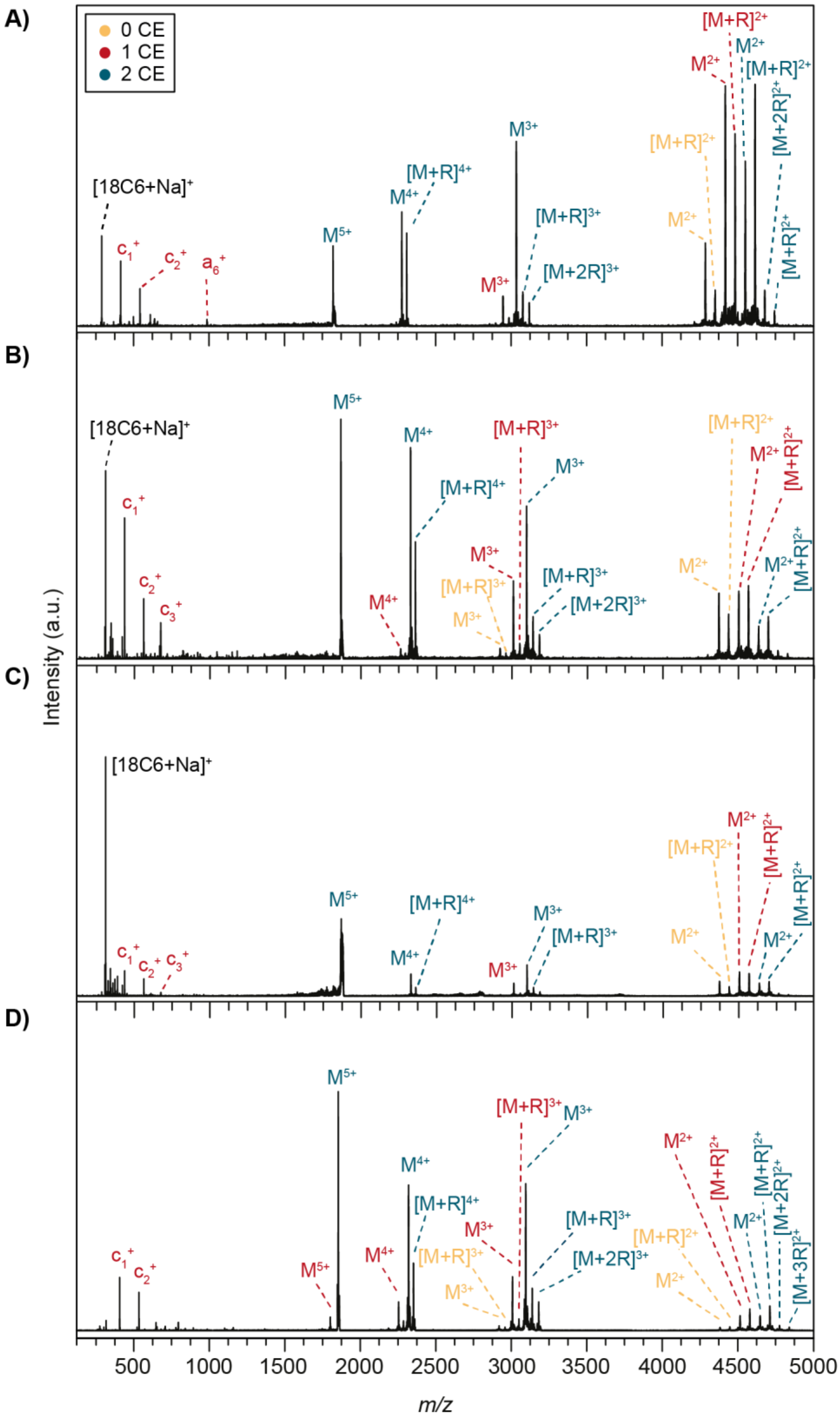


Figure B.3: Complete ETD spectra for 5+ precursor complex (M^{5+}) with two CEs. **A)** wild-type ubiquitin, **B)** K6-mutant, **C)** K11-mutant, **D)** noK-mutant.

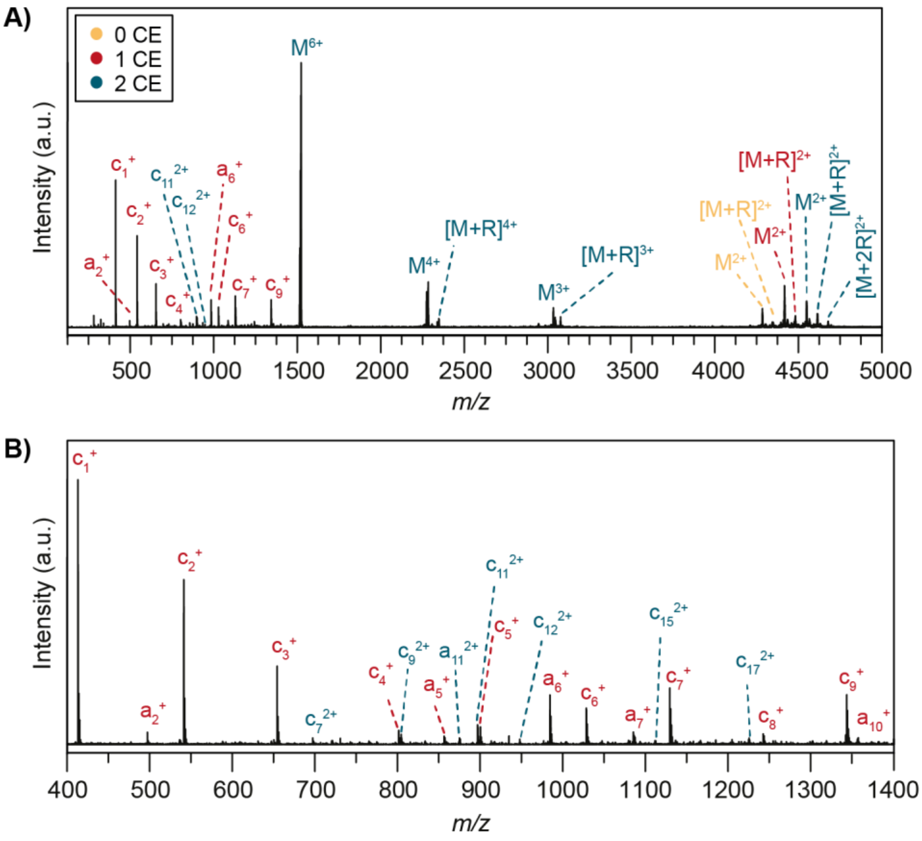


Figure B.4: A) Complete ETD spectrum for a 6+ precursor wt-complex with two CEs (M^{6+}). **B)** Extension into the *m/z* range 400 to 1400. Fragments with one and two CEs attached to wt ubiquitin are observed. Intense peaks are formed, which indicate N-terminal binding of the first CE (c_1^+, a_2^+, c_2^+, c_3^+, a_4^+, c_4^+, a_5^+, c_5^+), but also fragments with one CE indicating attachment to the N-terminus or to K6 (a_6^+, c_6^+, a_7^+, c_7^+, c_8^+, c_9^+, a_{10}^+) appear with reasonable intensity. The fragments $c_7^+$and c_9^+ with two CEs (green) show N-terminal and K6-binding, whereas the longer sequenced fragments (a_{11}^{2+}, c_{11}^{2+}, c_{12}^{2+}, c_{15}^{2+} and c_{17}^{2+} could also indicate binding to K11.

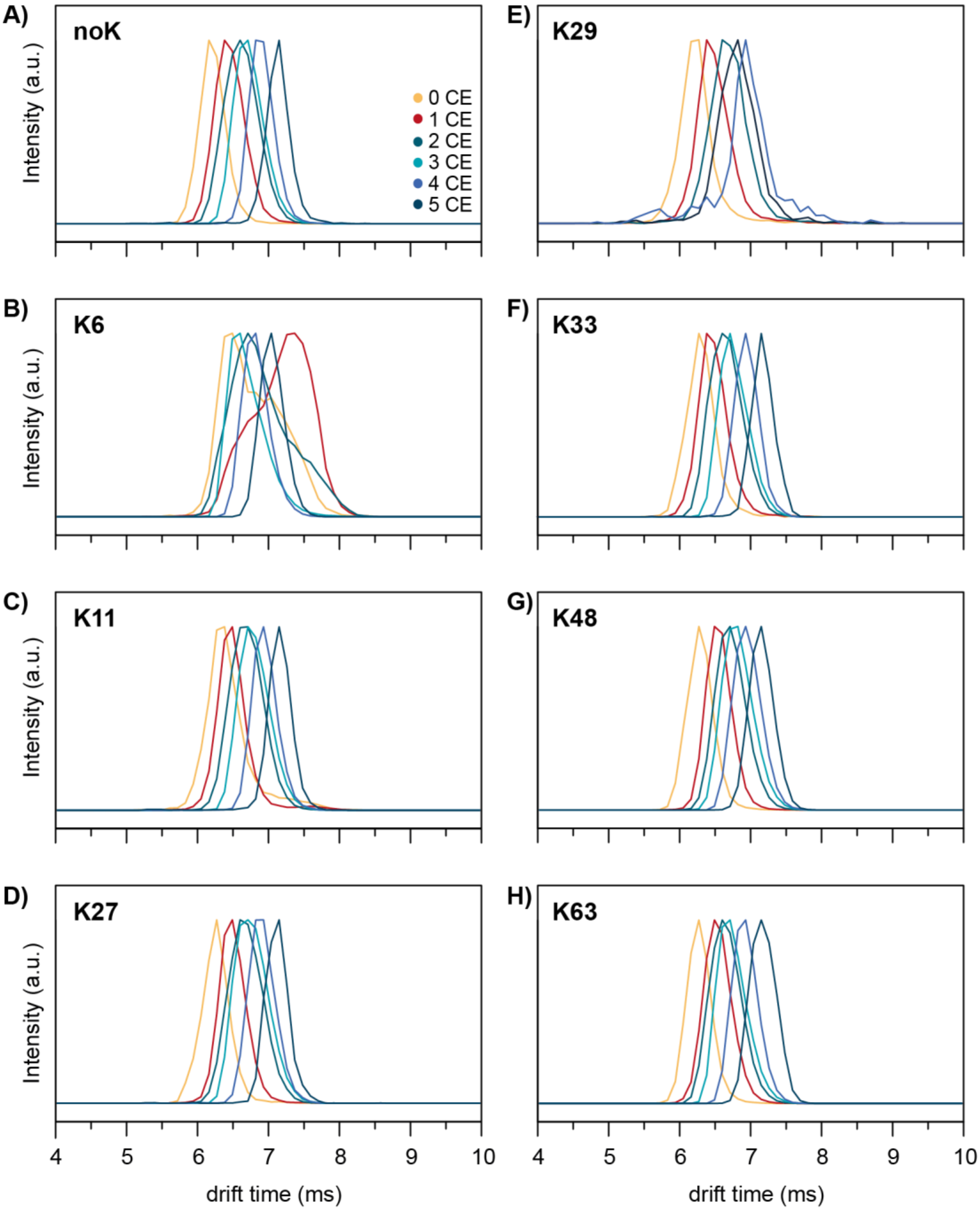


Figure B.5: Arrival time distributions (ATDs) with a wave velocity of 700 m/s show charge state 5+ for all possible lysine-to-arginine mutants **(A)** without lysine and **B) – H)** with one remaining lysine at a specific position. The different colors indicate different numbers of crown-ether (CE) molecules attached (from 0 to 5 CE complex). Only the K6-mutant shows the same conformational heterogeneity as the wt and structural compaction is observed upon attachment of CEs. ATDs of the other mutants are homogeneous and narrow. They are compact already without the addition of CEs and do not show a significant change in structure when CEs are attached.

Appendix C

Calmodulin complexes with Munc13 peptides

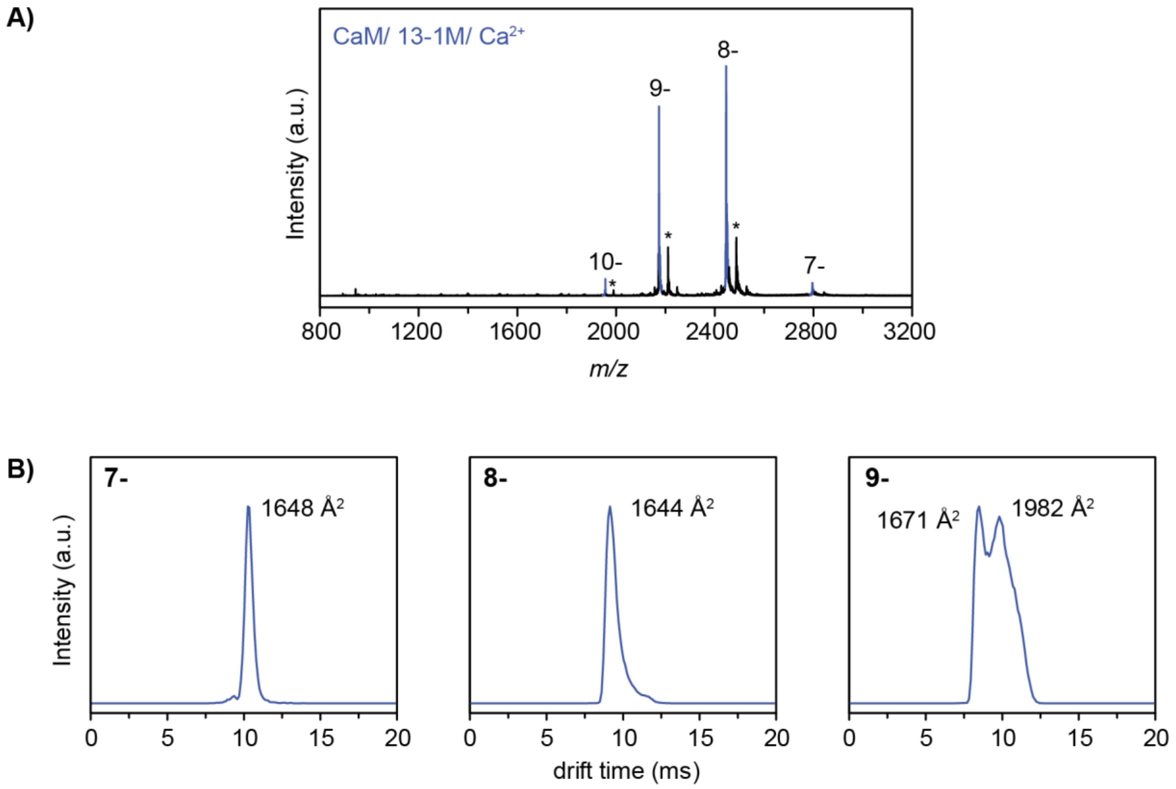


Figure C.1: A) nESI mass spectrum of CaM complex with peptide 13-1M and 150 μM Ca^{2+}. Complex formation occurs predominantly in a 1:1:4 CaM/peptide/Ca^{2+} stoichiometry. **B)** ATDs of 1:1:4 CaM/13-1M/Ca^{2+} complex and determined CCS values for the main conformations.

Table C.1: $^{DT}CCS_{He}$ values for 1:1:4 CaM/ peptide/ Ca^{2+} complexes with Munc13-1 peptides. The values correspond to an average of measurements in triplicate and the errors represent the standard deviation.

Charge state	**CaM/13-1/Ca²⁺**	**CaM/13-1C/Ca²⁺**	**CaM/13-1M/Ca²⁺***
7	1661 ± 11 Å²	1761 ± 13 Å²	1648 ± 2 Å²
8	1671 ± 4 Å²	1755 ± 8 Å²	1644 ± 14 Å²
9	1726 ± 14 Å²	1756 ± 10 Å²	1671 ± 23 Å²
9	1993 ± 16 Å²	1883 ± 24 Å²	1982 ± 29 Å²
10	Low intensity	2387 ± 74 Å²	Low intensity

*average of measurement in duplicate

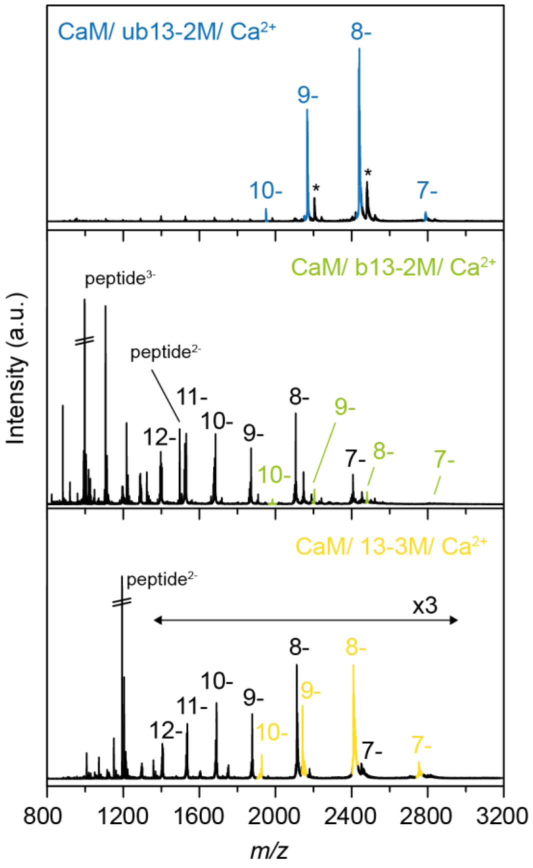


Figure C.2: nESI mass spectra of CaM/peptide/Ca^{2+} complexes (10 µM: 20 µM: 150 µM) with mutant peptides of ubMunc13-2 (ub13-2M, blue), bMunc13-2 (b13-2M, green), and Munc13-3 (13-3M, yellow). Mass spectra show complexes of a 1:1:4 stoichiometry (CaM/peptide/Ca^{2+}).

Table C.2: ^{DT}CCS$_{He}$ values for 1:1:4 CaM/peptide/Ca^{2+} complexes with Munc13 derived peptides. The values correspond to an average of measurements in triplicate and the errors represent the standard deviation.

Charge state	CaM/ub13-2/Ca^{2+}	CaM/ub13-2C/Ca^{2+}	CaM/ub13-2M/Ca^{2+}
7	1646 ± 13 Å^2		1642 ± 5 Å^2
8	1669 ± 2 Å^2	1734 ± 3 Å^2	1641 ± 2 Å^2
9	1719 ± 13 Å^2	1747 ± 5 Å^2	1681 ± 22 Å^2
9	Low intensity	[b] 1985 Å^2	1996 ± 11 Å^2
10	[a] 2047 ± 7 Å^2	2086 ± 25 Å^2	Low intensity
10	[a] 2181 ± 20 Å^2	Low intensity	[a] 2259 ± 5 Å^2
	CaM/b13-2/Ca^{2+}	**CaM/b13-2C/Ca^{2+}**	**CaM/b13-2M/Ca^{2+}**
7	1698 ± 20 Å^2	1766 ± 4 Å^2	
8	1723 ± 1 Å^2	1787 ± 4 Å^2	1708 ± 8 Å^2
9	1758 ± 24 Å^2	1863 ± 26 Å^2	1782 ± 43 Å^2
9	2027 ± 43 Å^2	2120 ± 57 Å^2	2004 ± 58 Å^2
9	2155 ± 55 Å^2		Low intensity
10	2359 ± 11 Å^2	2064 ± 39 Å^2	Low intensity
10	Low intensity	2348 ± 39 Å^2	Low intensity
	CaM/13-3/Ca^{2+}	**CaM/13-3C/Ca^{2+}**	**CaM/13-3M/Ca^{2+}**
7	1667 ± 16 Å^2	1711 ± 11 Å^2	1647 ± 8 Å^2
8	1700 ± 31 Å^2	1734 ± 5 Å^2	1674 ± 8 Å^2
9	Low intensity	1786 ± 30 Å^2	[b] 1701 Å^2
9	1970 ± 18 Å^2	[a] 1988 ± 3 Å^2	[b] 2019 Å^2
10	Low intensity	[a] 2108 ± 45 Å^2	Low intensity
10	[a] 2205 ± 75 Å^2	2339 ± 9 Å^2	[a] 2333 ± 34 Å^2

[a] average of measurement in duplicate, [b] CCS determination only once, conformer with low intensity

Table C.3: Experimental ^{DT}CCS$_{He}$ for the CaM/Ca$_4^{2+}$ complex and parameters for estimation of mass-to-CCS correlation for idealized spherical particles (globular conformation) and cylindrical-shaped particles (dumbbell conformation).

PDB	M (Da)	z	Ω_{exp} (Å^2)	Ω_{av} (Å^2)	d_{eff} (g/cm^3)	r_{eff} (Å)	V_{eff} (cm^3)	h (Å^2)
1PRW	16858	7	1531	1597	0.67	21.5	41886	
		8	1595					
		9	1664					
3CLN	16858	9	1913	1913	0.26	23.7	109188	62

Idealized spherical particles:

$$d_{eff} = \frac{M}{N_A} * \frac{1}{V_{eff}}$$

(C1)

$$V_{eff} = \frac{4}{3} * \pi * {r_{eff}}^3$$

(C2)

$$r_{eff} = (\frac{\Omega av}{\pi})^{1/2} - r_{He}$$

(C3)

With $N_A = 6.022 * e^{23}$, $r_{He} = 1$ Å.

Idealized cylindrical particles:

$$V_{eff} = \pi * {r_{eff}}^2 * h$$

(C4)

Table C.4: $^{DT}CCS_{He}$ values for 1:1:4 CaM/peptide/Ca^{2+} complexes with Munc13 derived peptides at activated conditions (charge state 8-).

Peptide	*m/z*	Collision Voltage (V)	CCS (Å^2)*
13-1	2430	60	2081
ub13-2	2444	30	2102
b13-2	2466	30	2130
13-3	2403	30	2102
13-1C	2610	60	2147
ub13-2C	2605	40	2134
b13-2C	2600	27	1951/2155
13-3C	2537	30	2116
13-1M	2446	60	2091
ub13-2M	2440	30	2100
b13-2M	2481	30	2127
13-3M	2404	30	2088

*CCS error < 0.5 %

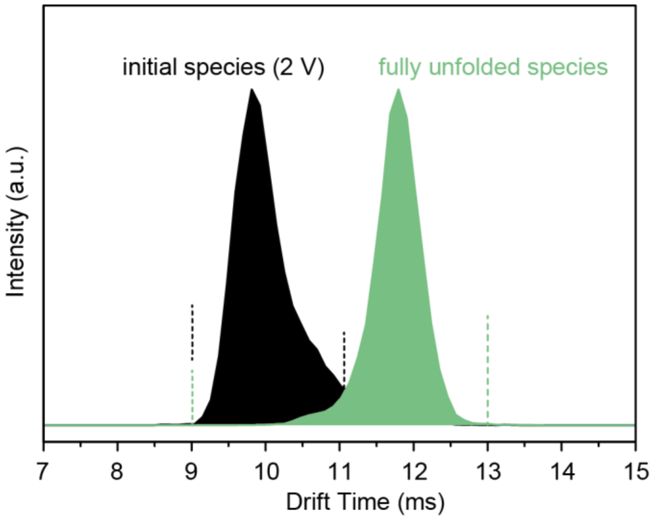


Figure C.3: Illustration for the determination of the relative unfolding of a species: ATDs of initial species at 2 V collision voltage (black) and fully unfolded species (starting ~ 40 V, green). For the calculation of the relative unfolding the integral of the initial ATD is divided by the integral of the total complex ATD. Integration limits are indicated in dotted black and green lines.

Table C.5: 50 % CIU and CID values for 1:1:4 CaM/peptide/Ca^{2+} complexes with Munc13 derived peptides. The values correspond to an average of measurements in triplicate and the errors represent the standard deviation. A Boltzmann function was employed for fitting the data.

Complex	50 % CIU (V)	50 % CID (V)
CaM/13-1/Ca^{2+}	15.3 ± 0.1	61.6 ±1.0
CaM/13-1C/Ca^{2+}	20.3 ± 0.6	64.7 ± 0.3
CaM/13-1M/Ca^{2+}	17.4 ± 1.7	62.7 ± 1.1
CaM/ub13-2/Ca^{2+}	16.5 ± 0.8	53.3 ± 4.7
CaM/ub13-2C/Ca^{2+}	17.3± 1.1	55.5 ± 0.9
CaM/ub13-2M/Ca^{2+}	16.7 ± 1.1	61.0 ± 3.2
CaM/b13-2/Ca^{2+}	18.0 ± 0.8	58.2 ± 0.8
CaM/b13-2C/Ca^{2+}	24.4 ± 0.6	60.6 ± 2.3
CaM/b13-2M/Ca^{2+}	17.2 ± 0.4	51.2 ± 0.9
CaM/13-3/Ca^{2+}	12.8 ± 1.7	16.8 ± 6.0
CaM/13-3C/Ca^{2+}	16.7 ± 1.0	16.3 ± 2.5
CaM/13-3M/Ca^{2+}	15.1 ± 1.6	35-40*

*The experiment did not lead to an unambiguous result, also after several repetitions. Fitting was not possible

List of Publications

M. Göth, V. Badock, J. Weiske, K. Pagel, B. Kuropka, *Critical Evaluation of Native Electrospray Ionization Mass Spectrometry for Fragment-Based Screening, ChemMedChem,* **2017**, *12*, 1201-1211.
doi: 10.1002/cmdc.201700177

M. Göth, K. Pagel, *Ion Mobility-Mass Spectrometry as a Tool to Investigate Protein-Ligand Interactions*, *Anal. Bioanal. Chem.* **2017**, *409*, 4305-4310.
doi: 10.1007/s00216-017-0384-9

N. Lipstein, **M. Göth**, C. Piotrowski, K. Pagel, A. Sinz, O. Jahn, *Presynaptic Calmodulin Targets: Lessons from Structural Proteomics*, *Expert Rev. Proteomics* **2017**, *14,* 223-242.
doi: 10.1080/14789450.2017.1275966

M. Göth, F. Lermyte, X.J. Schmitt, S. Warnke, G. von Helden, F. Sobott, K. Pagel, *Gas-Phase Microsolvation of Ubiquitin: Investigation of Crown Ether Complexation Sites using Ion Mobility-Mass Spectrometry*, *Analyst* **2016**, *141*, 5502-5510. → Cover picture: Analyst **2016**, *141*.
doi: 10.1039/c6an01377e

Curriculum Vitae

The curriculum vitae is not included in the online version

due to privacy policies.

.

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, 5. Oktober 2017