

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

A Study of Amino Acid - Metal Ion Aggregation Through Ion Mobility Mass Spectrometry and Computational Modeling

Permalink

<https://escholarship.org/uc/item/7mj4j6xw>

Author

Chary, Priyanka Padma

Publication Date

2023

Peer reviewed|Thesis/dissertation

University of California

Santa Barbara

A Study of Amino Acid – Metal Ion Aggregation Through
Ion Mobility Mass Spectrometry and Computational
Modeling

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy

in

Chemistry

By

Priyanka P. Chary

Committee In Charge:

Professor Michael T. Bowers, Chair

Professor Steven K. Buratto

Professor Mattanjah de Vries

Professor Stuart Feinstein

March 2023

The dissertation of Priyanka P. Chary is approved.

Mattanjah de Vries

Steven K. Buratto

Stuart Feinstein

Michael T. Bowers, Committee Chair

January 2023

A Study of Amino Acid – Metal Ion Aggregation Through Ion Mobility Mass Spectrometry
and Computational Modeling

Copyright © 2023

by

Priyanka P. Chary

Acknowledgments

The road to this degree has been a long and arduous one, and I could never have made it this far without the help and guidance of the people around me.

First, I would like to thank my graduate research advisor at UCSB, Dr. Mike Bowers. He gave me opportunities to branch out with my research and try multiple projects, and patiently guided me through the various setbacks I encountered. Thank you so much for constantly believing in me when even I was doubting myself.

Thank you to my fellow group members, for all your assistance and for making the graduate school experience more fun and fulfilling. I am especially grateful to Thomas Wyttenbach, who provided crucial guidance and advice on both the experimental and theoretical halves of my project, allowing me to make my concepts into reality.

Thank you to the staff at the Center for Scientific Computing at UCSB who helped me get the computational portion of my project off the ground, and tirelessly answered all of my Linux questions. Thank you as well to Dr. Christian Bleiholder at FSU, who went above and beyond to help me with any PSA-related issues. Additionally, I would like to thank my committee and everyone within the Chemistry Department who kindly assisted with the progression of my research when I reached out with questions. I have learned so much from so many people while I've been here, and hope to take that knowledge with me into the future.

Above all, I want to thank my family: my mom, my dad and my younger sibling. They have been there for me throughout everything, and I simply would not be here today without their unconditional love and support. Thank you for all the delicious food, for dragging me out of the lab to spend time out in nature, and for keeping me sane throughout the ups and downs of graduate school. And to my dad, thank you too for the countless hours you spent with me on video calls as I wrangled this dissertation into a state of completion. With your tireless encouragement, I have successfully made it to the end and can now cast this precious culmination of my years of research into the fires of Mount Doom.

Thank you Amma, Appa, and Meghana. I love you more than words can ever hope to express.

Curriculum Vitae:

Priyanka Padma Chary

Education:

- Bachelor of Arts in Chemistry, Scripps College, Claremont, CA (2010-2014)
- Doctor of Philosophy in Chemistry (expected), University of California, Santa Barbara, Santa Barbara, CA (2014-present)

Research Positions and Employment:

- Undergraduate Researcher at Scripps College (W.M. Keck Science Department, Claremont Consortium) (2010-2014)

Advisor: Dr. Anthony Fucaloro

Research Subject: Volumetric studies of aqueous alcohol solutions, observing the alcohol-water interface

- Research Assistant at Pomona College (Claremont Consortium) (2014)

Advisor: Dr. Masha Prokopenko

Research Subject: Analyzing the oxygen levels of ancient oceans through the measurement of nitrogen levels in coral samples

- Graduate Researcher at UCSB (2014-present)

Advisor: Dr. Michael T. Bowers

Research Subject: Amino acid/metabolite aggregation, specifically the effect of metal ions on amino acid aggregation pathways

- Teaching Assistant at UCSB (2014-present)

General Chemistry/General Chemistry Lab, Physical Chemistry/Physical Chemistry Lab, Biophysical Chemistry

Publications:

Fucaloro, A.F., Burgess, R., Chary, P., Hu, A., Kan, E., Lakliang, Y., Stewart, R. and Yoshino, A. (2015). A volumetric study of aqueous primary alcohols as functions of concentration and temperature. *Journal of Solution Chemistry*, 44 (11), 2236-2257. <https://doi.org/10.1007/s10953-015-0400-7>

Chary, Priyanka P. (2014). A Volumetric Study of Aqueous Butanol Solutions. *Scripps Senior Theses*. 319. https://scholarship.claremont.edu/scripps_theses/319

Awards:

- Outstanding Service to the Department during the 2018-2019 Academic Year
- Individualized Professional Skills (IPS) Grant Awardee (2020)

Abstract

A Study of Amino Acid – Metal Ion Aggregation Through Ion Mobility Mass Spectrometry
and Computational Modeling

by

Priyanka P. Chary

Protein aggregation has long been a topic of interest within the scientific community due to its connection to various neurodegenerative disorders such as Alzheimer's Disease and Parkinson's Disease. While amyloidogenic fibrils were long thought to be the primary toxic species involved in these disorders, in recent years it has been shown that smaller, intermediary oligomers are more toxic than the mature fibrils. This has prompted additional study into these intermediary oligomers, and subsequently the potential aggregating behaviors of small metabolites such as nucleobases and individual amino acids. In this dissertation I will focus on the aggregation of select amino acids, and examine the effects of various alkali metal cations on the oligomer formation pathway and their potential role in stabilizing and promoting these aggregates. This has been done through a combination of experimental and computational techniques, namely ion mobility mass spectrometry and molecular dynamics.

Part 1: Introduction

The aggregation of proteins into oligomers and amyloidogenic fibrils, and the correlation of these aggregates to various neurodegenerative diseases, has been an area of great interest in recent years. For example, extensive research has shown that the protein amyloid beta forms toxic oligomeric structures which have been found to be correlated to the onset of Alzheimer's disease.^{1,2} Initially focus was placed primarily on amyloidogenic fibrils as the species responsible for the negative symptoms of these neurodegenerative disorders. However, there has been a growing shift towards viewing smaller intermediary oligomers as the predominant toxic species. Now, even proteins which have not been found to form fibrils can still be of scientific interest. Additionally, individual "metabolites" such as nucleobases and free amino acids may exhibit similar aggregating behaviors, as proposed by the "generic amyloid hypothesis". An overview of the generic amyloid hypothesis will be given presently, while its application to the aggregation of amino acids in particular will be discussed in more detail in the next chapter of this dissertation.

The potential negative consequences of protein aggregation in a medical context have been thoroughly investigated in the past. The generic amyloid hypothesis addresses the aggregation of metabolites to form toxic oligomers, and their subsequent relevance to the medical community. However, a greater understanding of metabolite aggregatory pathways can also benefit the field of materials chemistry. Certain small metabolites have been found to self-assemble into ordered structures. Diphenylalanine, for example, assembles into nanotubes which are both thermally and chemically stable. The relative stability of these nanotubes makes them suitable for nanomaterials development as well as other applications

in biotechnology such as diagnostic and therapeutic applications.^{3,4,5} Of course, in order to replicate or modify these metabolite structures for industrial purposes, both knowledge of the intermolecular forces that cause this self-assembly and an understanding of the metabolite assembly pathway are required.

Small metabolites additionally have a third potential use as a model system for more complex systems. Most biological systems of interest are too large to be effectively analyzed by common lab techniques. A large protein, for example, must be broken up into fragments to be run through an average mass spectrometer. One such example is amyloid beta. The full-length amyloid precursor protein (APP) is nigh-impossible to characterize through such techniques and has less direct relevance to aggregation. Therefore, most work is done on the commonly occurring and more epidemiologically relevant peptide fragments A β 1-40 and A β 1-42, cleaved from the precursor protein by various enzymes *in vivo*.⁶ Both of these fragments display varying aggregatory behavior and assemble via different pathways, demonstrating the significance of removing even a few amino acids from the peptide chain.⁷

Additionally, key fragments such as A β 25-35 are even less complicated to study, and analysis of these fragments has been just as helpful in elucidating the amyloid beta assembly mechanism and understanding what specific parts of the protein induce aggregation of toxic end products.^{8,9} By examining the individual components of an aggregating polypeptide, a greater understanding of the overall mechanism behind its aggregating behavior can be found, which is crucial in developing methods to inhibit this aggregation.

Generic Amyloid Hypothesis

Protein aggregation has sparked interest within the scientific community due to potential links to disease as well as the potential application of aggregates for future materials. Though historically studies regarding aggregatory behavior were focused solely on proteins and polypeptides, both of these topics can apply to small metabolites as well.

In a 2015 paper by Shaham-Niv et al., the authors introduced the concept of the “generic amyloid hypothesis”, giving a name to a phenomenon which had already been observed in multiple past experiments.¹⁰ Proteins with no known correlation to neurological diseases, such as myoglobin, were observed to aggregate into amyloidogenic structures.^{11, 12} Additionally, phenylalanine was observed to aggregate in a similar way to these proteins and form amyloidogenic fibrils. These phenylalanine aggregates were conclusively found to be harmful through cytotoxicity assays.¹³ The propensity of phenylalanine to aggregate and form fibrils is relevant to our understanding of phenylketonuria – a genetic metabolic disorder in which the body is no longer able to produce the phenylalanine hydroxylase enzyme, which takes phenylalanine and converts it to tyrosine. Without sufficient quantities of this enzyme, phenylalanine builds up within the body in high concentrations.¹⁴ An excess of phenylalanine as seen under these conditions would be enough to kickstart the oligomerization process, leading to the toxic fibrillar structures which were observed by Shaham-Niv et al.

With that in mind, the authors of this study put forward the hypothesis that such aggregation was not specific to phenylalanine, but in fact a generic phenomenon: if any similar metabolite were to exist in the kind of anomalous concentrations that would lead to phenylalanine aggregating, they too might have the potential to aggregate into potentially

toxic structures. As an aside, metabolites are classified as including free amino acids and nucleobases, and many of these metabolites have been linked to various metabolic disorders, although the exact mechanisms by which the metabolites cause or influence these disorders remain unknown.¹⁵

Ion Mobility Mass Spectrometry

Ion mobility mass spectrometry (IM-MS) was utilized to study the small metabolite aggregates that are described in this dissertation. IM-MS can be used to identify the composition of a sample by determining the experimental cross sections of the component ions, with a certain cross section correlating to a given ion's mass and charge.

Using electrospray ionization, ions are formed at the source and injected into the instrument. The particles pass into an ion funnel, where they are collected before being pulsed into the drift cell. Ions travel down the length of the drift tube, pulled through by a weak electric field applied to the cell. At the same time, the ions undergo collisions with buffer gas molecules within the drift tube. Once the ions have reached the end of the drift tube, they are collected in another ion funnel, and subsequently pulsed into the quadrupole mass filter. Finally, the ions reach the detector at the end of the instrument, and the time it takes for each ion to hit the detector, its "arrival time" t_a , is recorded.

With the quadrupole mass filter, certain species can be selected for through the application of a specific RF voltage that corresponds to the desired mass-to-charge (m/z) ratio. By selecting for a specific mass, an arrival time distribution (ATD) plot is created. Any cluster with that specific mass-to-charge ratio will show up on the ATD, arranged in order of

the time it takes for each of those clusters to reach the detector. Alternatively, one can choose not to utilize the mass filter, allowing all particles to pass unimpeded through the quadrupole and to the detector, resulting in a full mass spectrum of all species in solution. Through analysis of a combination of mass spectra and ATDs, the individual species present in solution can be separated out and identified.

Two factors affect the time it takes for a given ion to pass through the instrument to the detector: its size and its charge. At the same time as the ions are pulled through the drift tube by the applied electric field, they collide with helium buffer gas molecules present within the drift tube. Larger species have a greater surface area, and will thus experience more collisions than a more compact species with an identical charge. Additionally, for two particles of the same size, the ion with the greater charge will experience the pull of the applied electric field more strongly, and will pass through the drift tube faster than the less charged ion.

By combining these two factors, any ion within a sample solution can be identified. For a selected mass-to-charge ratio, ions of different sizes and charges will separate out, appearing as distinct peaks in the ATD plot of arrival time versus signal intensity. The center of each peak, where the signal intensity is at a maximum, corresponds to the arrival time t_a of that ion. The arrival time is plotted as a function of the pressure-to-drift-voltage ratio, p/V , and the slope of the resulting graph gives the mobility K . By converting the mobility K to the reduced mobility K_o , the collision cross section σ of the ion can be derived.¹⁶

To demonstrate this derivation mathematically, we begin by defining the velocity at which the ions travel through the drift tube, v_d :

$$\vec{v}_d = K \cdot \vec{E} \quad (\text{Equation 1})$$

The velocity is equal to the force exerted on the ions by the applied electric field E , multiplied by the mobility K . K accounts for the delaying effect of the opposing frictional force which is exerted on the ions as they collide with buffer gas molecules in the drift tube. Traditionally, the reduced mobility K_o is used instead of K to avoid the dependence of the latter term on the temperature and pressure of the buffer gas.

$$K_o = \left(K \cdot \frac{p}{760 \text{ torr}} \cdot \frac{273.15\text{K}}{T} \right) \quad (\text{Eq. 2})$$

As stated previously, the mobility and thus the reduced mobility can be derived from the slope of the graph of t_a versus p/V (written in the form of $y = mx + b$):

$$t_a = \left(\frac{l^2}{K_o} \cdot \frac{1}{T} \cdot \frac{273.15\text{K}}{760 \text{ torr}} \right) \cdot \frac{p}{V} + t_o \quad (\text{Eq. 3})$$

where l is the length of the drift cell, and t_o is the time spent outside the drift cell before reaching the detector.

Once derived, K_o can be used to find the collision cross section of a given ion, σ using the Mason-Schamp equation:

$$K_o \approx \frac{3q}{16N} \left(\frac{2\pi}{\mu k_b T} \right)^{1/2} \cdot \frac{1}{\sigma} \quad (\text{Eq. 4})$$

where N is the buffer gas number density, μ is the reduced mass of the species that are undergoing collision (the ion and the buffer gas molecule), and k_b is Boltzmann's constant.

The collision cross section of a given oligomer is of significant interest, as it allows for not just the identification of that oligomer when done in tandem with peak fitting, but also gives information about the specific three-dimensional structure of an oligomer. Two clusters with the same mass and charge, but with one in a more compact conformation than the other, can be separated out and identified based on their respective collision cross sections.

Molecular Dynamics/Computational Modeling

Generally, computational modeling can be used to refer to any method which uses computers to simulate and analyze a system, both within and outside of the field of science.¹⁷ One such method is molecular dynamics (MD), a scientific computational technique which has been used in the past for a vast range of applications. These include but are not limited to the elucidation of molecular structures, determining thermodynamic values for systems of interest, and observing the dynamics of a system over some time range.^{18, 19}

The appeal of MD lies in the ability to characterize systems which might be difficult to set up experimentally, or to determine information that cannot be easily obtained from an experiment. For example, one might be able to derive the collision cross section of a polypeptide through IM-MS, or analyze its secondary structure (random coil, beta sheet, etc.) through circular dichroism, but not be able to elucidate specific details of the sample's molecular structure.²⁰ Or it might be of interest to find whether a particular polypeptide oligomer would have a compact or extended structure in solution. In these situations, MD can be used in conjunction with experimental techniques to obtain relevant conformational information, which in turn allows for greater confidence in the experimental results.

Simulated Annealing

In brief, simulated annealing is a method to seek out the lowest-energy, most stable conformation of a computationally modeled oligomer. Though even this algorithm may never be able to reach the global minimum of the potential energy surface, simulated annealing can certainly get closer than standard MD as it is able to bypass relatively high-energy local minima which would otherwise trap the system in an unfavorable and unrealistic conformation.

Simulated annealing gets its name from the algorithm's resemblance to the process of heating solid materials to a high temperature, then annealing them through a repeated cycle of melting and cooling in order to form the most stable and low-energy crystal structure once the material has reached its freezing point.²¹ The general simulated annealing procedure works in a similar way. First, a molecular structure is subjected to high temperatures, allowing for greater access to otherwise inaccessible parts of the potential energy landscape. Then the temperature is decreased in set intervals, giving time between each interval for the system to re-equilibrate, until some final temperature is reached.

The basic mathematical theory behind the simulated annealing algorithm is briefly summarized as follows. The system starts at equilibrium. An atom within the system is picked at random to be displaced by a slight amount, and the new energy of the overall system is calculated. If the change in energy of the system, ΔE , is negative (the system decreased in energy), the atomic displacement is accepted and the new structure is used for the next random displacement. However, if ΔE is positive (the system experienced a net energy gain), the new conformation has some probability for the structure to be accepted, based on a Boltzmann distribution:

$$P = e^{-\Delta E/k_B T} \quad (\text{Eq. 5})$$

This procedure is repeated multiple times as the temperature is gradually decreased, resulting in what should be a low-energy and stable final conformation.^{22, 23, 24}

In a later section of this dissertation, I will describe the specific procedure by which simulated annealing was performed on metabolite clusters. From this computational data, theoretical collision cross sections for each cluster were derived, allowing for the construction of a map of the potential energy of a given cluster versus its cross section across a hundred runs.

Software

The GROMACS molecular dynamics software package was used for all computational calculations within this dissertation, and all simulated annealing used GROMACS' built-in algorithm. GROMACS is an open-source software that was originally developed by the Berendsen group at Groningen University in 1995, and its functionality has been maintained and updated since that time by a community of developers.²⁵ GROMACS has been traditionally used for simulations of biological systems, and provides a high-performance method to carry out molecular dynamics simulations, trajectory analysis, free energy calculations, and other such analytical techniques.²⁶

All computations within GROMACS were done using the OPLS-AA force field, which is an all-atom force field developed by William Jorgenson et al. using potentials derived from liquid state dynamics (hence the name, which stands for “Optimized Potentials for Liquid Simulations”).²⁷ As the name would indicate, OPLS-AA has been primarily used

in the past to characterize solvated systems. However, this force field has been successfully used to model systems such as small molecules in the gas phase as well.^{28, 29}

Theoretical cross sections were calculated by applying the Projected Superposition Approximation (PSA) method to the oligomer structures obtained through molecular dynamics. An overview of this method was first put forth in a 2011 journal article by Christian Bleiholder, Thomas Wyttenbach and Michael T. Bowers, and further characterized and elaborated on in several subsequent papers.^{30, 31, 32, 33, 34}

PSA was developed in order to improve on the accuracy of precursor methods such as Projection Approximation (PA), which runs quickly but becomes increasingly inaccurate for larger molecules. On the other hand, PSA maintains enough computational efficiency that the calculations can be carried out on a reasonable timescale, in comparison to more accurate but much slower methods such as the trajectory method (TJM). The PSA method builds on the previous PA calculations, in which the model is improved by the introduction of an additional shape factor which accounts for the structural features of the analyte when determining its effective collision cross section.

Instrumentation

A mass spectrometer with an electrospray ion source was used to collect the experimental data presented herein. This mass spectrometer was adapted from a previous instrument, both of which were constructed in-house. The precursor ESI mass spectrometer has a 5 cm long drift cell and a nanospray ionization source (Figure 1.1). There are several

components to this instrument: the source where the sample is injected into the instrument, the ion funnel, the drift cell, the quadrupole, and the detector.³⁵

The instrument which was used for the majority of data collection is a higher-resolution mass spectrometer with a 200 cm long drift cell (Figure 1.2). The various components of the instrument are similar to the first, with the exception of the proportionally longer drift cell. The added length of the drift tube allows for more sensitive measurements, since the increased drift time gives the ions a chance to separate out to a greater extent based on their size and charge before they reach the detector. This higher resolution instrument worked especially well for the small metabolite clusters described in this dissertation, as the increased length of the drift tube compensated for their relatively small size and allowed for accurate mobility measurements.

Both instruments make use of nanospray ionization, a subset of electrospray ionization. Generally in electrospray ionization, a liquid sample is ejected at a set flow rate from a needle positioned close to the source inlet. Due to the application of an electric potential between the needle and instrument, the sample is ejected as a stream of charged droplets. The solvent in the droplets evaporates, and the resulting gas-phase ions are pulled towards the source and into the drift tube. Electrospray ionization has long been favored for IM-MS since it allows the sample to undergo a more gentle ionization. This preserves high-charge clusters while minimizing fragmentation, allowing for more of the original solution structure to be preserved in the gas phase. Nanospray ionization has a lower flow rate and higher sensitivity than conventional electrospray, which makes it more efficient for use on costly or otherwise precious samples.^{36, 37}

For nanospray ionization using the above experimental setup, the needles containing the sample solutions were prepared by hand: glass capillary tubes were pulled to form sharp tips and coated in gold to increase their conductivity. Helium was used as the drift cell buffer gas for all experiments.

Figure 1.1: Schematic of the original ESI mass spectrometer. The length of the drift cell is 5 cm.

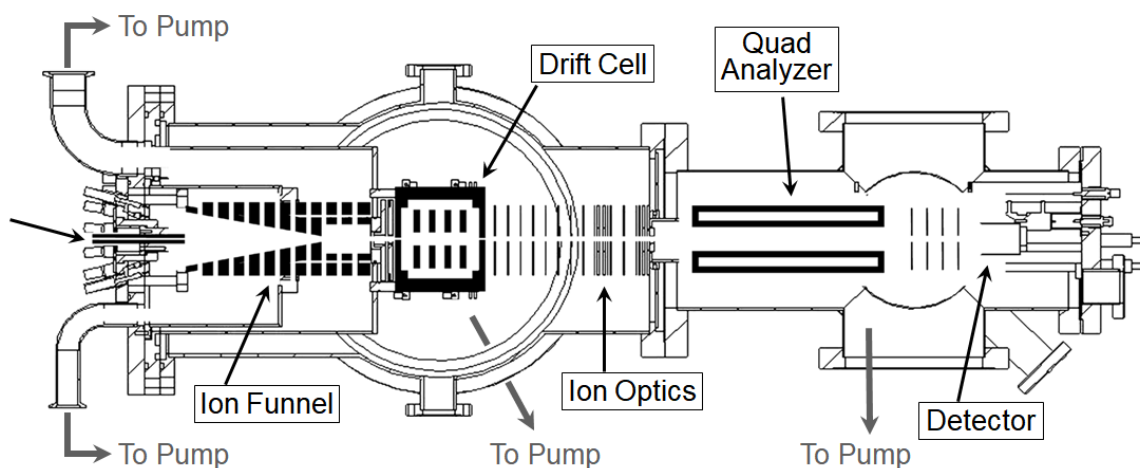
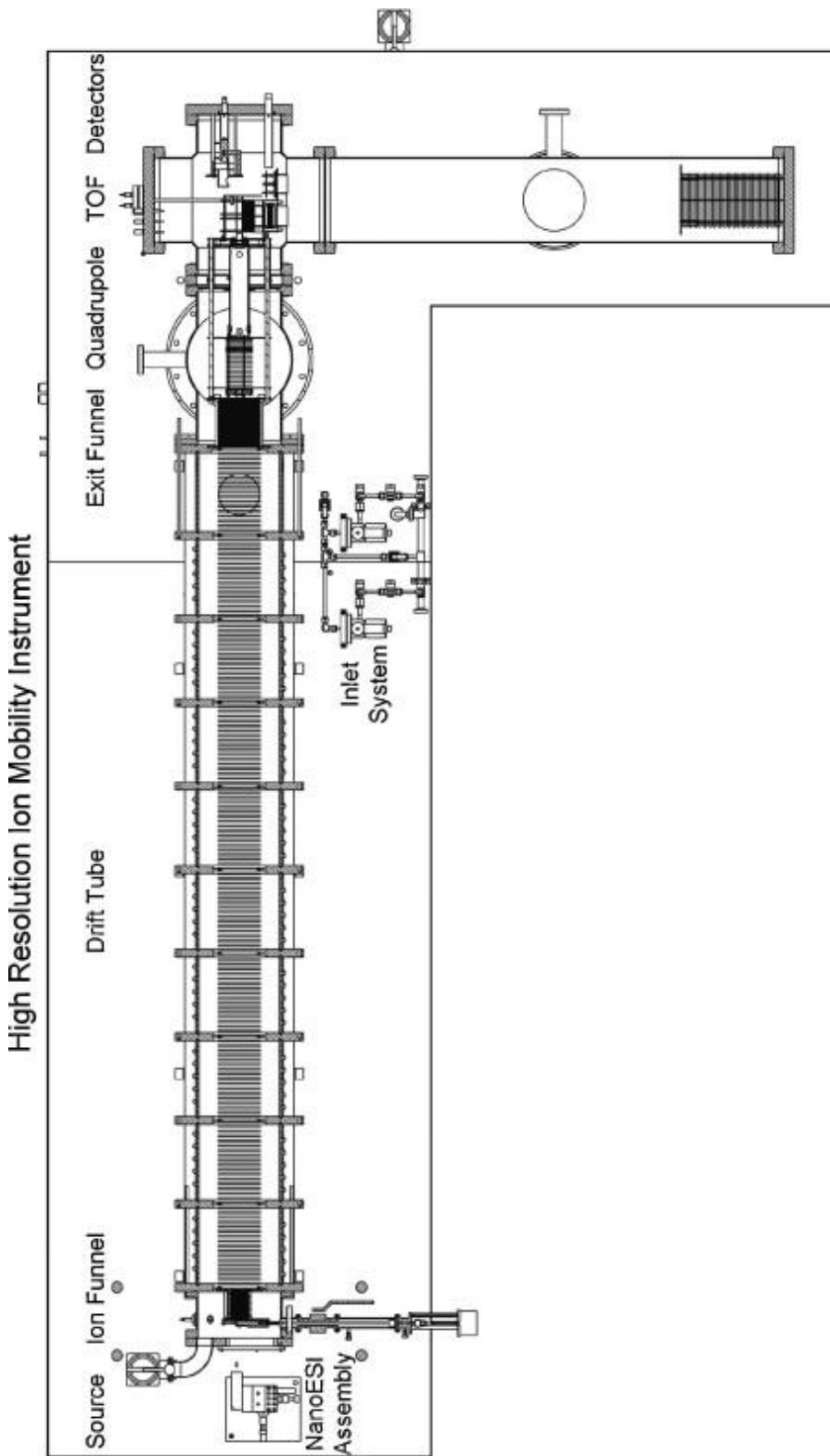


Figure 1.2: Schematic of the high-resolution instrument. Although the instrument has a time-of-flight (TOF) drift tube built in, only the quadrupole component of the instrument was utilized for the experiments laid out in this dissertation. The image is to scale.



Computation

Several Linux systems were used during the span of this research project to run computations and collect theoretical cross sections. The first and most prominent was the Pod computer cluster, located at the UCSB Center for Scientific Computing (CSC) at the California NanoSystems Institute (CNSI).

The second and third machines were personal computers equipped with the Ubuntu operating system, which were acquired in the latter half of this project. The GROMACS open-source software was downloaded from the developer website, compiled, and built to run on these computers. These individual systems are, by definition, less computationally robust than the CSC Pod cluster. However, due to the ability to run simulations instantaneously on local Linux systems as opposed to submitting the input files to a queue, use of these machines greatly decreases the wait times required to begin an MD simulation, which can take more than a week on the Pod cluster.

These two machines were also used to run a desktop version of the PSA software, the source code for which was kindly provided to me by Christian Bleiholder, one of the developers. While initially the public PSA webserver at FSU was used for cross section calculation, the immense lag resulting from running a hundred separate cross section calculations for each individual oligomer on a remote server quickly made the latencies untenable. As with GROMACS, running the PSA software locally allows for the same calculations to be run in a fraction of the time.

Part 2 – Amino Acid Aggregation

The phenomenon of protein aggregation has been extensively studied in past research. Correlations have been shown between the aggregating behaviors of certain proteins such as amyloid beta and α -synuclein to neurodegenerative disorders such as Alzheimer's Disease and Parkinson's Disease.^{38, 39, 40} Particular attention was paid to unraveling the pathways by which these proteins aggregated.⁴¹ Additionally, the relative toxicities of the products of these pathways, ranging from small oligomers to extended structures to mature fibrils, were probed through cytotoxicity studies to determine which of these pathways would result in harmful end products.

In addition to that of proteins and short peptide fragments, aggregation of small metabolites was also studied at this time. Many of these metabolites, encompassing individual amino acids and nucleobases, fall within the scope of the "Generic Amyloid Hypothesis" put forth by Shaham-Niv et al, which has been summarized in the introduction of this dissertation.⁴² The most well-known of these amino acids is phenylalanine.

Phenylalanine has been conclusively linked to a metabolic disorder known as phenylketonuria. This disorder results from a mutation in the phenylalanine hydroxylase (PAH) gene which causes decreased functionality. PAH is a crucial component of the metabolic pathway in which phenylalanine is converted to tyrosine.⁴³ In patients affected with phenylketonuria, phenylalanine builds up in excessive concentrations within the body due to a decreased ability to metabolize it, leading to various negative health effects including neurodegenerative disorders.^{44, 45, 46, 47} This high-concentration buildup of phenylalanine has been observed to lead to oligomerization and ultimately the formation of

amyloidogenic fibrils, an occurrence which so far has not been observed in any other individual amino acid. These fibrils have been shown to exhibit toxicity similar to fibrils formed by more well-known aggregating proteins such as amyloid-beta, making phenylalanine of particular interest from a pathological standpoint.⁴⁸ It must be pointed out that phenylalanine aggregation has also garnered interest for positive reasons, namely the potential for diphenylalanine to assemble into nanotube structures. These nanotubes are remarkably stable, allowing for potential use in various materials applications. Because of its relevance to these two disparate applications, phenylalanine has been by far the most well-characterized amino acid when it comes to aggregation.^{49,50}

Similar studies have also been conducted on the amino acid tryptophan. Tryptophan is involved in a number of key metabolic pathways, including the formation of serotonin.⁵¹ As with phenylalanine, genetic mutations that are associated with metabolic disorders such as hypertryptophanemia can disrupt these pathways. This leads to greatly increased concentrations of tryptophan as the body can no longer adequately process it.⁵² Under these conditions, the formation of oligomers from free monomeric tryptophan is encouraged, and the existence of these toxic assemblies has been experimentally demonstrated.⁵³

In many of these past studies, a clear emphasis was placed on amino acids which were found to assemble into amyloidogenic fibrils, as that was originally thought to be the main toxic species in protein aggregation. However, proteins can also self-assemble into smaller, soluble oligomers as intermediates to fibril formation, or as an end product via alternate pathways.^{54,55} In recent years, it has been shown that while amyloidogenic fibrils are indeed toxic, these smaller clusters can be equally toxic, if not more so, compared to fibrils.^{56,57} Following the generic amyloid hypothesis, a similar result can be predicted for

individual amino acids. Small oligomer intermediates could theoretically also play a role in the many negative side effects associated with metabolic disorders such as phenylketonuria. With this in mind, any metabolite with the potential to form these toxic oligomers can therefore be considered to be of interest, even if their specific propensity for fibril formation is low.

In comparison to the more high-profile phenylalanine, many of the other amino acids have been relatively neglected as potential topics for future study. Preliminary testing revealed limited large-oligomer formation for several of these amino acids, or indeed any oligomer formation beyond dimerization, so little work has been carried out on these metabolites. However, upon further experimentation, some evidence of further aggregation into clusters has been observed for a select number of amino acids.⁵⁸ Two of those amino acids, cysteine and proline, will be focused on in greater detail presently.

In addition to determining differences in aggregation between various amino acids, the effect of trace metal ions in solution (as might be found in vivo), and their possible influence on the aggregation process, is currently being looked at in greater detail. A large assortment of metal ions exist within biological systems, with functions ranging from protein structure stabilization to cellular signaling. Among these are cations such as Na⁺, K⁺ and Ca²⁺. Understanding the way that these metal ions interact with proteins and peptides in these biological systems can be more easily accomplished through an analysis of their influence on the aggregation of free amino acids in isolation.⁵⁹

In a private communication regarding yet-to-be-published data, a group of scientists, including Gert von Helden at the Fritz Haber Institute of the Max Planck Society, examined how various alkali metal ions mediated the aggregation pathway of proline oligomers. Four

specific alkali metal ions – Li^+ , Na^+ , Rb^+ and K^+ – were added to aqueous solutions of proline in a 5mM:1mM ratio of amino acid to metal ion. Using a hybrid quadrupole/time-of-flight mass spectrometer with a nano-electrospray ionization (n-ESI) source, collision cross sections for the proline oligomers in each solution were obtained and compared. Large alkali metal cations such as potassium and rubidium were found to favor the formation of tetramer, octamer and dodecamer proline clusters, while small cations such as lithium instead favored the formation of trimer and hexamer proline clusters. Finally, sodium ions were observed to favor not only trimer and hexamer formation, but tetramer formation as well, most likely due to their medium size which allowed them to mediate both pathways. Notably, all observed clusters were heteroligomers – proline clusters coupled to metal ions.⁶⁰

The correlation of my data to these results will be discussed in the second half of this chapter. While many of the same clusters were observed to form in my own experiment, the influence of the oligomer formation pathways were not as clear-cut as had been seen in previous work. In general, the presence of metal ions did cause the oligomer formation pathways to change, and heteroligomers were seen in abundance alongside pure clusters for every single solution examined. However, correlations of certain metal ions to a specific pathway were not observed. While fluctuations in peak intensity were observed in multiple cases, the overall distributions of oligomers in solution appeared to depend more on the identity of the amino acid than on the identity of the complexing metal ion. These results are consistent with previous studies of pure proline and cysteine, which have been previously seen to form noticeably different peak distributions in mass spectra.

Despite this, interesting conclusions can still be drawn regarding the various oligomerization pathways of certain cysteine and proline clusters complexed with various metal ions.

Experimental Results: Cysteine and Proline

The goal of this particular experiment was twofold. First, to observe the favored oligomers formed by each amino acid both individually and in conjunction with each metal cation. This would allow us to determine the potential pathways by which these species are formed. Second, to characterize these species through modeling and determine the potential effects of the presence of metal cations on oligomer structure.

Before beginning analysis of the following data, a key caveat regarding the experimental results needs to be taken into consideration. As stated in an earlier section of this thesis, all experimental data was collected on a high-resolution electrospray mass spectrometer. At the time this data was collected and extending up to the present day, this particular instrument was afflicted with an irregular mass cutoff which often caused the ion signal to decrease greatly above a certain value. This mass cutoff varied from day to day, but was generally located around 1000 amu. The reasons for this mass cutoff are still being investigated, but may have to do with a technical problem in the instrument's exit funnel.

Unlike many of the larger systems studied by fellow group members, the majority of my amino acid clusters are primarily located below this mass cutoff. This is especially true in the case of cysteine, with its predominant hexamer peak located at 726 amu. However, a drop-off in signal can be observed in several of the following mass spectra, almost certainly

due to this cutoff since peaks have previously been observed at those locations. In cases where clusters had been observed to form around a mass of 1000 amu in previous experimental data, such as the singly charged octamers for both proline and cysteine, these clusters may have presently formed in solution but were not observable in the mass spectrum. The aforementioned limitations of the instrument in this situation should be kept in mind when analyzing the following data in comparison to past experiments.

Despite this caveat, only the largest clusters in these systems would be affected. The majority of cysteine and proline clusters observed in this experiment fall below this cutoff and can be considered reliable.

Proline Oligomerization

We begin with proline. In keeping with past experiments, solutions of proline with the various metal ions were prepared in water, with a ratio of 5mM amino acid to 1mM metal ion. On its own, proline forms a large number of higher-order oligomers, with clusters up to a singly charged dodecamer $[12\text{Pro}]^{+1}$ having been observed (Figure 2.2). On the mass spectrum, the peaks corresponding to the $[6\text{Pro}]^{+1}$ hexamer to the $[9\text{Pro}]^{+1}$ nonamer are most prominent. This broad assortment of proline oligomers from low to high mass is consistent with previous findings. When it comes to the ATDs, these mass-to-charge peaks can be further resolved, and higher-order aggregates above the dodecamer are clearly shown to appear (Figure 2.3). This aligns with the Generic Amyloid Hypothesis, which predicts the formation of these potentially toxic higher-order oligomers.

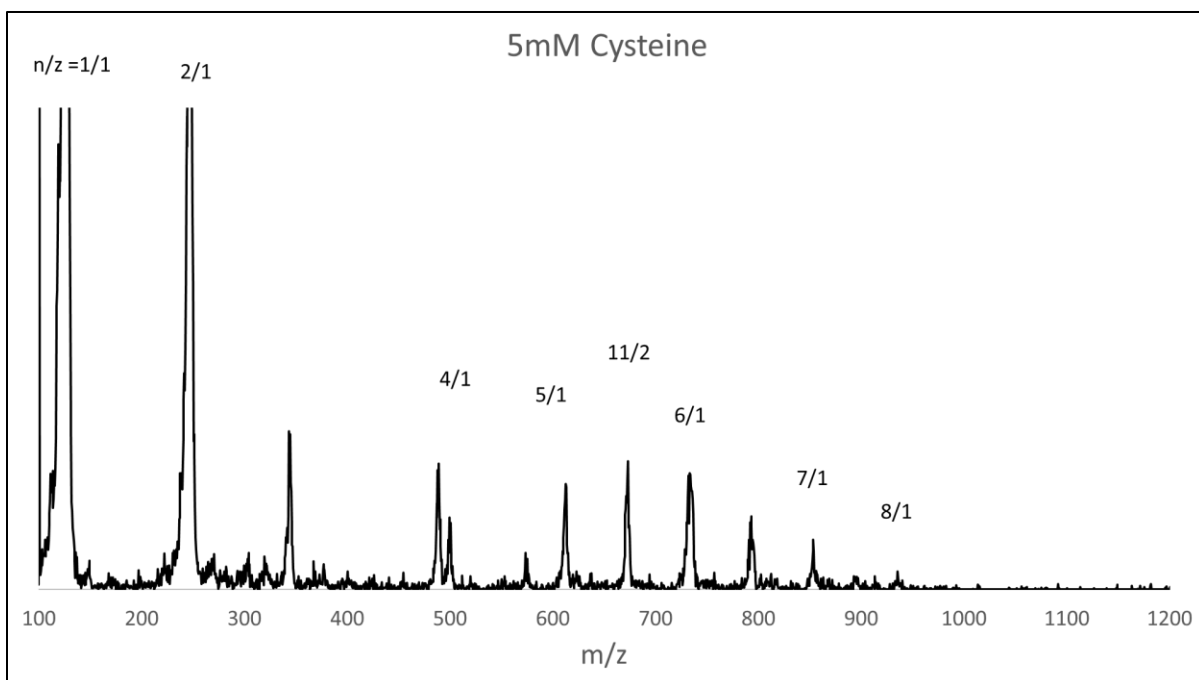


Figure 2.1. Mass spectrum for a solution of 5mM cysteine in water. The x-axis corresponds to the mass-to-charge ratio, while the y-axis represents signal intensity. The peaks are labeled using the corresponding n/z values, where n is the oligomer number and z is the charge. The spectrum has been zoomed in to better show the higher order peaks, as the high-intensity monomer and dimer peaks would otherwise throw off the scaling.

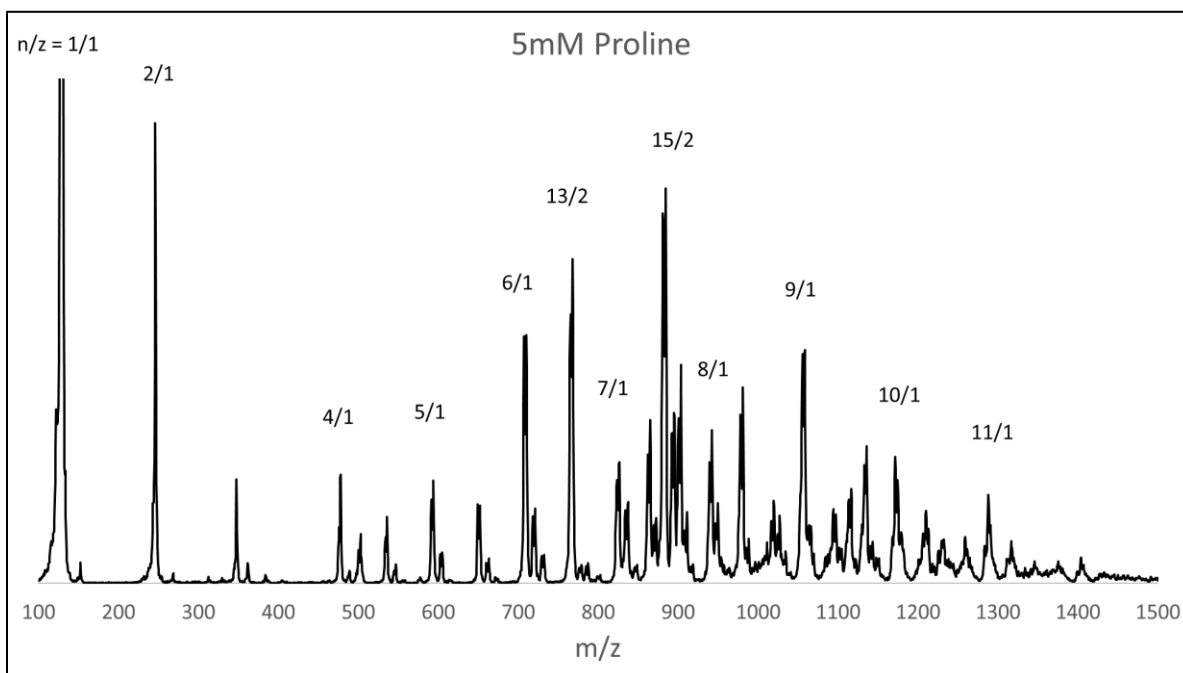


Figure 2.2. Mass spectrum for a solution of 5mM proline in water.

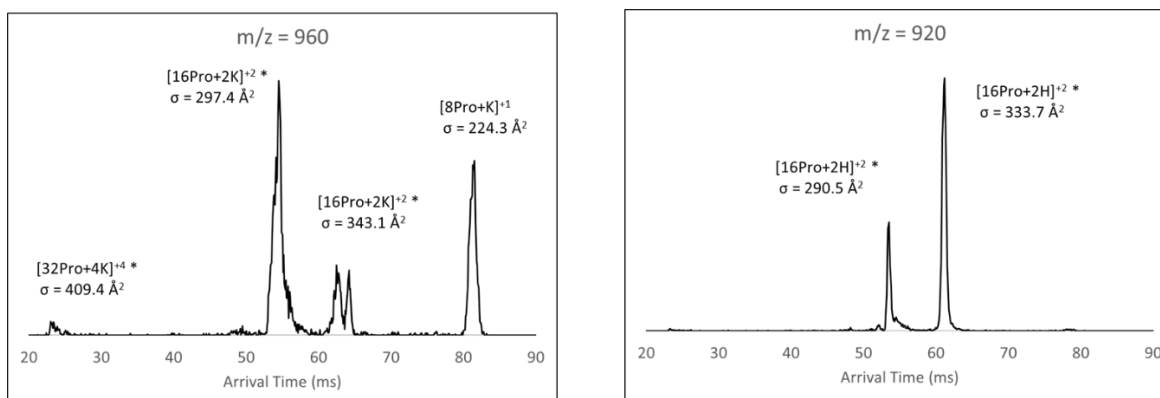


Figure 2.3. An ATD for a solution of proline and potassium, showing the formation of what has been tentatively assigned as a quadruply-charged 32mer potassium heteroligomer (left), and the corresponding homoligomer equivalent (right). While unlabeled, a small signal can be seen in the pure oligomer ATD at the same location, indicating that this oligomer is not formed solely due to the presence of the metal ion. Higher-order clusters such as those shown above are seen to form for both pure and metal-complexed proline, in larger numbers than is observed for cysteine. Similar behavior is seen for all four metal cations, not just potassium.

Proline and Lithium

Upon the addition of lithium ions to a solution of aqueous proline, a similar distribution was observed in comparison to pure proline. Peaks up to the singly charged octamer $[8\text{Pro}+\text{Li}]^{+1}$ were present in the mass spectrum, with the most prominent higher-order species being the singly charged hexamer $[6\text{Pro}+\text{Li}]^{+1}$ (Figure A.5, see Appendix). However, when it came to arrival time distribution (ATD) plots, the signal strength of the largest heteroligomer peaks decreased to the point where they could no longer be observed when the corresponding mass values were selected for. In contrast, pure proline octamer and hexamer peaks were observable in the ATDs, as were smaller heteroligomers such as the tetramer $[4\text{Pro}+\text{Li}]^{+1}$ and dimer $[2\text{Pro}+\text{Li}]^{+1}$ clusters.

Potentially due to the small size of the lithium cation relative to the proline clusters, the larger proline-lithium heteroligomers appear to be more unstable than either the

corresponding pure proline oligomers or the smaller heteroligomers. The one exception to this trend is the singly charged trimer species, which appears prominently in the mass spectrum but subsequently cannot be seen in ATDs.

Proline and Sodium

For an aqueous solution of proline and sodium, peaks up to the nonamer were observed to form in the mass spectrum. As was previously seen by von Helden et al., a prominent tetramer peak was present in the mass spectrum for the proline-sodium solution in comparison to both the pure proline solution and the proline-lithium solution, indicating that the presence of sodium cations encourages tetramer formation.

ATDs were successfully collected for all proline-sodium peaks, even for the largest heteroligomers such as the octamer $[8\text{Pro}+\text{Na}]^{+1}$ and nonamer $[9\text{Pro}+\text{Na}]^{+1}$. Unlike lithium, the presence of sodium seems to further stabilize, rather than destabilize, the formation of larger heteroligomers.

Proline forms an assortment of oligomers both in isolation and in the presence of sodium. However, when taking all of the ATDs into consideration, the relative intensities of the dimer, tetramer and octamer heteroligomer peaks at each specific mass are greater than the intensities of the corresponding pure proline peaks at their specific mass. The addition of sodium promotes an oligomerization pathway by which proline monomers form dimers, these dimers form tetramers and subsequently go on to form octamers.

Both singly and doubly charged octamers ($[8\text{Pro}+\text{Na}]^{+1}$ and $[8\text{Pro}+2\text{Na}]^{+2}$) are present in solution. In the proline-sodium solution, of the two species with a mass of 464 amu

(corresponding to the pure oligomer), the doubly charged octamer and the singly charged tetramer have similar intensities. However, at a mass of 483 amu (corresponding to the proline-sodium heteroligomer), the proportion of the doubly charged octamer is much higher compared to the singly charged tetramer. Sodium may act as a stabilizer in the formation of the $[8\text{Pro}+2\text{Na}]^{+2}$ cluster.

Proline and Potassium

From the mass spectrum recorded from the solution of potassium and proline, heteroligomers were observed up to at least the hexamer. Proline-potassium heteroligomers were observed up to the octamer in the collected ATDs. The hexamer and tetramer ($[6\text{Pro}+\text{K}]^{+1}$ and $[4\text{Pro}+\text{K}]^{+1}$) both predominated over their pure proline equivalent, based on the relative intensities of the ATD peaks.

Proline and Calcium

The first three metal cations used in this experiment are all alkali metals (Li^+ , Na^+ and K^+). Calcium is the only alkaline earth metal used, and is therefore the only one of the tested metal ions with a charge of +2. The peak distribution observed in the proline-calcium mass spectrum is noticeably different from the peak distribution seen for the other three metals, which differ slightly in composition from each other but follow the same general layout. However, this is an expected deviation. While calcium has a similar atomic radius to potassium, the increased charge will of course shift the peak distribution towards lower mass-to-charge ratios. Despite this, many of the same oligomers appear for calcium as for the other

three metal ions, having the same size but an extra charge. Based on initial peak assignments, oligomers past the doubly charged decamer are visible in the mass spectrum. On the other hand, while peaks corresponding to these proline-calcium heteroligomers do appear to be present in the mass spectrum, these clusters cannot be observed through subsequent mass selection. These clusters appear initially but, due to their high charge, may lack the stability needed to persist in the drift tube. Of the oligomers for which ATDs could be recorded, the most prominent heteroligomer was the hexamer $[6\text{Pro}+\text{Ca}]^{+2}$.

Proline is also capable of forming a singly charged calcium heteroligomer, provisionally assigned as $[n\text{Pro}+\text{Ca}-\text{H}]^{+1}$ for multiple values of n , where n is the oligomer number. While the existence of this cluster is certainly implausible compared to other proline/metal heteroligomers, this particular assignment fits best with the experimental evidence. This will be elaborated on in more detail for cysteine/calcium below, as the same phenomenon occurs there as well.

Cysteine Oligomerization

In comparison with proline, which has been observed to naturally form numerous high-order oligomers in solution, cysteine has been somewhat neglected due to its apparent lack of large oligomer formation, the notable exception being a “magic cluster” hexamer $[6\text{Cys}]^{+1}$. The mass spectrum corresponding to a solution of pure cysteine lacks the characteristic raised baseline associated with amino acids which aggregate to a greater extent, such as phenylalanine and proline (Figure 2.1). However, I have found in previous experimentation that pure cysteine does have the potential to form oligomers above the

hexamer, even extending up to the decamer $[10\text{Cys}]^{+1}$, though the hexamer is still the predominant peak. Cysteine can also form a doubly charged dodecamer, $[12\text{Cys}]^{+2}$, which is included in the same mass-to-charge peak as the magic cluster hexamer in the mass spectrum (Figure 2.1). Lastly, while this is observed to a lesser extent in comparison to proline, cysteine can also form potentially toxic oligomers that are even larger than the dodecamer, up to even an 18mer (Figure 2.4). The main topic of inquiry is whether the addition of metal cations might further encourage large oligomer formation to a similar extent as proline.

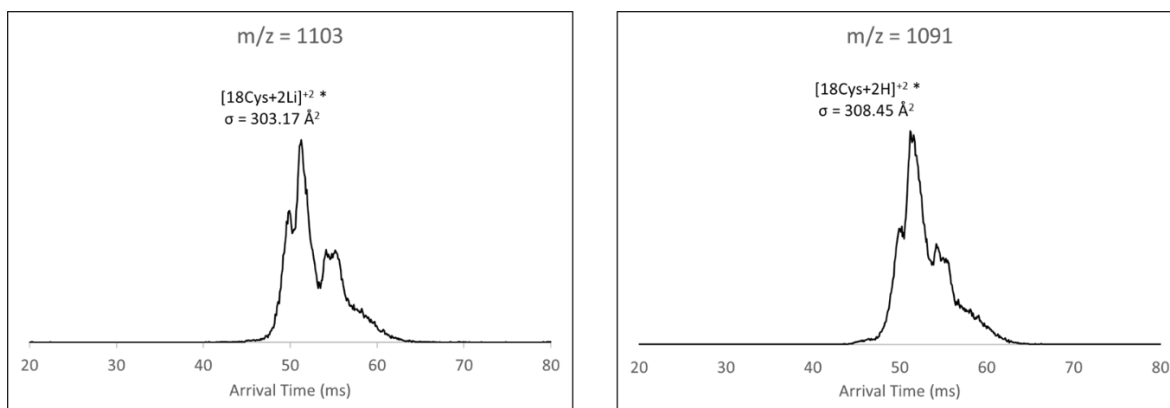


Figure 2.4. An ATD for a solution of cysteine and lithium, showing the formation of a doubly-charged 18mer lithium heteroligomer (left) and the corresponding homoligomer equivalent (right). While cysteine mostly forms clusters up to the dodecamer, higher-order clusters such as the above are seen to form for both pure and metal-complexed cysteine. Similar behavior is seen not just for lithium, but for sodium, potassium and calcium as well.

Cysteine and Lithium

Based on the ATDs, the most prominent cysteine-lithium heteroligomers were the hexamer $[6\text{Cys}+\text{Li}]^{+1}$ and the nonamer $[9\text{Cys}+\text{Li}]^{+1}$. While pure cysteine tetramer and octamer peaks were observed through mass selection, the corresponding heteroligomers were

not, suggesting that a tetramer/octamer oligomerization path may occur for pure cysteine, but is destabilized when lithium is added to the solution – or at least that cysteine prefers to undertake the hexamer/nonamer oligomerization path when complexed with lithium. One possible explanation is that the nonamer peak is predominantly the doubly charged 18mer rather than the singly charged nonamer, and thus results from hexamer stacking with little to no trimer presence. The process to confirm this hypothesis through simulated annealing is currently ongoing.

Cysteine and Sodium

When sodium ions were added to a solution of cysteine, clusters up to the hexamer were observed in the mass spectrum, although at low intensities (Figure A.2). In comparison to proline, where the addition of sodium stabilized large cluster formation, cysteine formed fewer high order clusters, and even the magic cluster hexamer was not clearly visible. However, the species up to and including the hexamer were still observable through mass selection with the quadrupole. While this may be due to a loss of signal in the larger mass regions of the spectrum due to the mass cutoff described earlier in this section, it is more likely due to the presence of sodium potentially slowing down the formation pathway for larger oligomers, since the hexamer peak was still visible in subsequent ATDs. Alternately, the pathway by which sodium encourages high-order oligomer formation in proline may have been proline-specific and did not function in the same way for cysteine.

In the mass spectrum, the presence of a significant peak at a mass corresponding to a higher-charge oligomer, $[9\text{Cys}+\text{Na}]^{+2}$, may indicate that sodium aids the aggregation of

cysteine via a trimer-stacking mechanism to form larger clusters up to the dodecamer. This would be similar to what was observed for proline, where sodium promotes aggregation of both trimers and tetramers through different pathways. The broad dodecamer peak seen in both of the hexamer ATDs indicates the presence of multiple isomers, further supporting the existence of several different aggregatory pathways. However, when comparing the pure cysteine hexamer ATD to that for the corresponding cysteine-sodium heteroligomer, the ratio of the singly charged hexamer $[6\text{Cys}+\text{Na}]^{+1}$ to the doubly charged dodecamer $[12\text{Cys}+\text{Na}]^{+2}$ actually increases. Therefore it appears more likely that sodium disrupts dodecamer formation and stabilizes the hexamer instead.

Cysteine and Potassium

Cysteine and potassium form clusters up to the decamer, as can be seen in the mass spectrum. Both heteroligomers and homoligomers are observable (Figure A.3). The lack of a prominent dimer peak in the mass spectrum combined with the presence of high-intensity peaks corresponding to larger masses would seem to suggest that potassium promotes aggregation starting from trimers and tetramers with the addition of monomers, rather than dimer stacking. However, when examining the ATDs, the intensity of both the dimer heteroligomer and homoligomer is quite high in comparison to the monomer.

From the ATDs, heteroligomers were observed up to the octamer. This matches with the predicted effect of potassium, which had been previously observed to promote an octamer aggregation pathway. However, no tetramer heteroligomers were observed upon mass

selection. This could mean that the predominant aggregation pathway does in fact involve dimer stacking to form hexamers as an intermediate stage of the octamer pathway.

Cysteine and Calcium

For the previous three metal ions, with some minor variations in intensity and distribution, cysteine forms a similar variety of clusters between all three. With calcium, too, cysteine forms analogous oligomers, the only difference being a +2 charge instead of +1. However, there is one apparent phenomenon which distinguishes the cysteine/calcium solution from all others: the existence of a $[n\text{Cys}+\text{Ca}-\text{H}]^{+1}$ cluster.

The exact assignment of this cluster is subject to change. This is not a commonly occurring structure. It does not appear for proline/calcium, and by its nature would not be expected to appear for any singly charged metal ion. The formation of the structure itself is also implausible: somehow, a hydrogen must be ejected from the oligomer and the resulting imbalance of charge stabilized by the presence of the calcium ion. However, cysteine/calcium clusters corresponding to the exact mass-to-charge ratio for a given oligomer number are observed both in the mass spectrum and in ATDs, and at multiple different oligomer numbers: $n = 4, 6$, and so on. Additionally, the relative ATD peak intensities indicate that they exist in large numbers. Clearly, some species is present in solution that matches the criteria for this $[n\text{Cys}+\text{Ca}-\text{H}]^{+1}$ oligomer, and the math just doesn't work if we try to set up the calculations assuming a more conventional structure such as $[n\text{Cys}+\text{Ca}]^{+2}$. The most salient piece of evidence for the existence of this type of cluster comes from examining arrival times in the ATDs, using the cysteine/calcium tetramer as an

example (Figure 2.5). The arrival time for the peak estimated to correspond to the $[4\text{Cys}+\text{Ca}-\text{H}]^{+1}$ oligomer is around 47.3 ms. The arrival time for the pure cysteine tetramer, $[4\text{Cys}+\text{H}]^{+1}$ is around 47.0 ms. Using the Mason-Schamp equation (Eq. 4), the experimental collision cross sections are estimated to be 141.8 \AA^2 and 139.8 \AA^2 , respectively. The addition of the calcium is evidently offset by the decreased charge, resulting in the structure behaving similarly to the pure cysteine tetramer. Identical results are observed for the doubly charged octamer peak. As there is the same general placement of peaks in the ATDs compared to the pure cysteine equivalents, the mystery oligomer must be similar in size.

As a result, I have cautiously concluded that this mystery oligomer is most likely to have the form $[\text{nCys}+\text{Ca}-\text{H}]^{+1}$, which places it at the correct mass-to-charge ratio while also having an arrival time distribution comparable to the pure cysteine cluster $[\text{nCys}+\text{H}]^{+1}$.

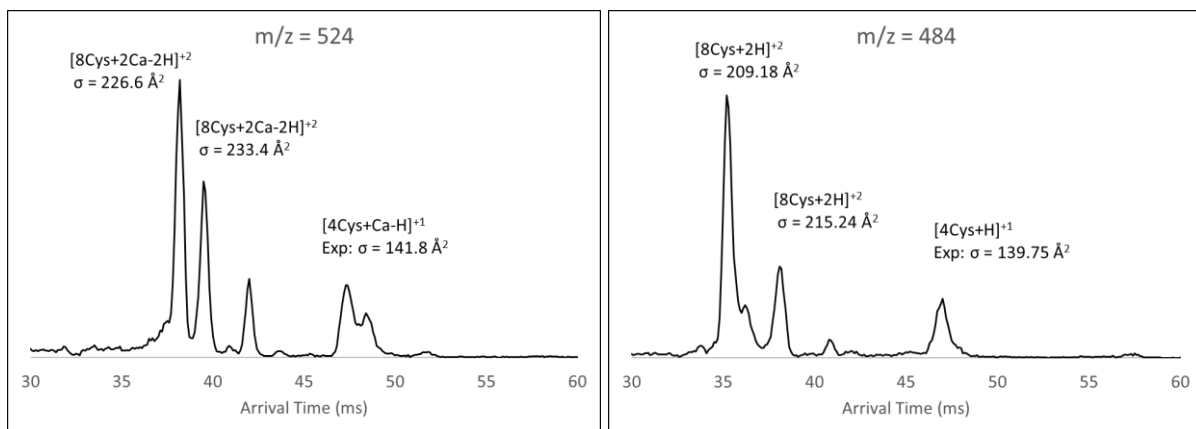


Figure 2.5. A comparison of two ATDs taken at the mass-to-charge ratio corresponding to the cysteine/calcium tetramer (left) and the pure cysteine tetramer (right). In all future ATD comparisons of pure oligomers and heterooligomers such as this, the ATDs were taken from the same solution and recorded on the same day, in this case within minutes of each other.

Conclusions: Calcium Versus Other Metal Ions

The presence of calcium results in a peak distribution that is unlike that seen for all three other metal ions. However, aside from minor fluctuations in peak intensity, this falls into the range of what is expected when accounting for the extra charge on calcium. While proline and calcium behave straightforwardly, one unique occurrence with cysteine is the formation of several $[n\text{Cys}+\text{Ca}-\text{H}]^{+1}$ clusters. To determine the specifics, more structural analysis must be done. This is one area in which further MD simulations can help elucidate possible structural differences which would cause this phenomenon.

Conclusions: Potassium Versus Other Metal Ions

Potassium, when complexing with hexamers, has a small but unique structural effect. For all other metal ions, the singly charged hexamer and doubly charged dodecamer clusters appear roughly in the same spot, although they may have different distributions of the various isomers. However, in the case of potassium, the peaks for both the hexamer and dodecamer are observed in the corresponding ATDS to shift to the right by about 5 ms. Moreover, this shift is seen for both proline and cysteine hexamers. So far, this effect seems to be restricted to the hexamer, but further research into this phenomenon is ongoing.

The occurrence of these potassium heteroligomers at a later time point in comparison to their fellow metal ion counterparts may be due in part to steric effects. Potassium is much larger in comparison to its fellow alkali metals lithium and sodium. Based on its computationally modeled position within a framework of amino acids, its larger size would presumably push the surrounding amino acids further away. Calcium has a similar size but its

charge is +2, double that of potassium. Thus the calcium cation will exert a stronger pull on the amino acid framework around it, which should cause the overall hexamer size to be smaller and more comparable to sodium and lithium. As extended isomers will appear at later time points in the ATD than compact isomers, the experimental placement of these potassium-complexed hexamers lines up with the predicted structures. However, this observed shift is not so extreme that potassium can be conclusively shown to differ from the standard metal ion aggregation protocol.

Final Conclusions

The question posed in this dissertation is twofold. Does the presence of a specific metal ion affect the mechanisms by which cysteine and proline aggregate in solution? And if so, are cysteine and proline affected in the same way, or do they behave differently?

The addition of any metal ion to an aqueous solution of cysteine or proline has one clear and observable effect on the clusters that form. Heteroligomers are clearly present in each sample, either in the mass spectrum or the arrival time distribution plots. If the metal cations truly had no effect on the mechanisms of cluster formation, only homoligomers would exist in solution. However, while there are some observable variations between each set of amino acid/metal ion data, mainly having to do with peak intensity and distribution, each metal ion ultimately forms many of the same oligomers as the others. There is no clear-cut pattern showing that the presence of a specific metal causes a certain species, for example a tetramer or hexamer, to predominate over other species. The question of whether or not certain aggregation pathways are selected for or completely blocked off by the presence of a

specific metal ion cannot be conclusively answered without a wider sample of data to draw from.

In all cases, the four metal ions exhibited similar behavior when complexed with both proline and cysteine. While there were minor fluctuations, no conclusive difference in behavior could be seen. Calcium, as the only ion with a +2 charge, did appear from a surface-level perspective to form different clusters compared to the other three. However, upon further inspection, these differences are no more than what is expected for an ion with a charge greater than one. Specifically, drastic differences are only truly present in the mass spectrum, where the peaks are shifted to the left of the spectrum by the amount predicted due to the mass-to-charge ratio being divided by two. When taking the extra charge into account, it can be seen that calcium does form a similar range of structures compared to the other metals. The one exception to this is cysteine/calcium, which aside from forming the aforementioned species observed in all cases, additionally formed unusual singly-charged calcium heteroligomers which were not observed for any other sample.

In summary, it appears that the presence of metal ions as a whole does not significantly change the aggregation mechanisms undertaken by certain amino acids, not favoring the formation of certain oligomers over others. Additionally, the identity of the metal ion itself does not appear to have a clear-cut effect on aggregation. Accounting for the expected divergences resulting from using a doubly-charged cation such as calcium, any altered behavior is either due to the presence of a metal ion in general or due to the well-established difference in aggregating behaviors previously observed for cysteine and proline, not necessarily because of the influence of any specific metal ion. As little to no change in

oligomer variety is observed experimentally between solutions, what changes do occur are most likely amino acid-specific.

Although their aggregation pathways differ, both cysteine and proline are capable of forming large, potentially toxic oligomers, which matches up with the results predicted by the Generic Amyloid Hypothesis. Future work will be conducted in order to conclusively determine that the identity of a metal ion has no effect on aggregation, for instance expanding the variety of metal cations to include more alkaline earth metals, to determine whether the results seen for calcium follow a consistent pattern. Cytotoxicity assays will also help determine whether the large oligomers formed by cysteine and proline are in fact toxic.

Part 3: A Computational Study of Amino Acid Aggregation

The aggregation of free zwitterionic cysteine and proline was studied both individually and in the presence of select alkali and alkali earth metal cations. The experimental data regarding these oligomeric clusters is expanded on in the previous part of this thesis. This experimental data was supplemented by computational modeling performed using the GROMACS molecular dynamics software package in conjunction with the OPLS-AA force field, both of which are characterized in more detail in the introduction of this dissertation. Here I will elaborate on my methods for obtaining these comparable theoretical structures through the use of simulated annealing.

Experimental Setup

Previous simulated annealing computations done by the Bowers group used the AMBER software package on local computers. One such study was conducted on the small polypeptide bradykinin, which had been extensively experimentally characterized and was of biological relevance due to its role in injury response and inflammation within the body.⁶¹ For any system more complex than the simplest of small metabolites, experimental characterization alone is not sufficient. In these cases, theoretical methods such as molecular dynamics (MD) are used to verify experimental results. However, a polypeptide such as bradykinin can adopt a very large number of conformations, and simply generating a computational structure would run the risk of the molecule ending up in a relatively high-energy local minimum on the potential energy landscape, unable to move to lower energy conformations which might fit more closely with experimental data.

To avoid this problem, the theoretical bradykinin structure was subjected to a simulated annealing protocol, selecting for lower energy conformations and feeding them back into the simulated annealing loop as input structures, and generating a large number of conformers. The large sample size allowed for a more extensive sampling of the potential energy landscape, and the selective nature of the simulated annealing protocol decreased the chances of the molecule becoming trapped in a high-energy local minimum with little correlation to the observed experimental results.⁶² While computational methods were restricted by the limitations of the technology available at the time, the simulated annealing protocol generated by the Bowers group for experiments such as these was successful in providing theoretical bradykinin conformations which correlated closely to the previously obtained experimental data.

Unfortunately, this particular method using AMBER was no longer accessible at the time of this project, due to the required software no longer being available on our current computing platforms. Therefore, using GROMACS, I developed and coded my own procedure to carry out similar molecular dynamics (MD) simulations on amino acid oligomers of interest. The steps of this procedure are detailed below.

Structure Preparation

To begin with, an amino acid oligomer structure was constructed using the Avogadro molecule editor software. Depending on the experiment, at this time either the oligomer was kept in its pure state, or a metal cation was added to the structure. Just as for the experimental analysis, the metal cations used were sodium (Na^+), potassium (K^+), lithium (Li^+) and

calcium (Ca^{2+}). The oligomer structure was uploaded to a Linux system and prepared for simulation through a multi-step process. All steps after this point were done using GROMACS on Linux platforms (on the CNSI cluster servers at UCSB, as well as on local PCs running the Ubuntu Linux software).

First, the cluster was constrained to and centered within a box of variable size. Potentially due to the effect of long-distance intramolecular interactions with itself across the walls of the box, clusters occasionally broke apart and fragmented when subjected to unconstrained motion during an MD simulation. Therefore, in order to give an accurate final structure comparable to what was experimentally observed, the volume of the box was increased as necessary for certain simulations, typically those involving proline or calcium ions which seemed to be less stable in MD. Generally, the box side length ranged from 1-20 nm, with some simulations requiring a box with a side length of 50 nm or higher. The immense computational power required to run MD with a larger box volume limits the maximum side length that can be used.

After being placed within a box, the oligomer structure was then neutralized with chlorine counterions. Cysteine and proline both exist in their zwitterionic “neutral” state at physiological pH: cysteine and proline both contain a deprotonated, negatively charged carboxyl group ($\text{pK1} = 1.71$ for cysteine, $\text{pK1} = 1.99$ for proline) and a protonated, positively charged amine group ($\text{pK2} = 8.33$ for cysteine, $\text{pK2} = 10.60$ for proline). Cysteine has an additional thiol group which is protonated at physiological pH ($\text{pK3} = 10.78$) and ends up being neutral due to the hydrogen canceling out the negative charge on the sulfur. From these dissociation constants, it is simple to confirm that both amino acids will have a net neutral charge when all local charges are totaled up.^{63, 64} Therefore, when constructing the

computational models for this study, each simulated oligomer was built using zwitterionic free amino acids, giving the oligomer itself a net neutral charge. However, in order to counteract any metal cations and neutralize the overall system so that MD could be performed, one to two chlorine counterions were added to the corners of the box, frozen in place some distance away from the oligomer to avoid any potentially destabilizing interactions.

Finally, the system underwent a simple NVT energy minimization to stabilize the structure before MD was carried out.

Simulated Annealing

In order to obtain a sufficient set of theoretical cross sections, each structure was run through a loop consisting of two stages: a simulated annealing step followed by an energy minimization step. In the simulated annealing stage, the energy-minimized initial structure, taken from the last preparation step, was first heated to 800 K and underwent MD for 30 ps. The prolonged time interval at a high temperature increases the possibility for the structure to undergo high-energy conformational shifts, moving the oligomer to regions of the potential energy landscape that would normally be inaccessible behind high energy barriers.

After this initial high-temperature step, the system temperature was decreased in intervals of 100 K down to 0 K. At each new temperature, the structure underwent MD for 5 ps. Once the simulated annealing sequence had completed, the structure underwent an energy minimization at 0 K, which lasted approximately 15 ps. The duration of the full simulated annealing process was 80 ps in total, to allow sufficient time for the system to restabilize

between each decrease in temperature. The final output structure was saved to a separate location and then fed back into the loop as the input structure of the next iteration. In total, 101 energy minimized structures (runs 0-100) were collected for each specific oligomer (Figure 3.1).

In the process of developing this algorithm I noticed that after the first run, every subsequent output structure had a similar structure to the initial output structure (from run 0), with only minor fluctuations in potential energy. This suggests that once the oligomer moves from its initial high-energy conformation into a local energy minimum, the final structure is too stable and it is not able to climb out of the potential energy well even at 800 K. To improve structure diversity and increase the chances of reaching lower energy conformations, the 101 runs were split up into 4 sets of 25 each (26 for the last run). Run 0 began from the initial input structure, feeding the resulting output structure back into the loop as the input structure for Run 1, and continuing the loop normally until Run 24. However, Run 25 was started from the initial input structure just as Run 0 had been, continuing the same loop until Run 49, before being reset to the initial starting structure for Run 50, and then again for Run 75 (Figure 3.2).

This allowed the oligomer to explore multiple local minima on the energy landscape, with a better chance of approaching the global minimum, while still providing enough of an opportunity for the oligomer to actually reach the local minimum within each set of runs.

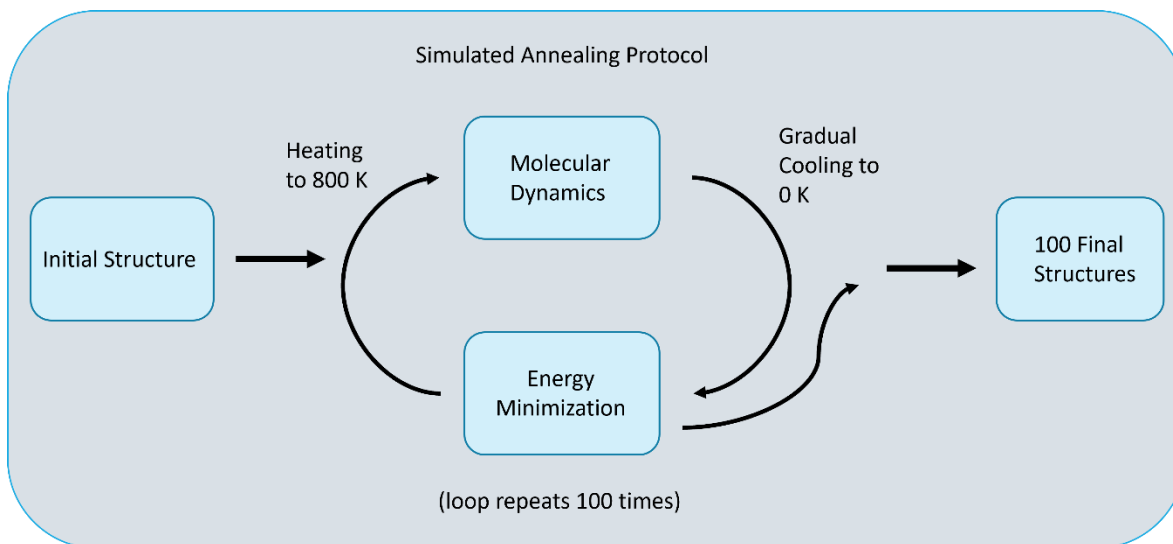


Figure 3.1: Flow chart depiction of the typical simulated annealing procedure, where the output structure of the previous step is used as the input structure of the next step.

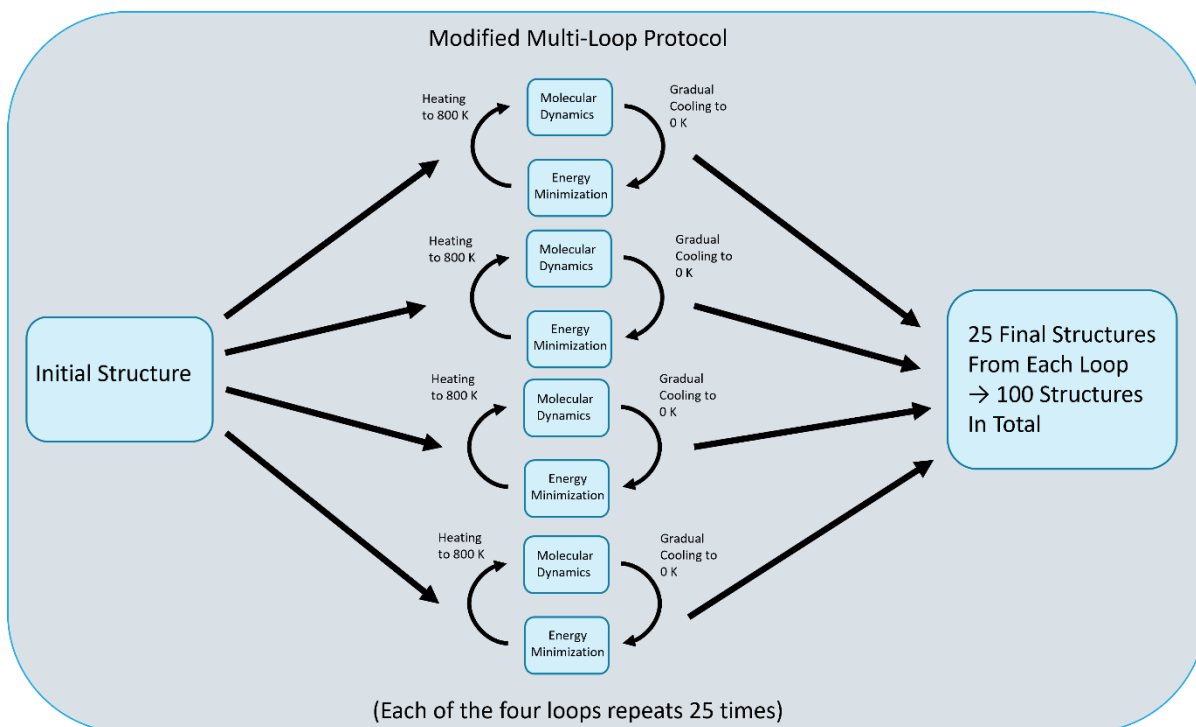


Figure 3.2: Flow chart depiction of my modified simulated annealing procedure, where progression was reset after each set of 25 runs.

Data Processing: PSA and potential energy scatter plot

Once the complete set of energy minimized output structures were obtained for each oligomer, the structures were edited to remove the counterions and fix some small file formatting issues from the simulated annealing process. Then, the prepared files were run through the Projection Superposition Approximation (PSA) software to calculate the theoretical collision cross sections for each structure. The parameters used to calculate the theoretical cross sections were chosen to match the experimental data: helium buffer gas, the corresponding drift cell temperature, and so on. From the resulting calculations, the PA and PSA cross sections were obtained.

Finally, the potential energies for each of the hundred plus output structures were graphed against their corresponding cross sections. From the resulting scatter plot, we were able to derive a general idea of how a particular oligomer's conformations were distributed across the energy landscape. Lower (more negative) potential energies correspond to more stable structures. With this scatter plot, the most probable theoretical conformations could be found and compared to the experimental cross section to identify which conformations actually formed under experimental conditions.

Proline and Cysteine Computations

Through the molecular modeling and computational analysis protocol described above, the molecular conformations of various oligomers of interest were constructed and subjected to molecular dynamics (MD), and their collision cross sections and energies determined. The most salient details from these computational experiments will be described

below, and all other data will be included in the appendix at the end of this dissertation. All structures were imaged using the Avogadro software, and relevant trajectories were examined using the VMD software.

One challenge when using GROMACS, and the simulated annealing algorithm used for MD, is that large and highly charged structures tend to rapidly fall apart under the extreme conditions of the initial high-temperature MD step. Additionally, the more charges a structure possesses, the more counterions must be added to maintain a neutral charge. However, these counterions can cause structural disruptions when in close proximity to the main structure. Careful maintenance is required to prevent this from occurring, and a balance must be struck between enlarging the box to accommodate more counterions and the exponential increase in computation time resulting from this enlargement. As a result, when using theoretical collision cross section data for peak assignment corresponding to large oligomers with multiple charges, in certain cases the cross section for the singly charged equivalent is used instead. If the peak assignment is correct, the assumed value will be within a reasonable range of the actual cross section.

All theoretical collision cross sections reported below are average values calculated from the hundred individual end structures for each oligomer which were obtained through the simulated annealing/energy minimization process. The individual cross section values have been presented below in a scatter plot of potential energy versus cross section for each oligomer.

Since this method merely samples the energy landscape based on the random SA algorithm, it is not guaranteed that every possible isomer that forms in will be observed in the scatter plot. For future work, the next step is to increase the probability of observing all

possible conformations using the SA method by improving the balance between the number of loops while still confirming that each loop successfully reaches equilibrium before it is reset.

Proline

We begin once again with proline. The experimental data discussed in the previous section of this dissertation provides details on which specific proline oligomers were observed in solution. An effort was made to model as many of these clusters as possible in order to determine pertinent structural information. Relevant oligomers of pure proline were modeled up to the dodecamer, and proline-metal oligomers were modeled up to the hexamer in most cases.

Through the use of the PSA method, the proline monomer was found to have an average theoretical collision cross section of 52.9 \AA^2 . Most of its output structures fell within the same region of the potential energy surface, as expected due to the limited degrees of freedom available for a single proline (Figure 3.3). Larger clusters, such as the pure proline hexamer, had more variation in conformation and energy, as can be seen within the scatter plot of potential energy versus cross section (Figure 3.4).

Present in the proline hexamer scatter plot are two distinct regions corresponding to at least two different isomers. These isomers are similar in cross section, and would most likely overlap in an ATD, but can be separated out into two conformations based on their differing potential energies (Figure 3.5). More stable structures have a more negative potential energy, but the most stable structure found through theoretical methods will not necessarily be the most common oligomer present in a real sample. In a comprehensive sampling of the entire

potential energy landscape, the oligomer observed most in MD should theoretically be the most common structure found in solution, and should also appear most prominently in the experimental spectrum.

In contrast with cysteine, which was a more well-behaved system, modeled proline clusters were frequently unstable and disintegrated under the high-temperature conditions of the simulated annealing algorithm. This occurred for pure proline clusters to a certain extent, but was especially true for proline-metal ion heterooligomers, which hindered the attempt to collect collision cross section data for such clusters. This may have been due to the inherent instability of proline oligomers in comparison to cysteine oligomers, as seen in their respective potential energies, especially with the addition of a possibly structurally destabilizing metal ion. Efforts to overcome this obstacle and determine more of these theoretical cross sections are currently ongoing.

Several averaged collision cross sections for proline, as well as their theoretical equivalents when available, are listed in the table below (Table 3.1).

Cysteine

Cysteine oligomers were modeled using the same SA method as for proline. The cysteine monomer was found to have an average theoretical cross section of 49.5 \AA^2 , a similar value to the proline monomer. Several averaged collision cross sections for cysteine, as well as their theoretical equivalents when available, are listed in the table below (Table 3.2).

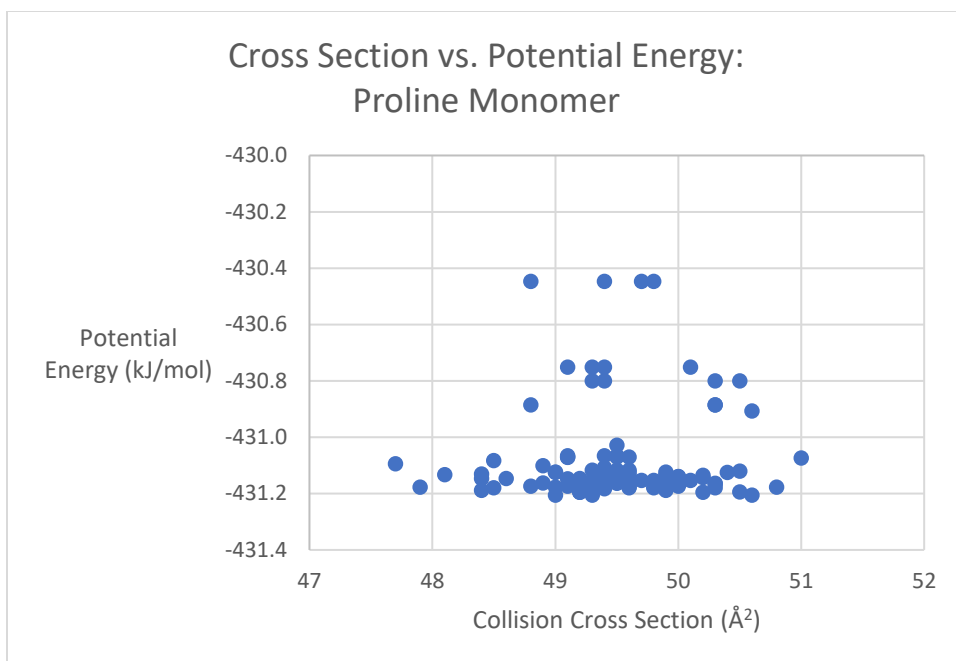


Figure 3.3. A scatter plot of potential energy versus CCS for the pure proline monomer, created using the hundred theoretical output structures from the simulated annealing process.

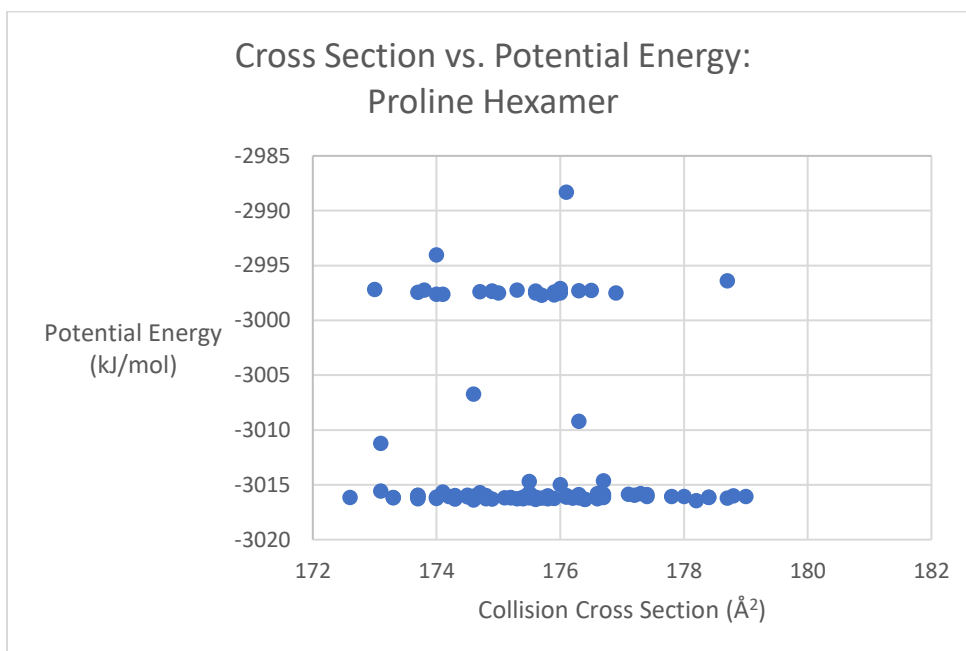


Figure 3.4. A scatter plot of potential energy versus CCS for the pure proline hexamer, created using the hundred theoretical output structures from the simulated annealing process.

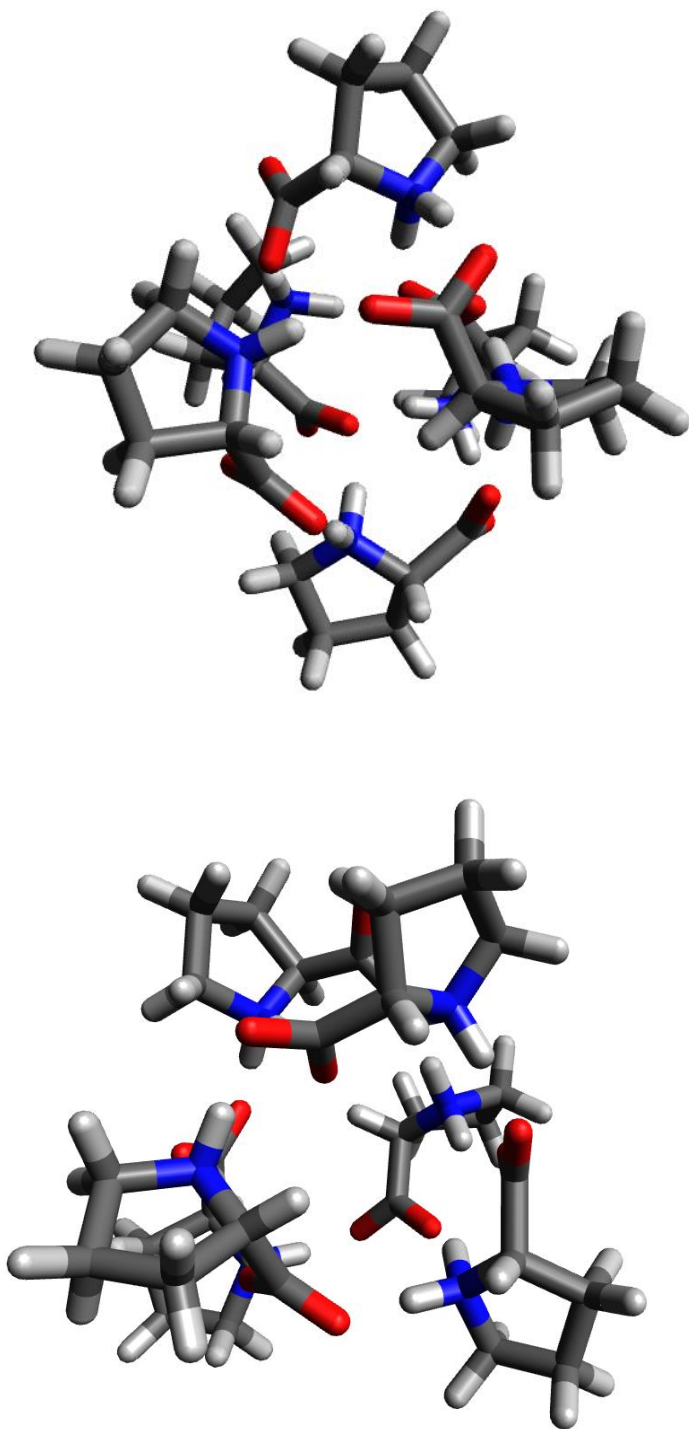


Figure 3.5. Structures for two of the distinct isomers formed by the proline hexamer during MD. Above is the higher-energy, more ordered hexamer structure which takes on a roughly octahedral configuration. Below is the lower-energy and more compact hexamer structure. Its disordered structure allows it to form a greater number of intermolecular interactions, which contributes to its increased stability in comparison to the octahedral structure.

Table 3.1. A list of proline oligomers, their experimental collision cross sections determined using IM-MS on a solution of 5mM proline in water, and the corresponding theoretical cross sections calculated using MD simulations and the PSA method.

Proline Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Pro}+\text{H}]^{+1}$	56.0	52.9
$[2\text{Pro}+\text{H}]^{+1}$	77.0	85.0
$[4\text{Pro}+\text{H}]^{+1}$	142.6	137.0
$[6\text{Pro}+\text{H}]^{+1}$	-	175.7
$[8\text{Pro}+\text{H}]^{+1}$	167.02	-
$[12\text{Pro}+2\text{H}]^{+2}$	285.7	-

Table 3.2. A list of cysteine oligomers, their experimental collision cross sections determined using IM-MS on a solution of 5mM cysteine in water, and the corresponding theoretical cross sections calculated using MD simulations and the PSA method.

Cysteine Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Cys}+\text{H}]^{+1}$	56.15	49.5
$[2\text{Cys}+\text{H}]^{+1}$	79.03	77.4
$[4\text{Cys}+\text{H}]^{+1}$	138.39	125.0
$[6\text{Cys}+\text{H}]^{+1}$	162.04	161.2
$[8\text{Cys}+\text{H}]^{+1}$	184.71	197.7
$[12\text{Cys}+2\text{H}]^{+2}$	262.01	261.8 ‘

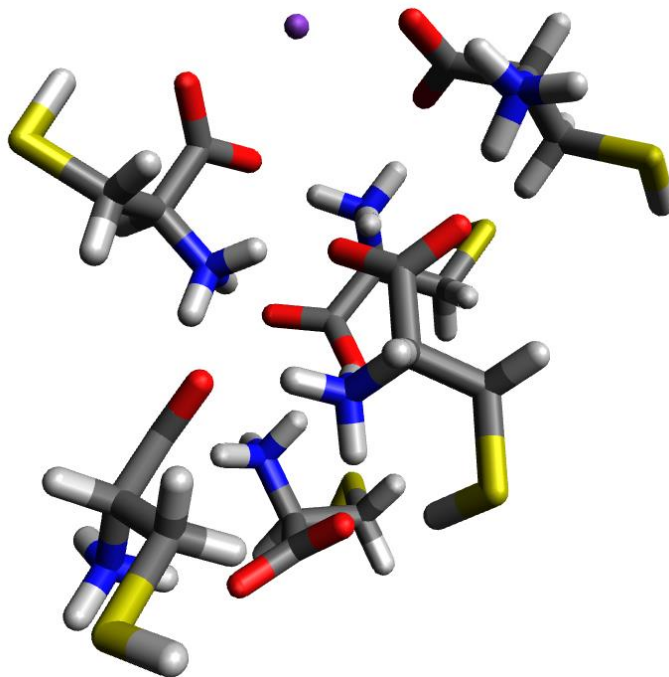


Figure 3.6. Representative output structure of the cysteine-potassium hexamer from the simulated annealing process, modeled using Avogadro. The cysteines are arranged in a disordered manner below the potassium ion, and are held together through a combination of hydrogen bonding and disulfide bonding.

Calcium Versus Other Metal Ions

The process of running simulations on cysteine and proline clusters complexed with calcium is still ongoing. Nonetheless, from the data at hand, some preliminary conclusions can be made. As an example, for the pure cysteine tetramer, the average potential energy over the 100 output structures is -2472.56 kJ/mol. For the cysteine/sodium tetramer, the average potential energy is -2762.25 kJ/mol. And for the cysteine/calcium tetramer, the average potential energy is -3470.02 kJ/mol, a noticeable decrease from either the pure or the sodium-complexed cysteine tetramer. The +2 charge on the calcium cation clearly stabilizes the resulting heteroligomer by a huge margin in comparison to the pure oligomer (Figure

3.7). Additionally, we can compare the proline/calcium tetramer, which has an average potential energy of -3207.24 kJ/mol. This value is significantly lower in comparison to the pure proline tetramer, with an average potential energy of -1933.6 kJ/mol: a difference of over 1000 kJ/mol.

This increased stability may explain some of the strange behavior exhibited by the two amino acids complexing with calcium: specifically, the formation of what have been tentatively assigned as singly charged calcium heteroligomers, the $[n\text{Cys}+\text{Ca}-\text{H}]^{+1}$ and $[n\text{Pro}+\text{Ca}-\text{H}]^{+1}$ clusters. An introduction to these unusual singly charged clusters as well as my reasoning for having assigned them as such have both been provided earlier, in the experimental data section of this dissertation. The existence of this specific class of heteroligomer is the main difference between oligomerization with calcium and oligomerization with the other three alkali metals. Because the cysteine/calcium tetramer is significantly more stable in comparison to the pure cysteine tetramer, this lends credence to the idea that calcium can somehow eject a hydrogen from one of the cysteine zwitterions in order to replace it and form a singly charged oligomer, even though this would normally be an implausible sequence of events. Similar results can be seen for proline/calcium oligomers.

A more focused analysis of this phenomenon will be carried out in the future with regards to proline/calcium oligomers, as experimentally these clusters are doubly charged and more analogous to proline/cysteine clusters complexed with the other metal ions.

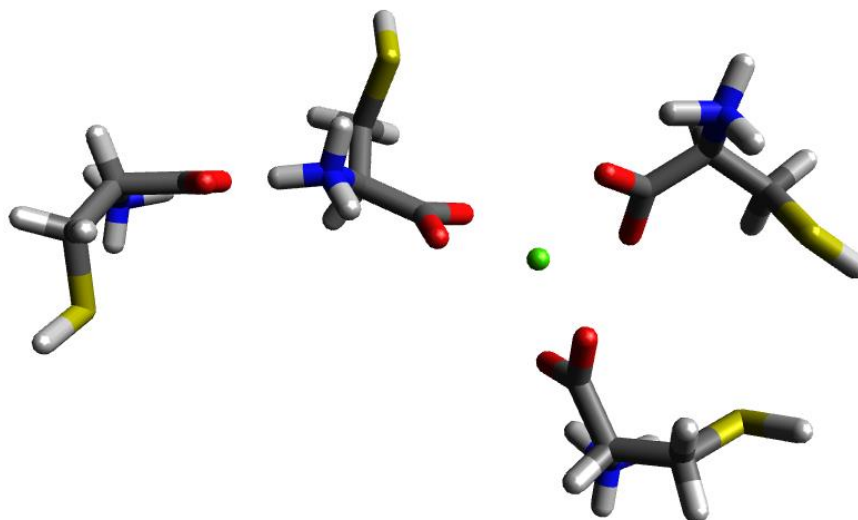


Figure 3.7. A representative structure of the cysteine-calcium tetramer. Based on the structure, the mechanism of tetramer formation – or at least, the mechanism for one potential tetramer pathway – would appear to involve the addition of a monomer to a preexisting trimer.

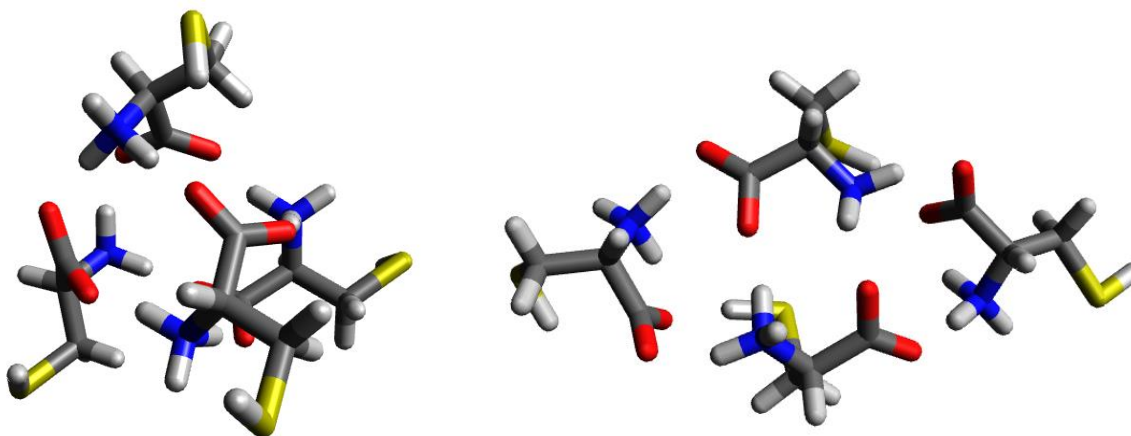


Figure 3.8. Two types of structures observed in the potential energy scatter plot for the pure cysteine tetramer: a roughly tetrahedral structure and a more planar, slightly less stable structure. Neither resemble the structure formed above for the cysteine-calcium tetramer, although the addition of a calcium to the planar structure could cause it to take on a similar configuration.

Potassium Versus Other Metal Ions

As elaborated on in a previous section of this dissertation, in the experimental data collected with regards to the hexamer, both cysteine and proline were found to form clusters with potassium, and these clusters took on a more extended structure relative to the other metal ion-amino acid hexamers. This trend had to be studied through MD as well in order to be fully confirmed.

As an example, the cysteine-potassium hexamer was compared to the pure cysteine hexamer as well as the cysteine-sodium hexamer. All three sets of structures had similar average potential energies, in the -3850 to -4150 kJ/mol range. In comparison to the much more drastic energy differences seen elsewhere, these values are relatively close. The pure cysteine hexamer logically had the smallest cross section distribution, with an average around 161.2 \AA^2 . Next was the cysteine/sodium hexamer, around 171.6 \AA^2 . Finally, the cysteine/potassium hexamer was around 175.3 \AA^2 . The latter oligomer does indeed have a larger average cross section, as expected due to its delayed arrival times, but further simulations must be conducted to compare the remaining metal ions with the current results. If potassium does have an unusual effect on amino acid aggregation beyond what is caused by the other metal cations simply being present in solution, then this effect appears to be more subtle than that of calcium.

Conclusions: Cysteine versus Proline

Through examination of the scatter plots for both the cysteine and proline pure hexamer, the disparity in potential energy between the two can be clearly seen (Figure 3.9). The cysteine hexamer has an average potential energy of -3855.79 kJ/mol. In comparison, proline has an average potential energy of -3077.95 kJ/mol. Not only does the proline hexamer have a higher potential energy, the difference between the two values is approximately 777.84 kJ/mol – a significant energy decrease from proline to cysteine.

This difference in energy and consequently stability can be most likely attributed to the difference in side chains for cysteine and proline. While cysteine has a relatively unobtrusive thiol group which can form stabilizing disulfide linkages with neighboring cysteines, proline possesses a bulky 5-membered ring, which limits the possible movements it can undergo and can result in steric clashes with other prolines in close proximity. This trend is supported by a comparison of other oligomers in each sample, with cysteine oligomers consistently having lower energies than their proline counterparts – the comparison above of pure proline and proline/calcium tetramers with pure cysteine and cysteine/calcium tetramers is one such example.

When reviewing the collected experimental data, it was a common occurrence for certain proline clusters to be present in the mass spectrum but not in subsequent ATDs. In contrast, cysteine clusters were less numerous but remained observable throughout the process of recording ATDs for each mass-to-charge ratio, discounting the usual drop in signal for larger species with the passage of time. From the data, cysteine appears to form less of a variety of clusters than proline, but the clusters it does form are more stable. This is backed up by the potential energy data, where cysteine clusters across the board have a lower

energy in comparison to their proline counterparts. This is observed both for pure cysteine/proline oligomers, and for metal ion heteroligomers as well. Taken in conjunction with the results derived experimentally, this is another piece of evidence supporting the conclusion that the identity of the amino acid has greater relevance to the oligomer formation pathway than the presence of a metal ion or lack thereof, or the specific identity of that metal ion. This is true for all cysteine clusters in comparison with their proline equivalent, not just the hexamer.

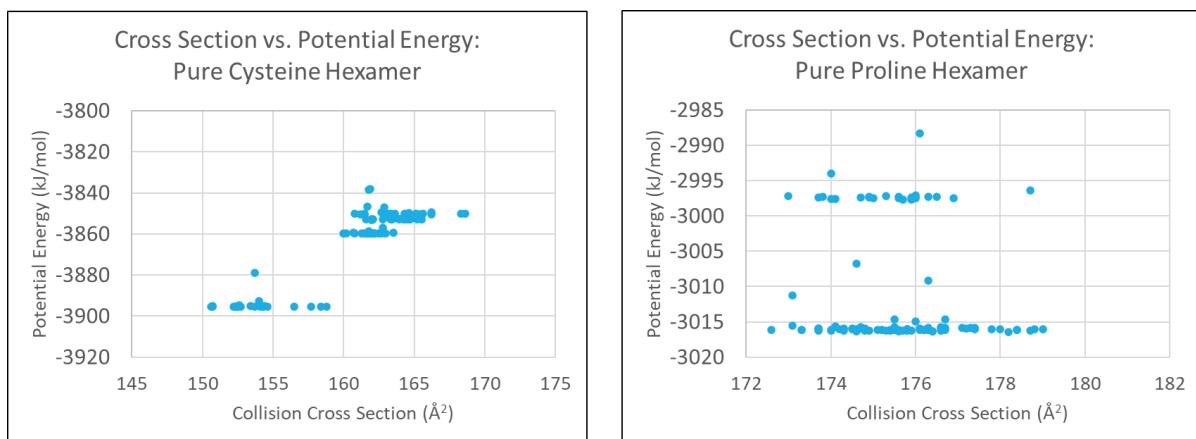


Figure 3.9. A comparison of two scatter plots of potential energy versus collision cross section for the pure cysteine hexamer and pure proline hexamer. All isomers of the cysteine hexamer are located at a lower potential energy than the proline hexamer isomers.

Conclusions: Pure versus Metal-Complexed Oligomers

In addition to the identity of the amino acid involved in aggregation, the addition of a metal to the sample will affect the stability of the resulting cluster. This can be observed in scatter plots of the pure cysteine hexamer compared to the cysteine-lithium heterologomer equivalent (Figure 3.10). While the pure cysteine hexamer has a range of -3820 to -3900 kJ/mol, the lithium heterologomer has a range of -4270 to -4320 kJ/mol, which is comparatively more negative and indicates greater stability. This increased stabilization of the heterologomer compared to the homologomer is observed for both cysteine and proline, for all oligomers aside from the hexamer, and for all four metal ions regardless of identity.

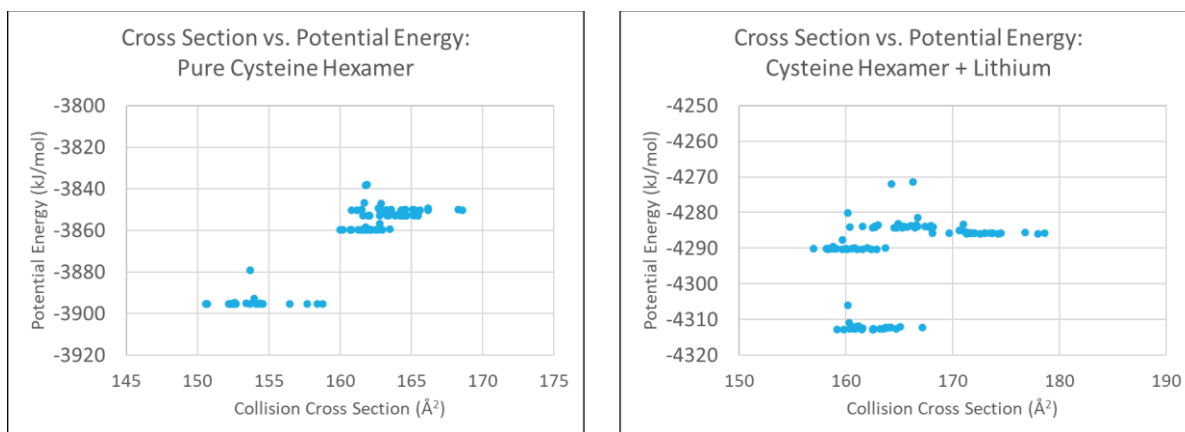


Figure 3.10. A comparison of two scatter plots of potential energy versus collision cross section for the pure cysteine hexamer in comparison with the lithium-complexed cysteine hexamer. The heterologomer has a more negative potential energy range compared to the homologomer.

Conclusions: Comparison of All Four Metal Ions

As previously stated above, the addition of any metal cation will increase the stability of the resulting heteroligomer in comparison to the pure oligomer. However, the identity of the metal used does make a difference in oligomer stability, as can be seen in the scatter plots (Figure 3.11).

Three of the metal ions have a +1 charge: lithium, sodium, and potassium. Of these, lithium has the most negative potential energy range, between -4250 and -4330 kJ/mol. Sodium has a range of -4140 to -4180 kJ/mol. Finally, potassium has a range of -4030 to -4090 kJ/mol. Since potential energy and stability are inversely proportional, lithium heteroligomers are most stable, followed by sodium heteroligomers and then potassium heteroligomers. This may be due to steric effects resulting from the increased ion size. From preliminary observations of the theoretical structures formed through simulated annealing, potassium has been solely observed in the outer regions of the heteroligomer, while lithium has additionally been observed in the center of an amino acid cluster. However, calcium clusters, with a charge of +2, have a range of approximately -4900 to -5000 kJ/mol. This is more negative than even that of lithium, making calcium heteroligomers the most stable of all. Evidently the effect of the increased charge on calcium predominates over any steric effects caused by its size, which is almost identical to potassium.

In summary, the trend shows that for metal ions of a certain charge, as ion size increases, heteroligomer stability decreases. However, as ion charge increases, heteroligomer stability will increase as well. This is true for both cysteine and proline oligomers as well as oligomers other than the hexamer.

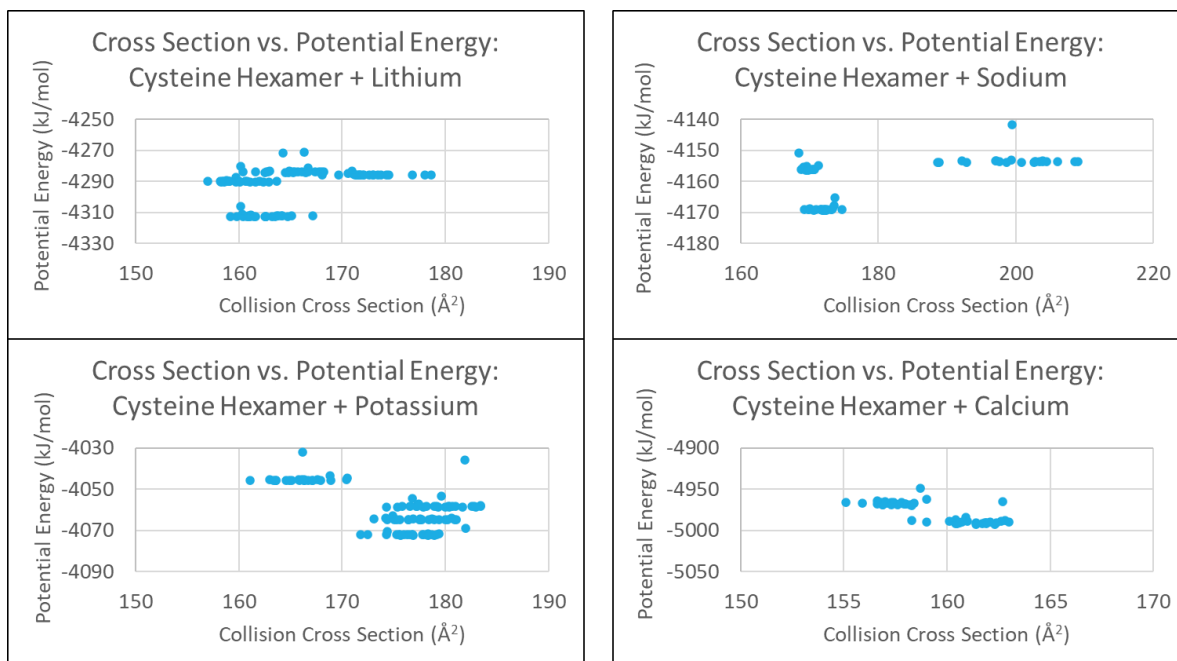


Figure 3.11. A comparison of scatter plots of potential energy versus collision cross section for the cysteine hexamer complexed with lithium, sodium, potassium and calcium.

Final Conclusions

The concept behind this experiment was to computationally generate a number of oligomer conformations and determine which of these conformations were most energetically favorable and therefore most likely to form. We were subsequently able to compare the cross sections of the theoretical structures to the experimental cross section data and from there, deduce which conformations were present in our experimental samples. Additionally, through a comparison of the simulated structures themselves in parallel with the experimental peak distribution, some preliminary estimations of the mechanism of formation could be made.

In summary, we can draw several key conclusions from the comparative stabilities observed in the potential energy scatter plots above. First, proline clusters are always higher energy and thus less stable than their cysteine counterparts. This holds true when comparing pure and metal-complexed amino acid clusters, and for comparing heteroligomers of all four metal ions. Second, for both proline and cysteine, a cluster which is complexed to a metal ion will always be more stable than its pure amino acid counterpart. This is observed for both proline and cysteine, and for any of the four metal ions. Finally, the four metal cations can be compared to each other. At a given charge level stability decreases as metal ion size increases, with lithium being most stable, followed by sodium and then potassium. However, the presence of higher-charged metal ions, such as calcium, results in increased stability when compared to ions with a lower charge.

Looking back to the experimentally derived results in Part 2 of this dissertation, we see the presence of higher-order oligomer formation for both cysteine and proline, in alignment with my extension of the Generic Amyloid Hypothesis to include not just amyloidogenic fibrils, but large oligomers as well. It has been demonstrated from the computational data that the presence of metal ions in the same location where amino acid aggregation would be taking place *in vivo* would stabilize the resulting oligomers. If these oligomers are indeed toxic, this would amplify any negative side effects and damage caused within the body. Therefore, I conclude that knowledge of the mechanisms of metal ion heteroligomer formation has relevance to the ongoing process to characterize and understand metabolic disorders.

Part 4: Appendix

Included in this section are all of the figures and tables which could not be included in the main portion of the dissertation, but were still used to formulate the conclusions which were presented there.

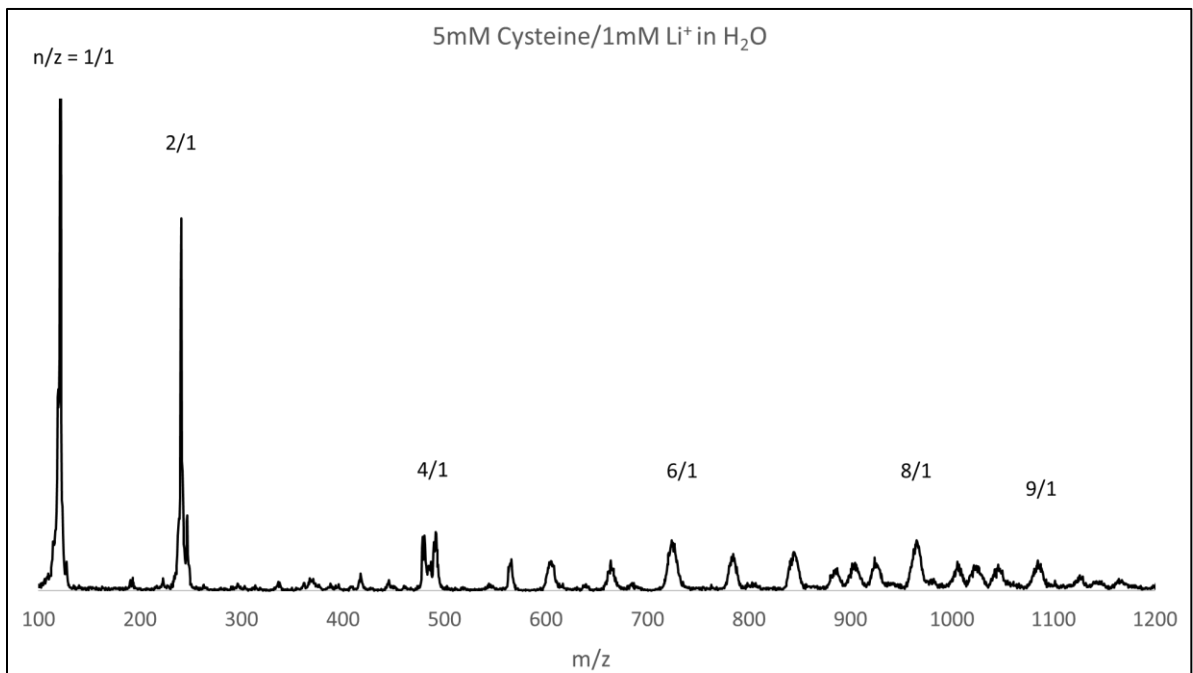


Figure A.1. Mass spectrum of an aqueous solution of 5 mM cysteine and 1 mM lithium ions. The x-axis corresponds to the mass-to-charge ratio, while the y-axis corresponds to signal intensity. Each peak is labeled with its corresponding n/z value, where n is the oligomer number and z is the charge.

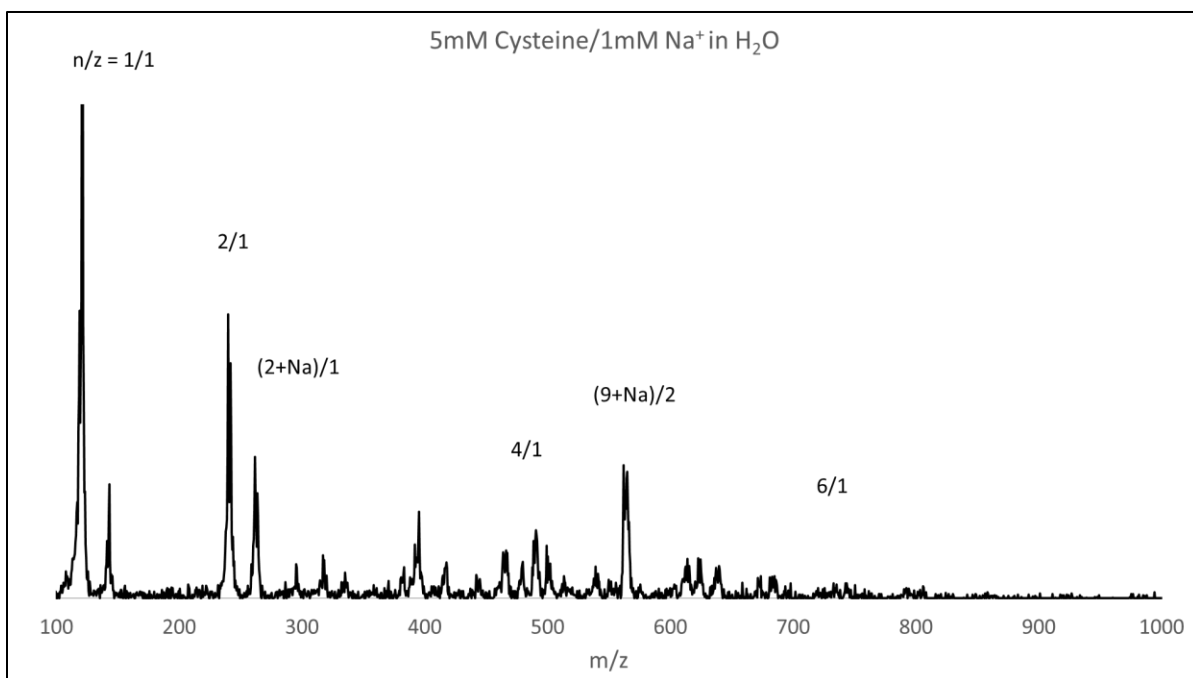


Figure A.2. Mass spectrum of an aqueous solution of 5 mM cysteine and 1 mM sodium ions.

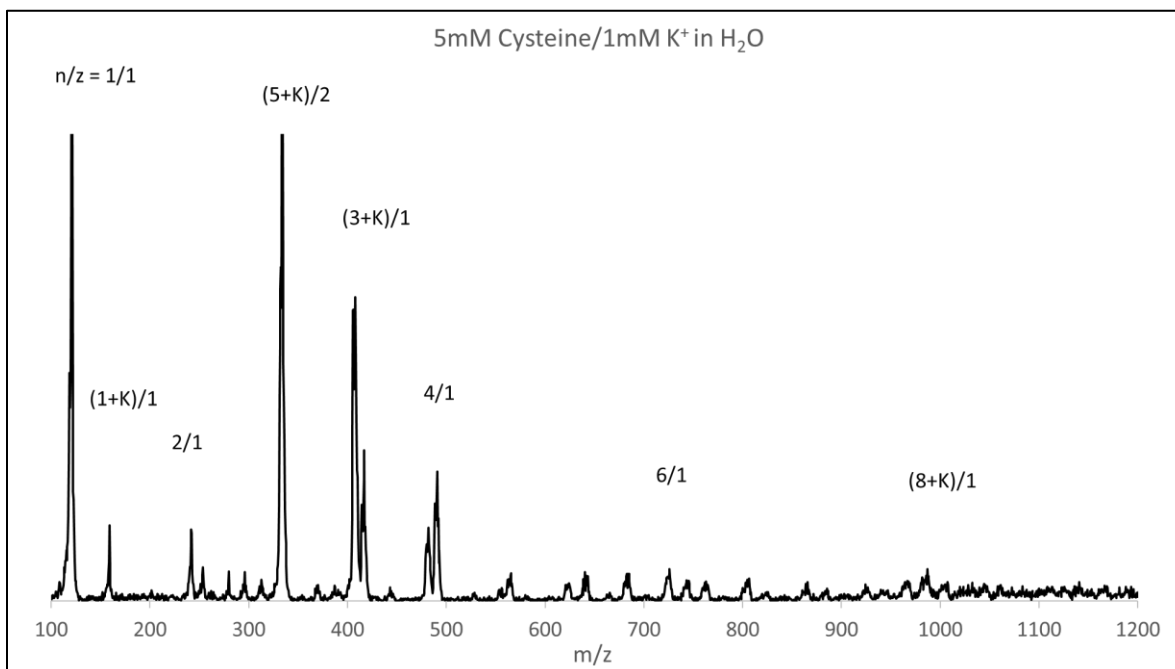


Figure A.3. Mass spectrum of an aqueous solution of 5 mM cysteine and 1 mM potassium ions.

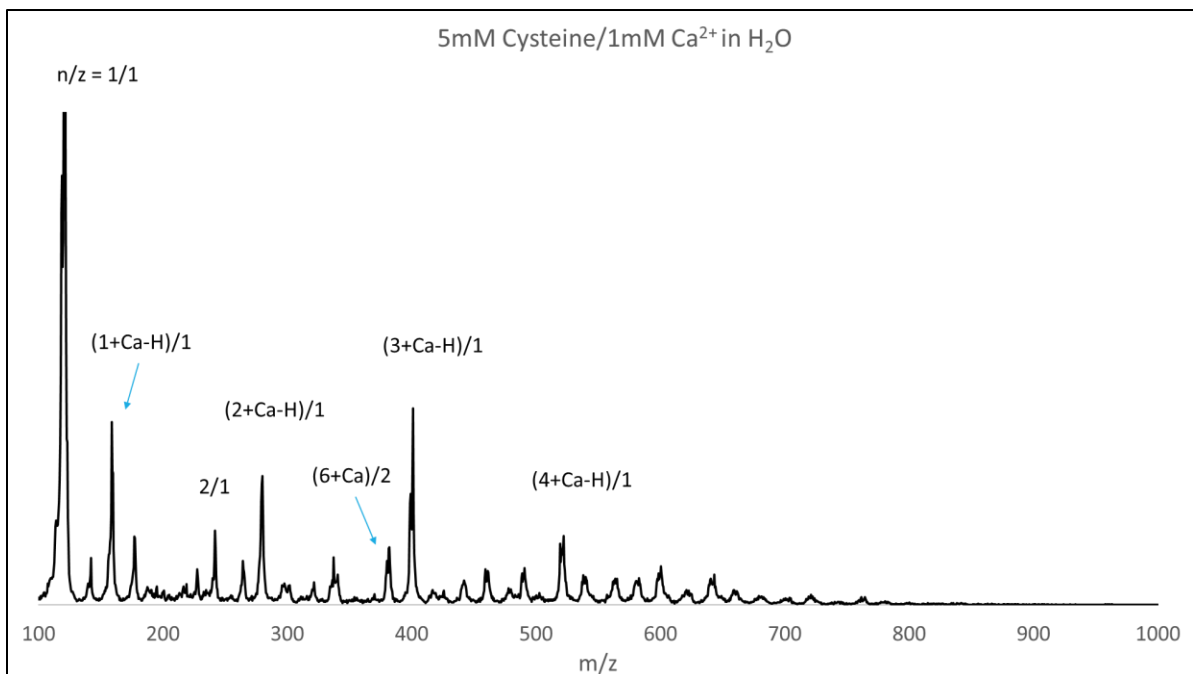


Figure A.4. Mass spectrum of an aqueous solution of 5 mM cysteine and 1 mM calcium ions. The signal dies down after the hexamer peak. This could be due to a mass cutoff issue but is more likely to be due to calcium limiting the formation of high-order singly charged clusters, as the same trend is not seen in the case of proline (see Figure A.8). Cysteine was expected to form standard doubly charged clusters of cysteine held together with a single calcium, similar to those seen both for proline/calcium as well as cysteine with the other metals. However, the peaks which formed did not match that assignment, and the only calculated assignment that did match the observed peaks assumed a singly-charged oligomer with one hydrogen removed to offset the charge on the calcium. These peaks have been tentatively identified as singly-charged oligomers, but are still subject to change until further experimentation can be carried out.

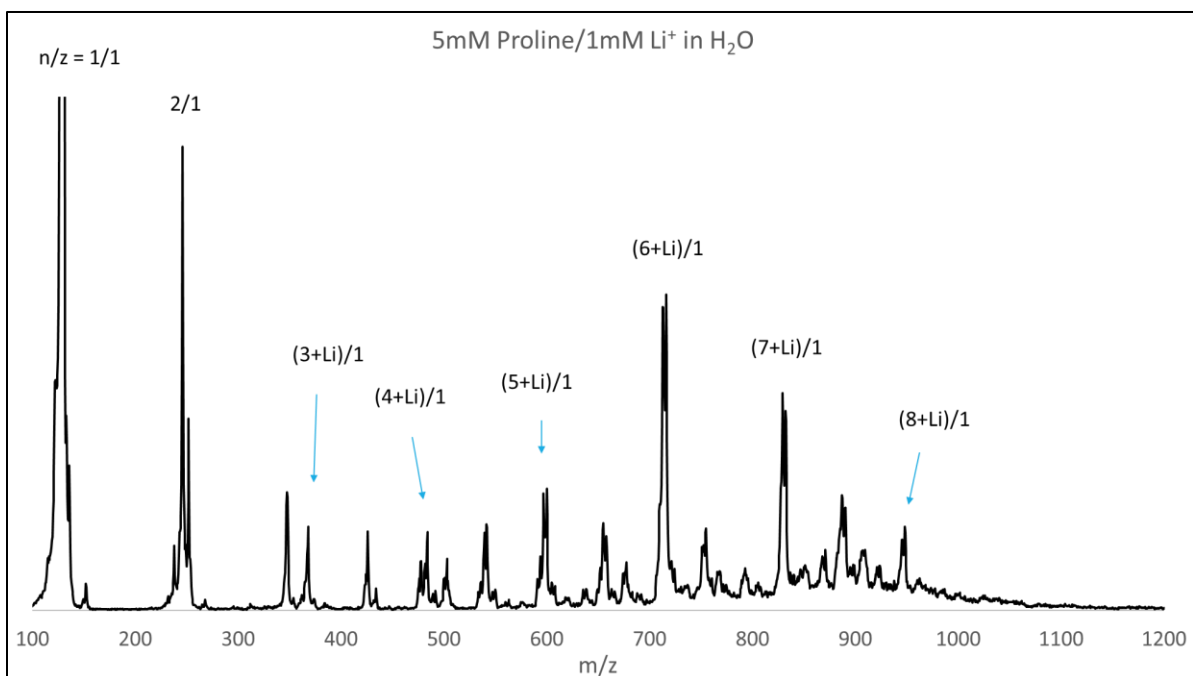


Figure A.5. Mass spectrum of an aqueous solution of 5 mM proline and 1 mM lithium ions.

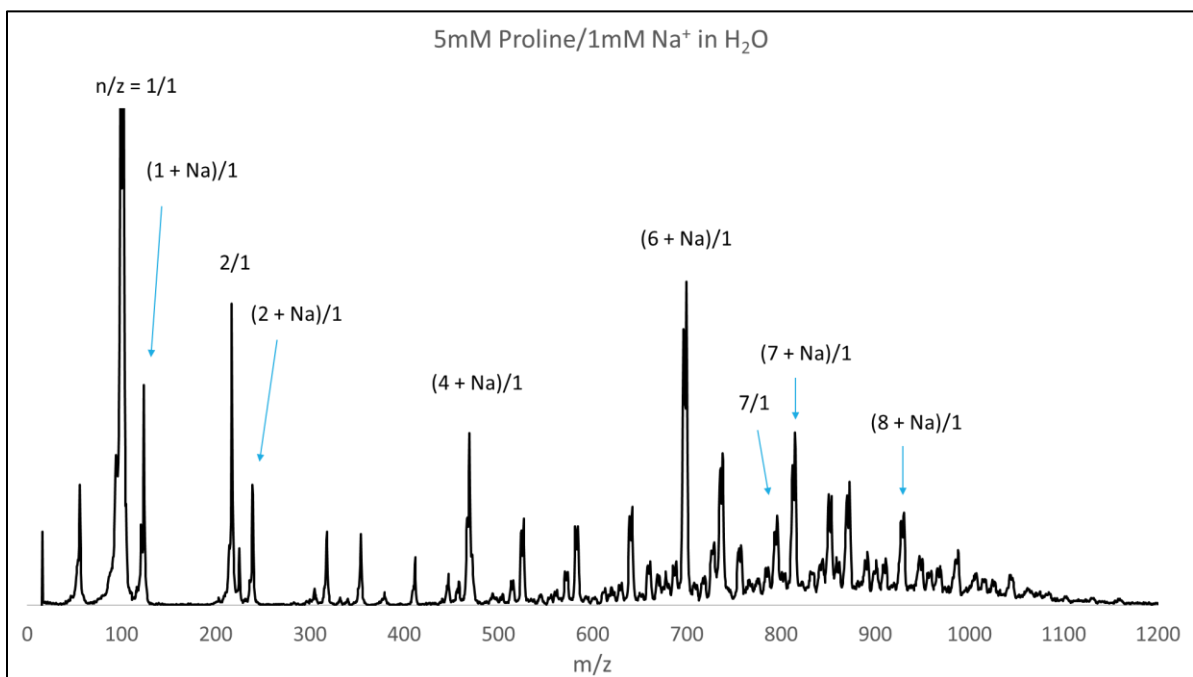


Figure A.6. Mass spectrum of an aqueous solution of 5 mM proline and 1 mM sodium ions.

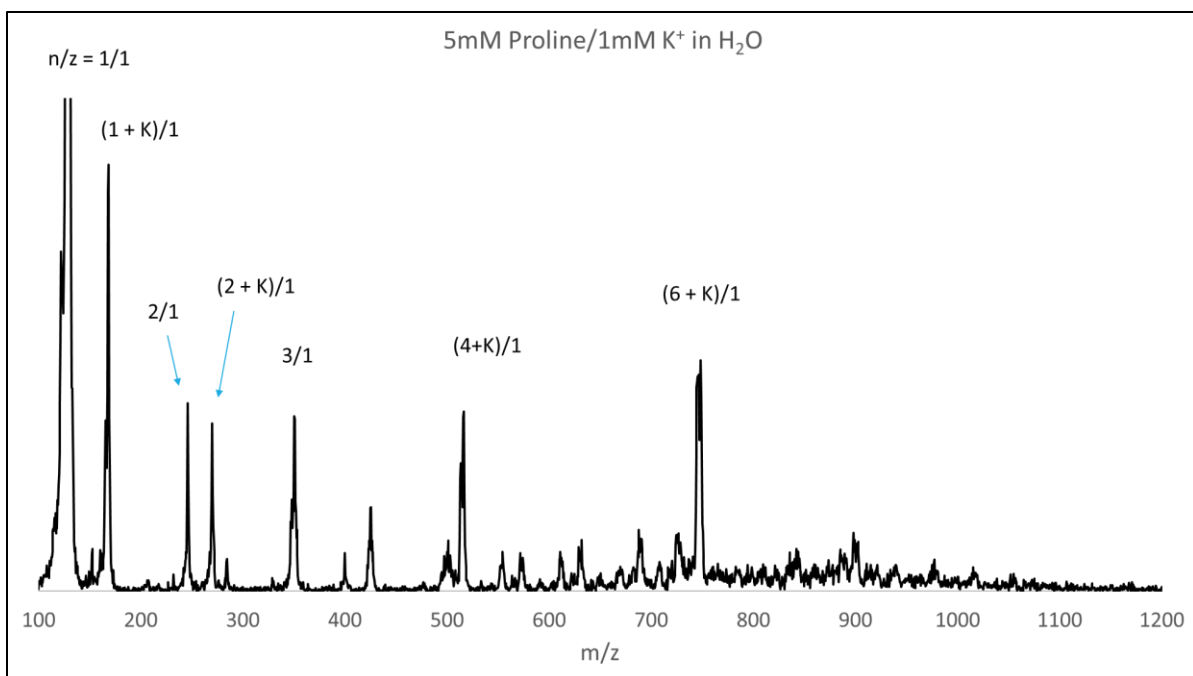


Figure A.7. Mass spectrum of an aqueous solution of 5 mM proline and 1 mM potassium ions.

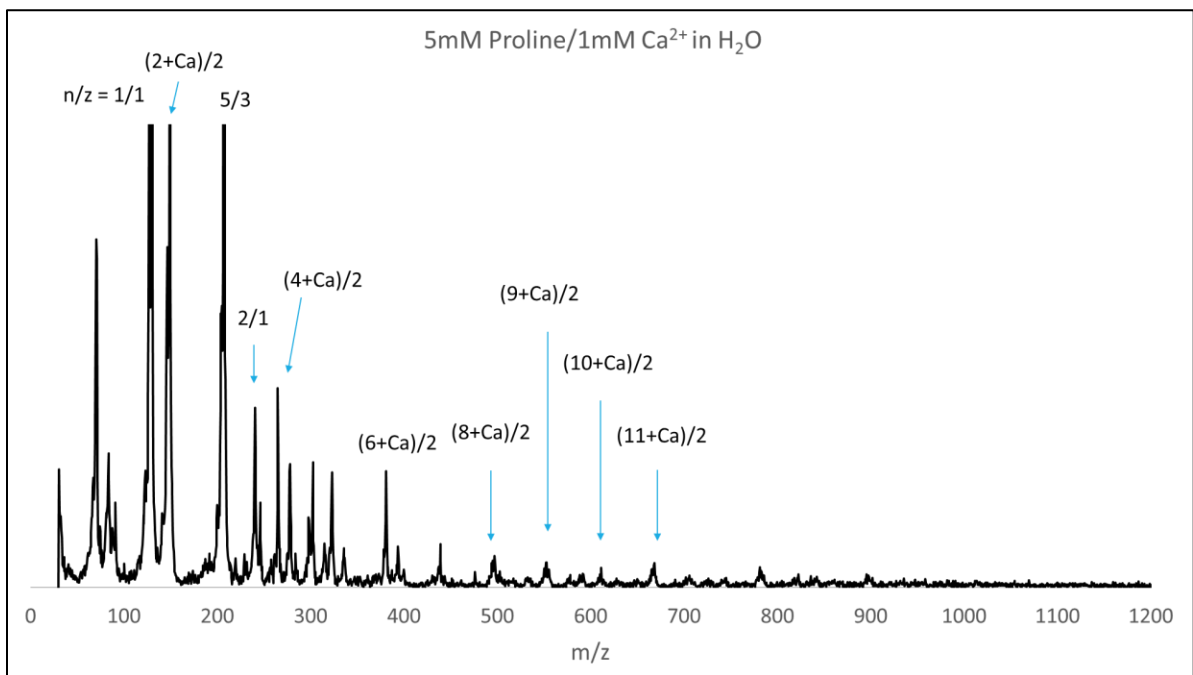
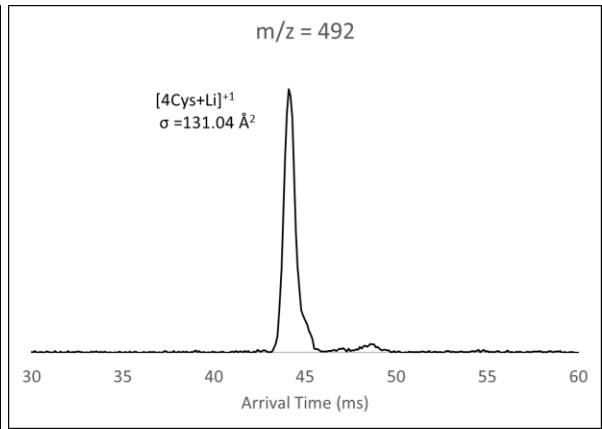
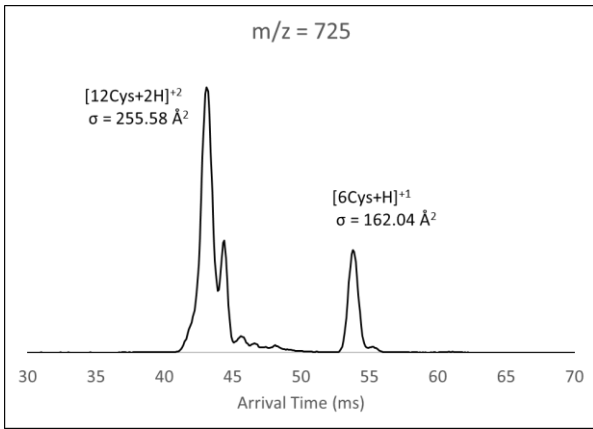
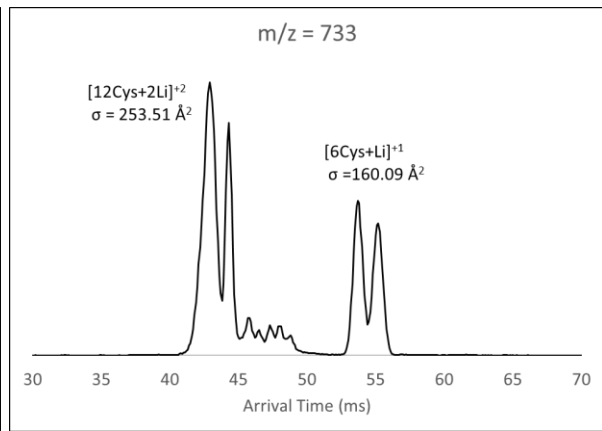
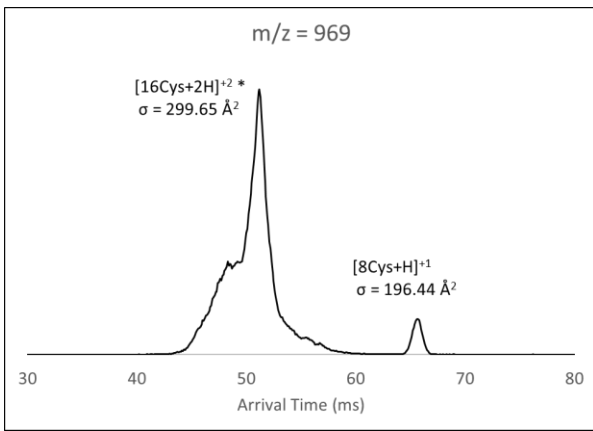
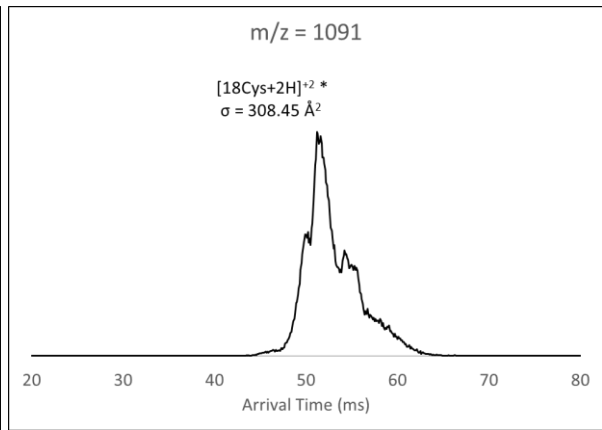
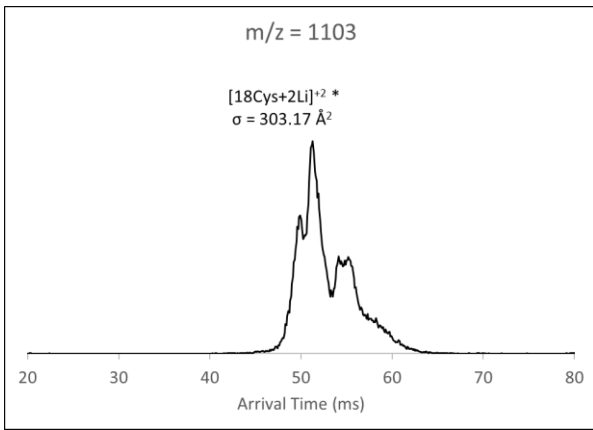


Figure A.8. Mass spectrum of an aqueous solution of 5 mM proline and 1 mM calcium ions. Unlike cysteine, proline does appear to form higher-order clusters analogous to the singly charged large oligomers seen for other metals. However, due to the increased charge on the calcium, proline-calcium clusters that are the same size as these singly charged oligomers will appear at an earlier arrival time in the ATD.



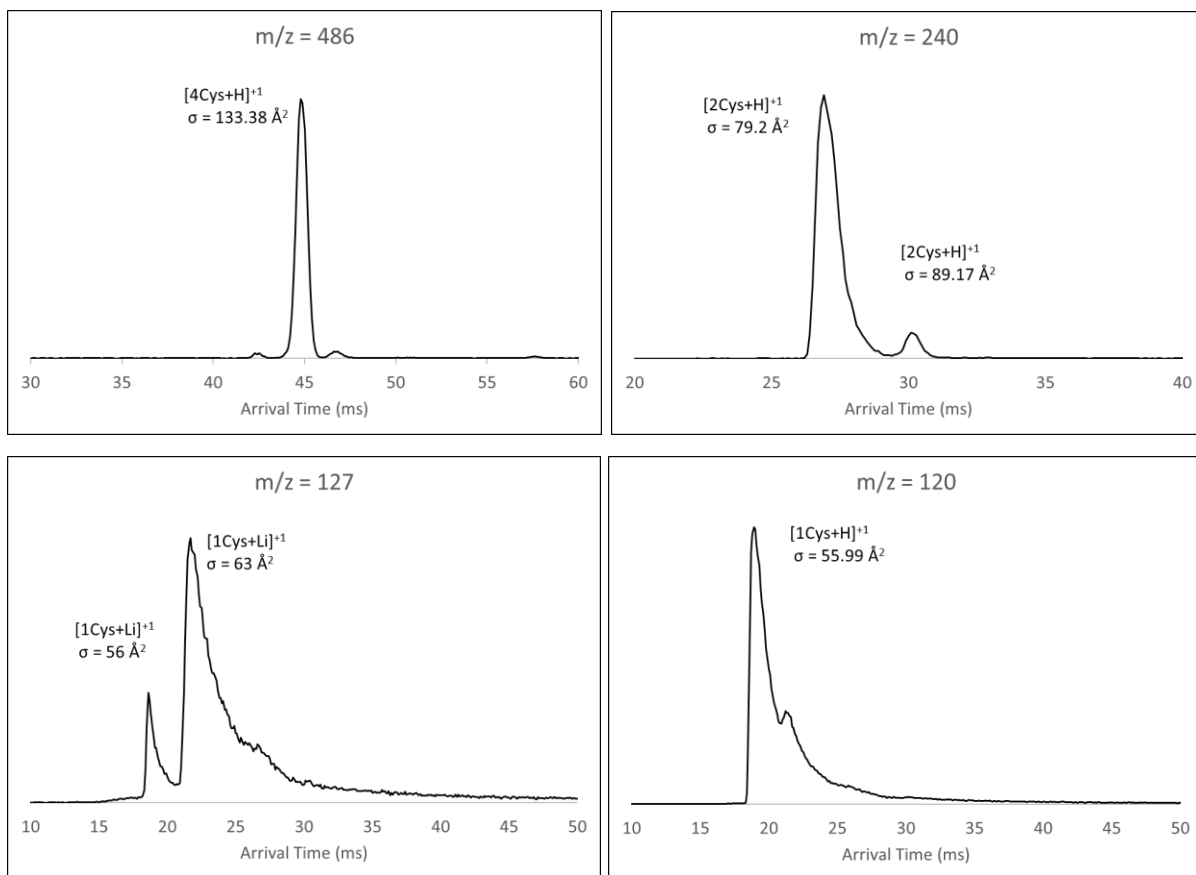
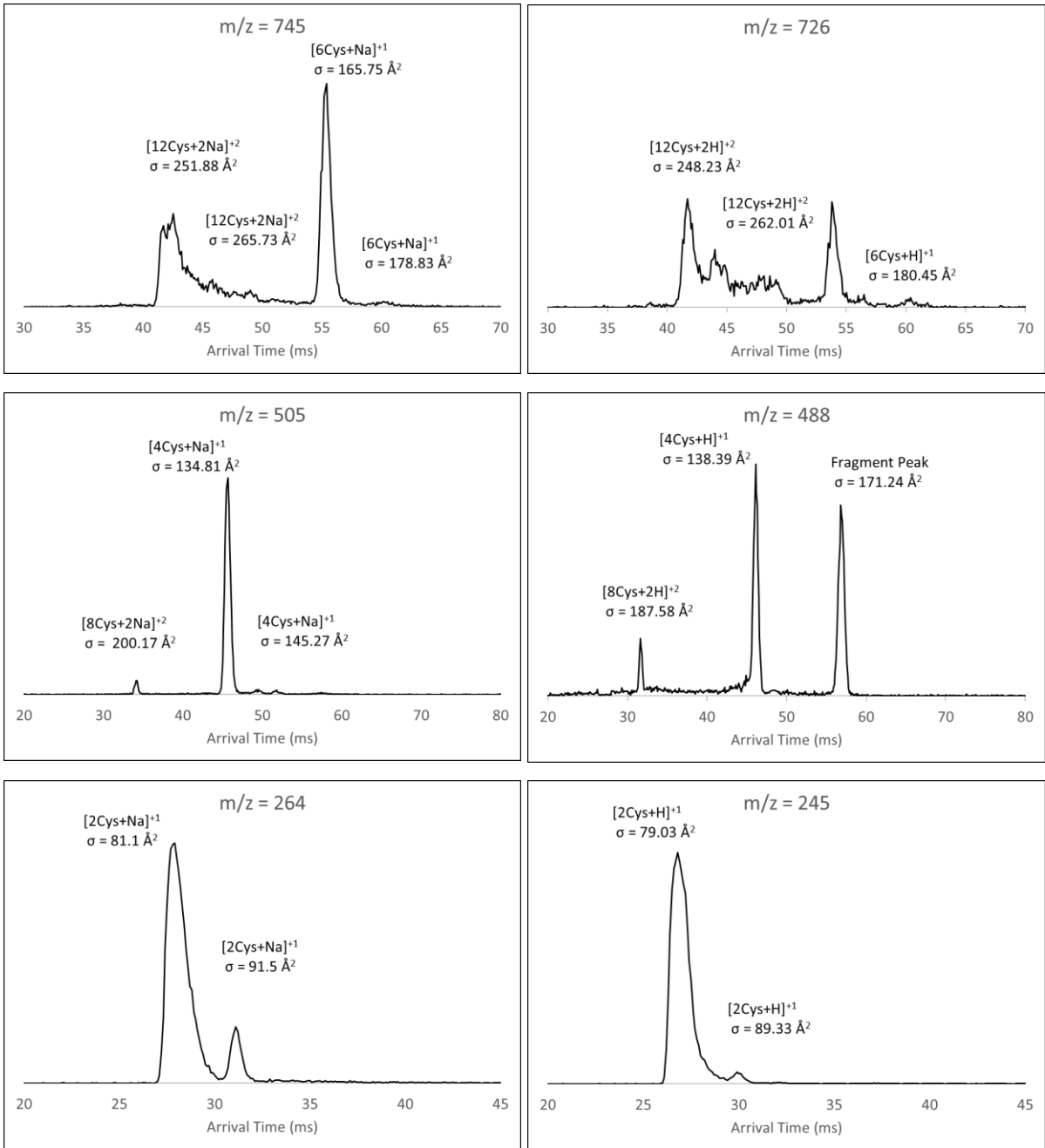


Figure A.9. Arrival time distribution (ATD) plots for oligomers of cysteine and lithium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS. All ATDs for a given amino acid/metal ion solution were taken from the same solution and recorded on the same day. For the higher mass-to-charge ratios, some preliminary peak assignments have been made based on the experimental data, with computational follow-up experiments still ongoing. These preliminary assignments have been marked with an asterisk where applicable.



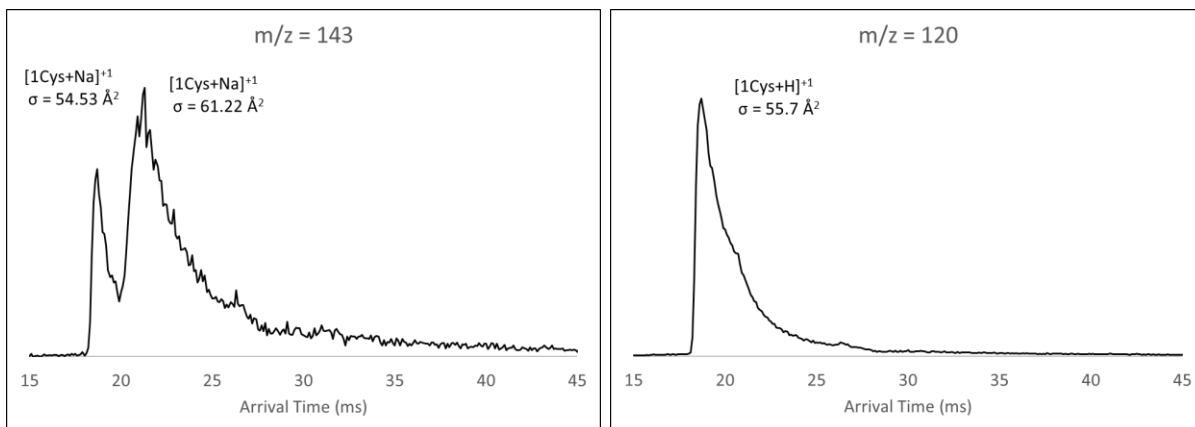
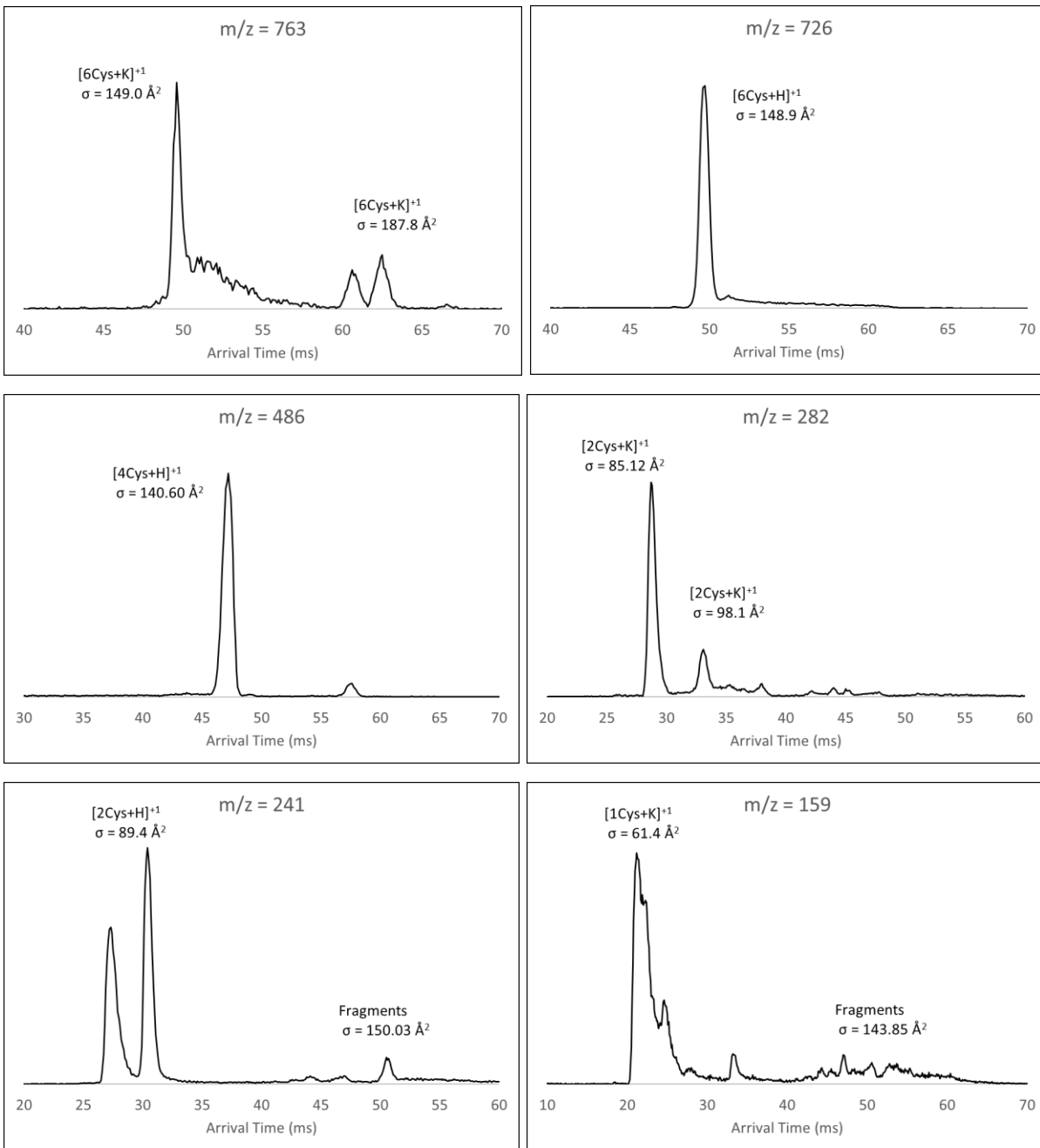


Figure A.10. Arrival time distribution (ATD) plots for oligomers of cysteine and sodium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS.



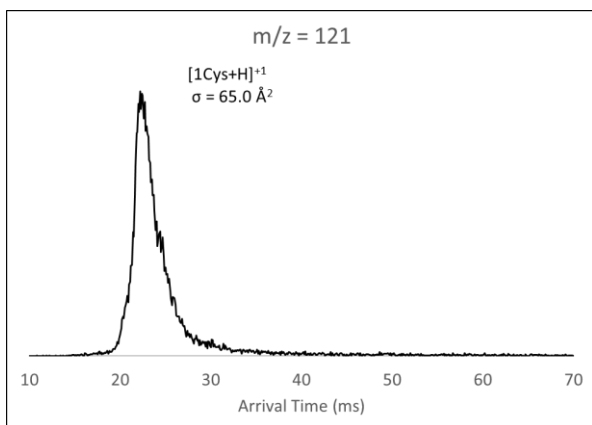
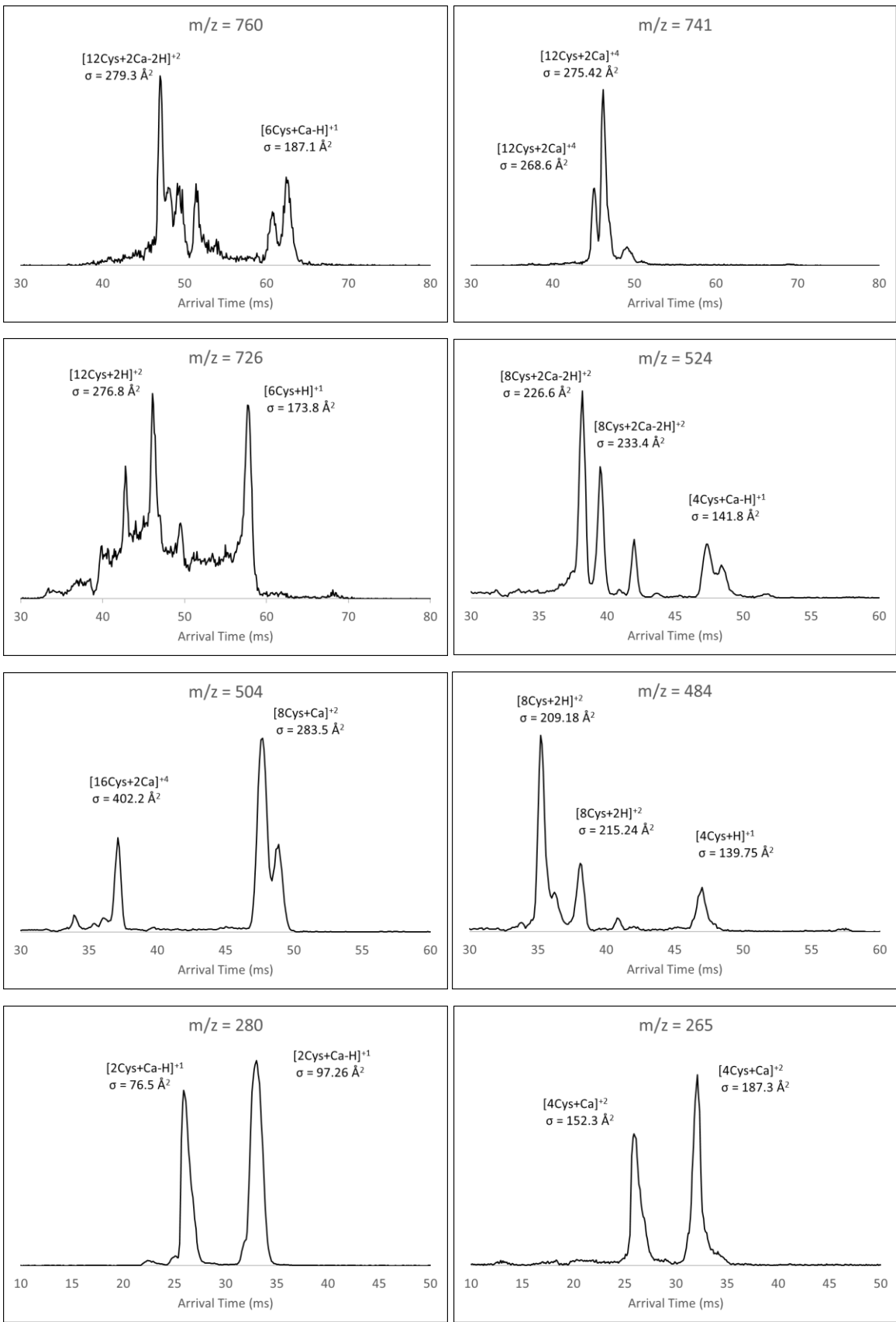


Figure A.11. Arrival time distribution (ATD) plots for oligomers of cysteine and potassium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS.



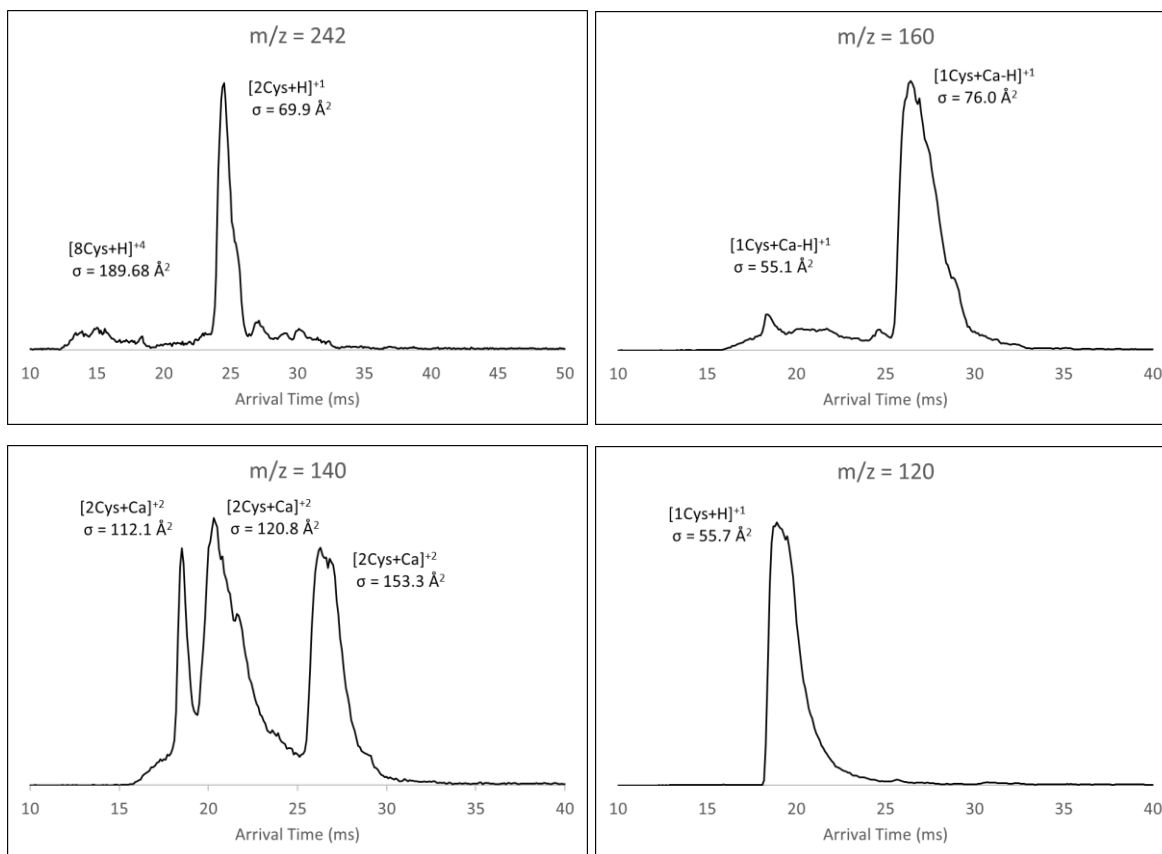
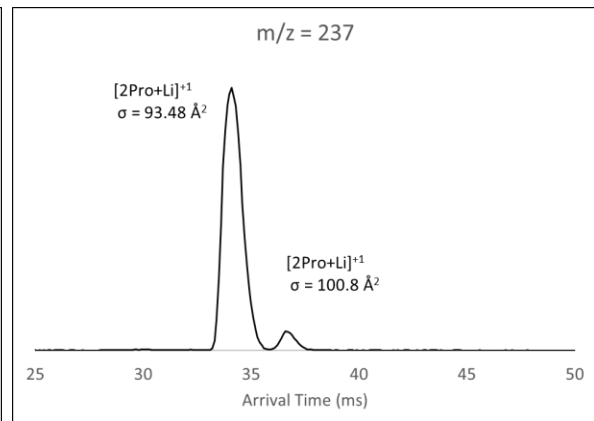
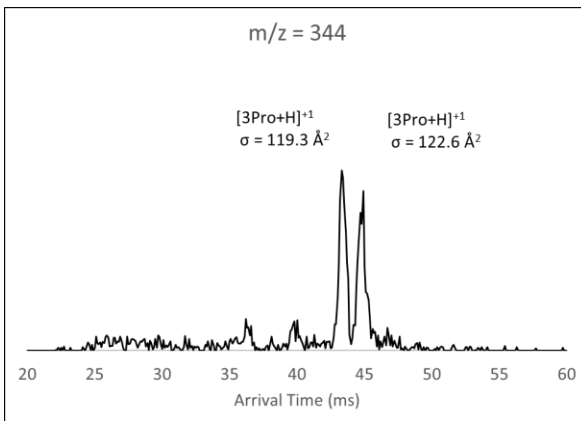
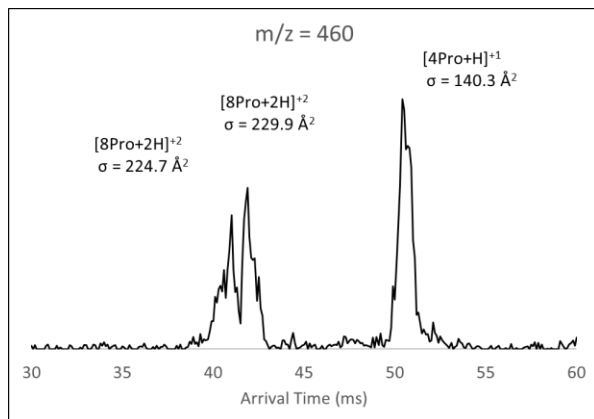
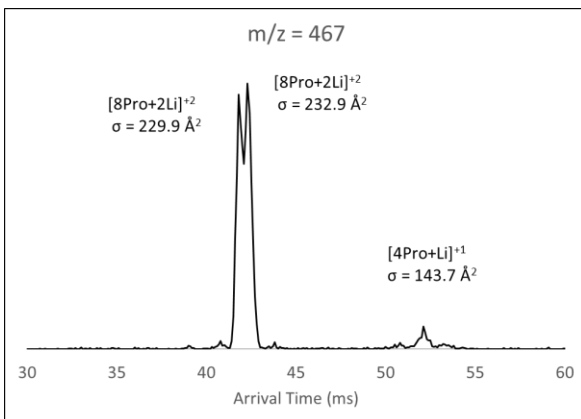
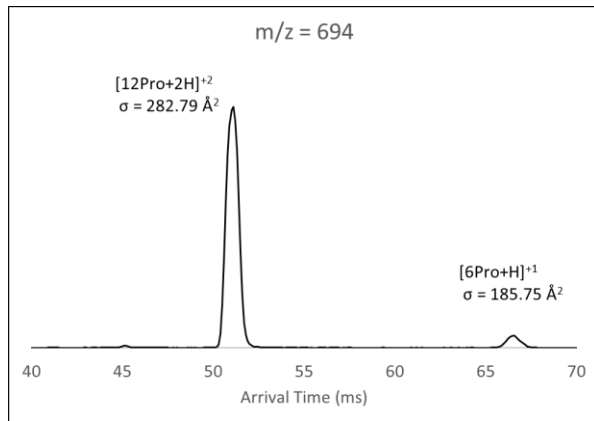
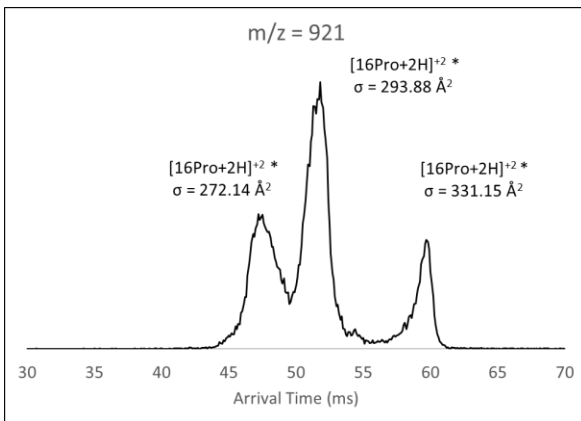


Figure A.12. Arrival time distribution (ATD) plots for oligomers of cysteine and calcium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS. For the higher mass-to-charge ratios, some preliminary peak assignments have been made based on the experimental data, with computational follow-up experiments still ongoing. These preliminary assignments have been marked with an asterisk where applicable.



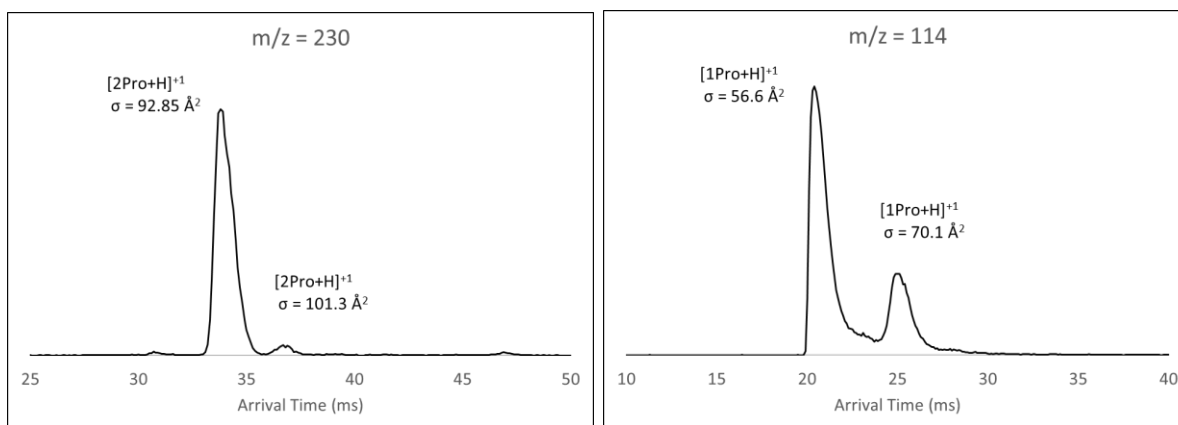
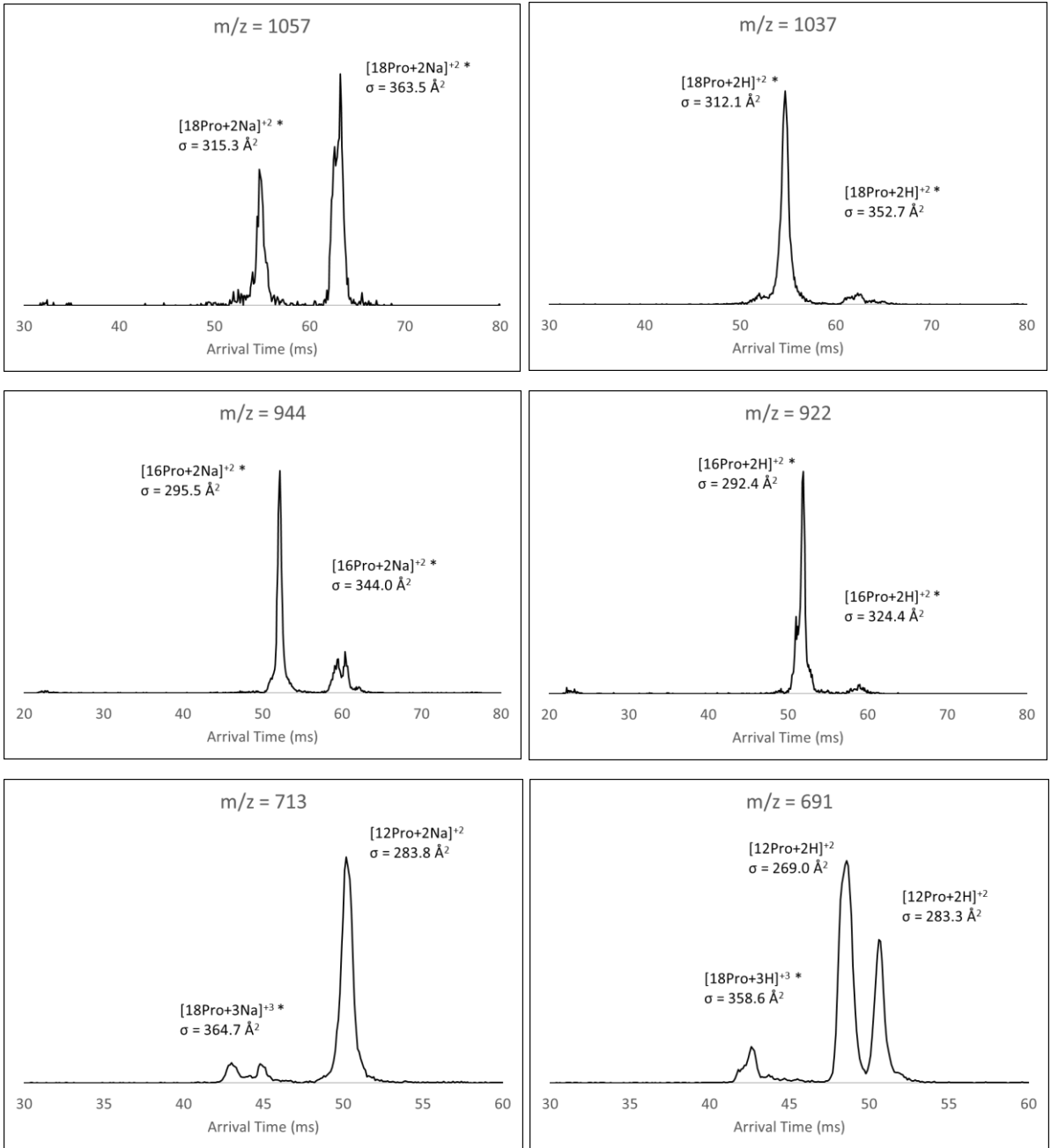


Figure A.13. Arrival time distribution (ATD) plots for oligomers of proline and lithium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS. For the mass-to-charge ratio corresponding to the singly charged proline octamer ($m/z = 921$), some preliminary peak assignments have been made based on the experimental data, with computational follow-up experiments still ongoing. These preliminary assignments have been marked with an asterisk.



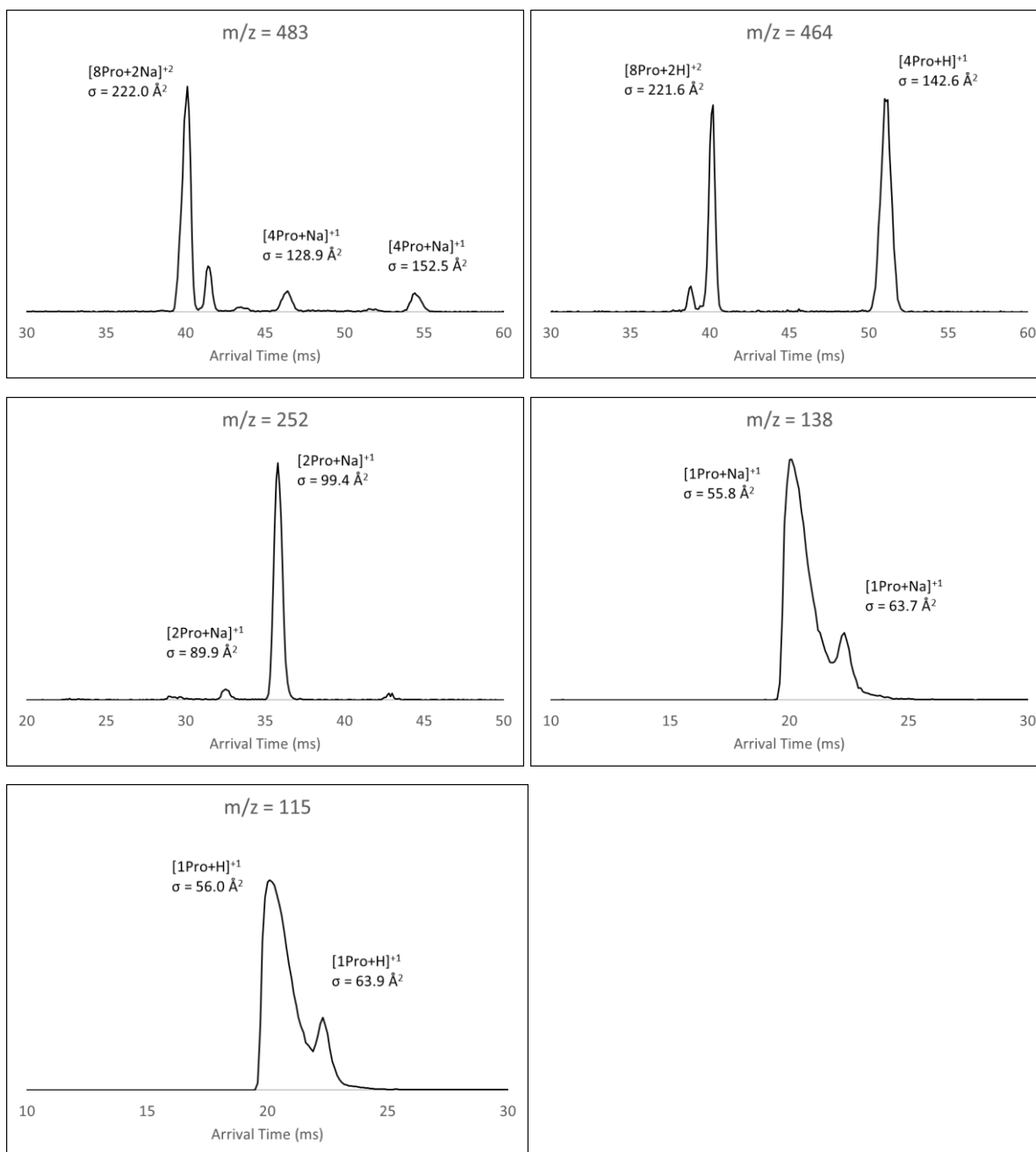
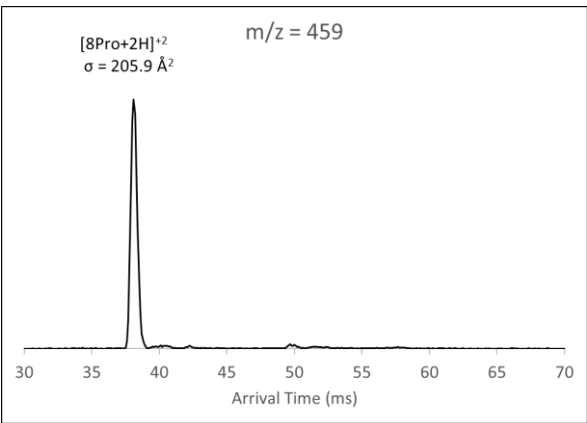
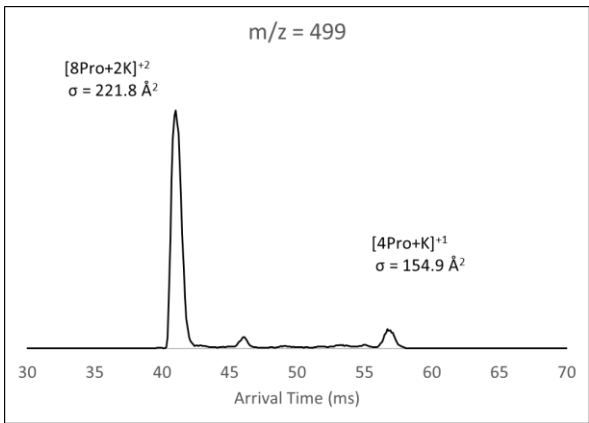
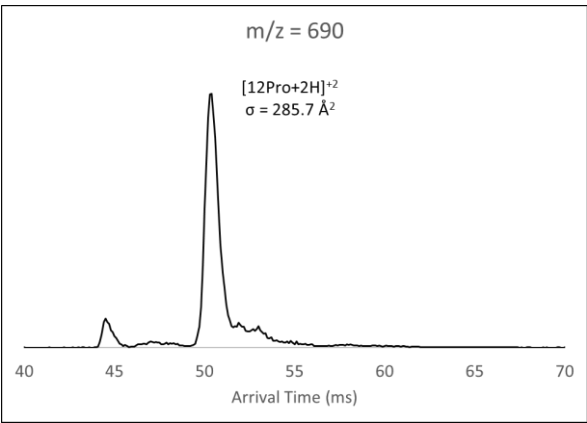
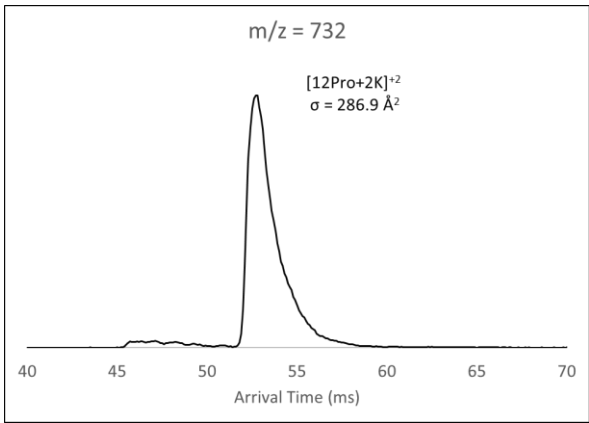
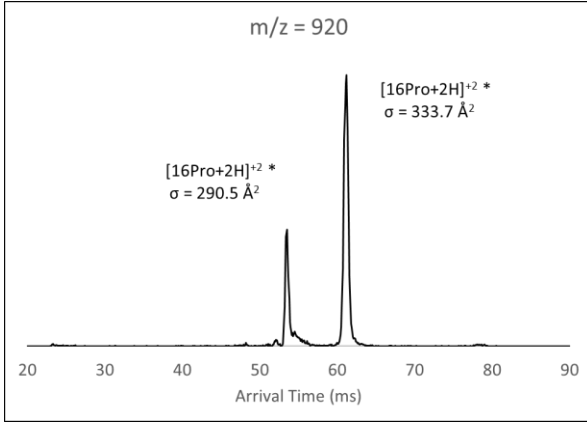
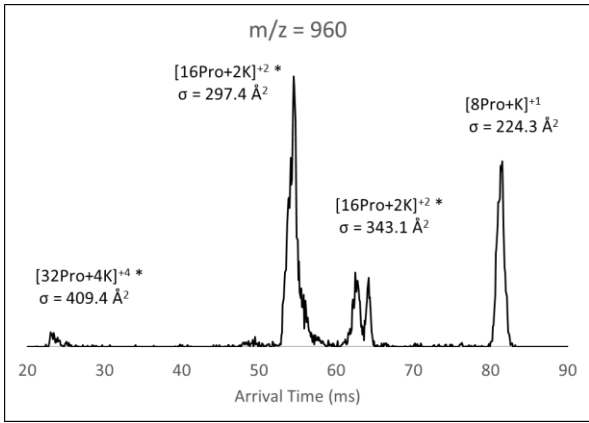


Figure A.14. Arrival time distribution (ATD) plots for oligomers of proline and sodium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS. For several of the higher mass species, some preliminary peak assignments have been made based on the experimental data, with computational follow-up experiments still ongoing. These preliminary assignments have been marked with an asterisk where applicable.



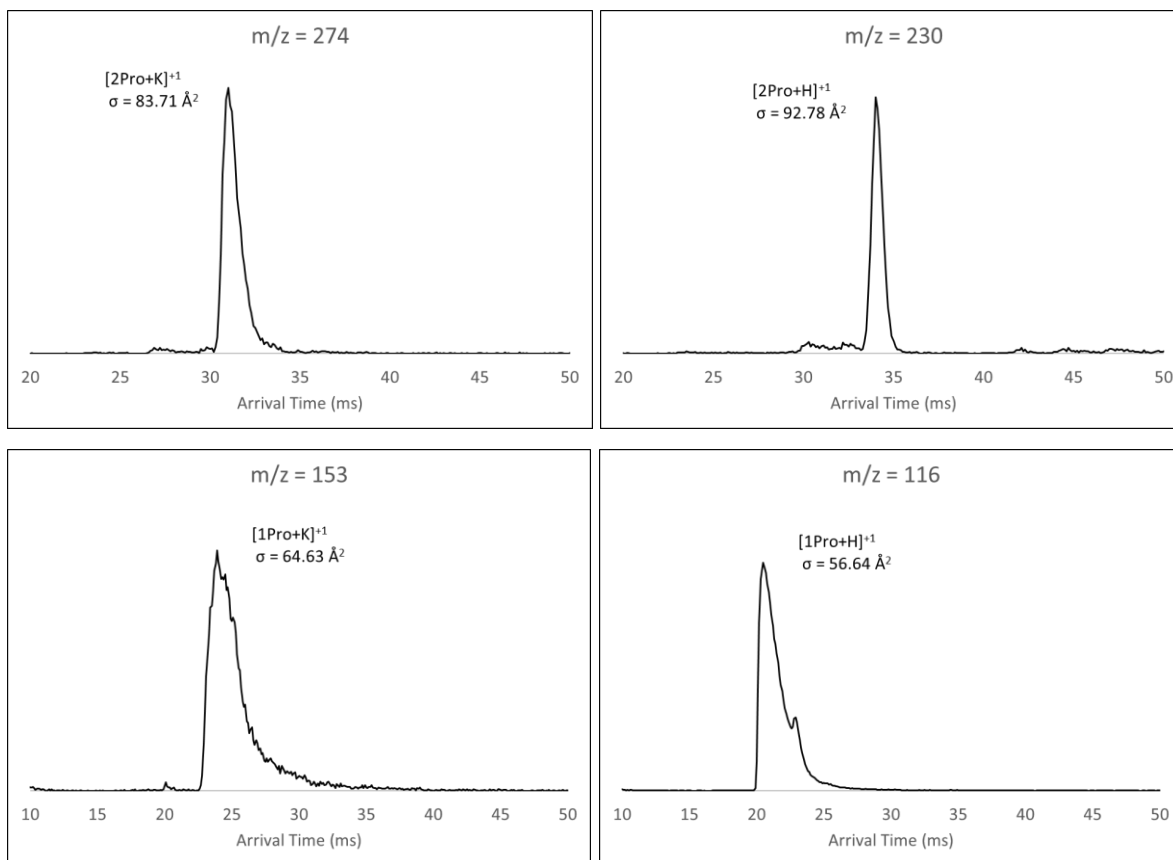
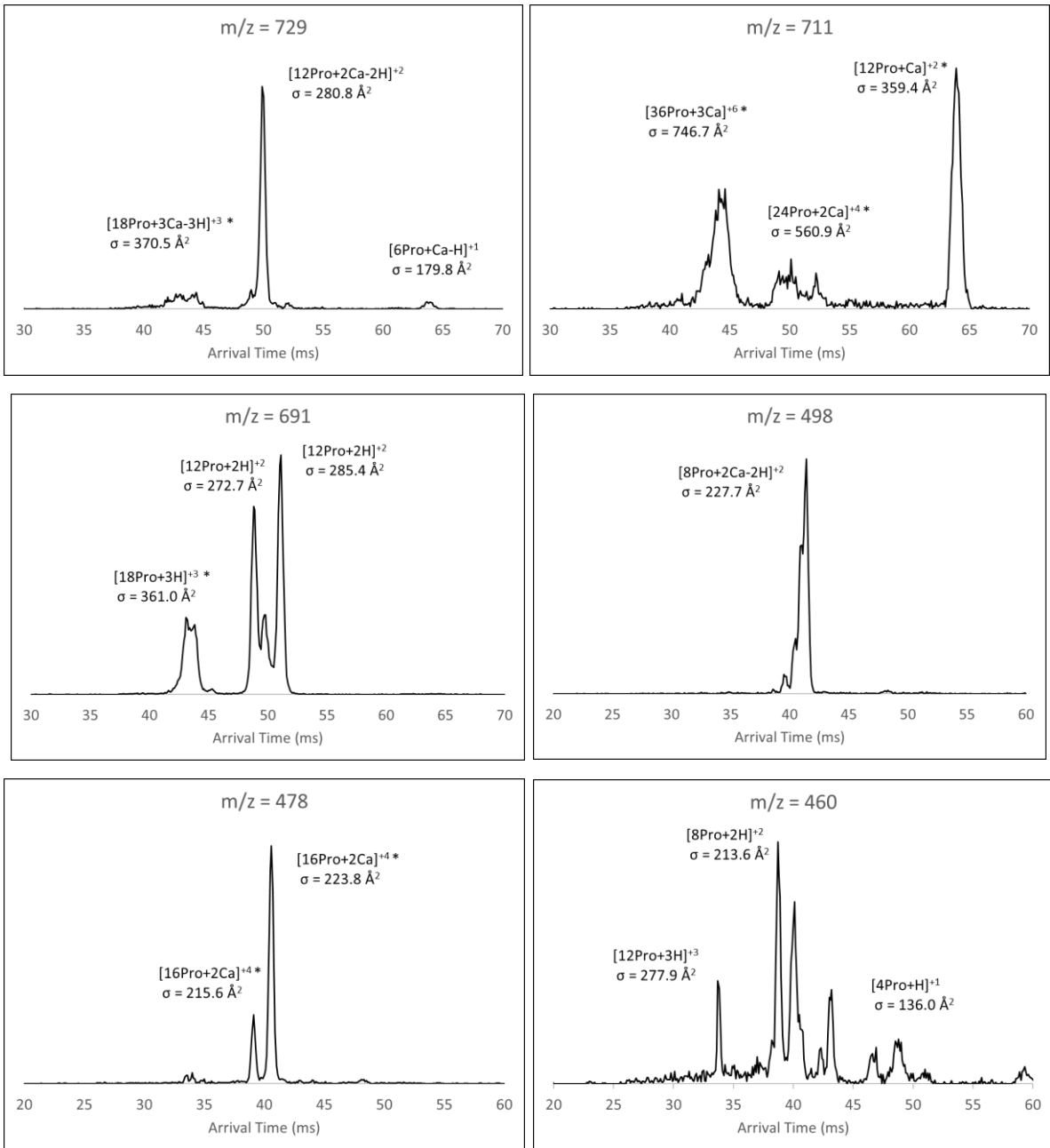
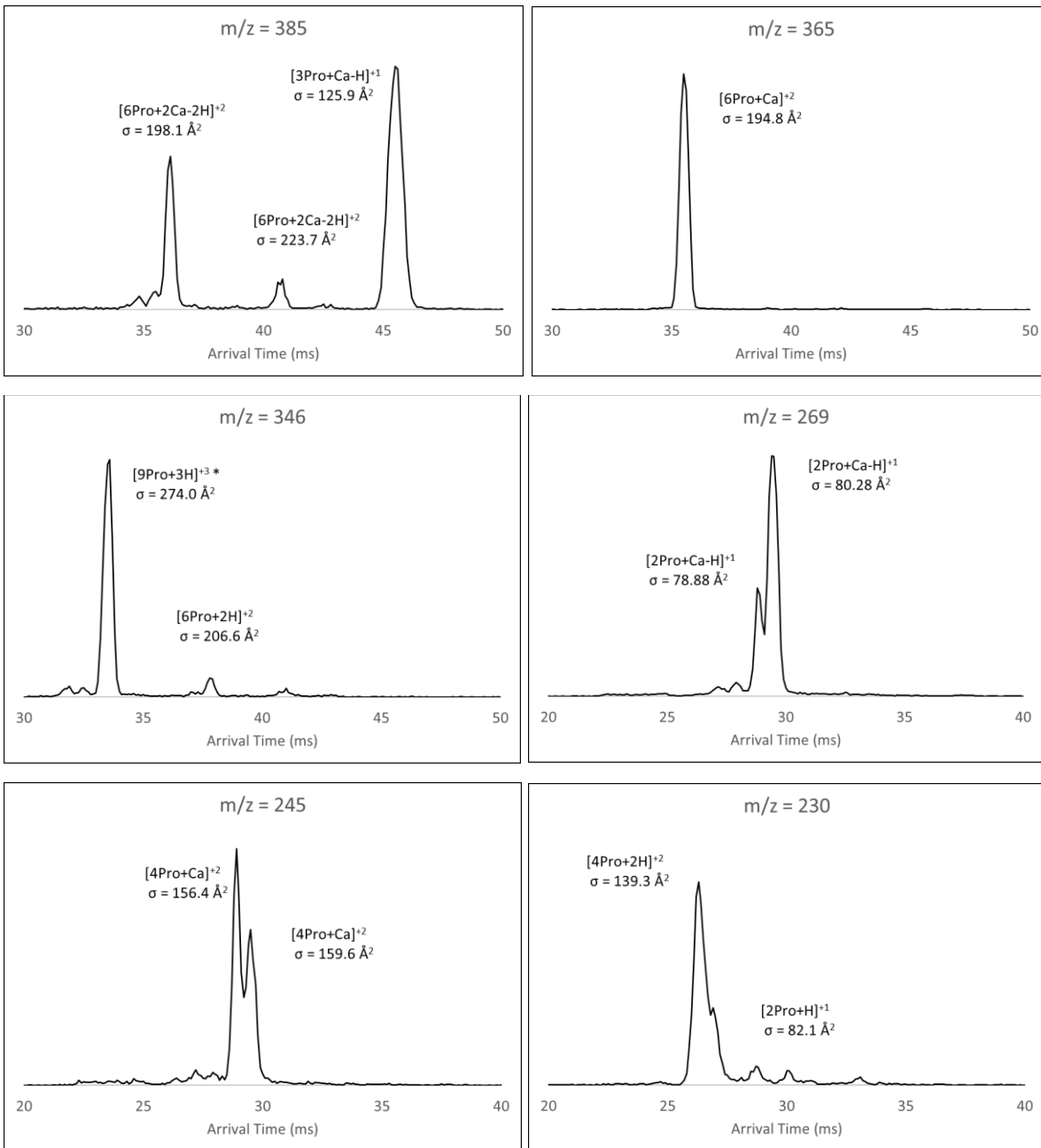


Figure A.15. Arrival time distribution (ATD) plots for oligomers of proline and potassium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS. For the highest mass-to-charge ratios, some preliminary peak assignments have been made based on the experimental data, with computational follow-up experiments still ongoing. These preliminary assignments have been marked with an asterisk.





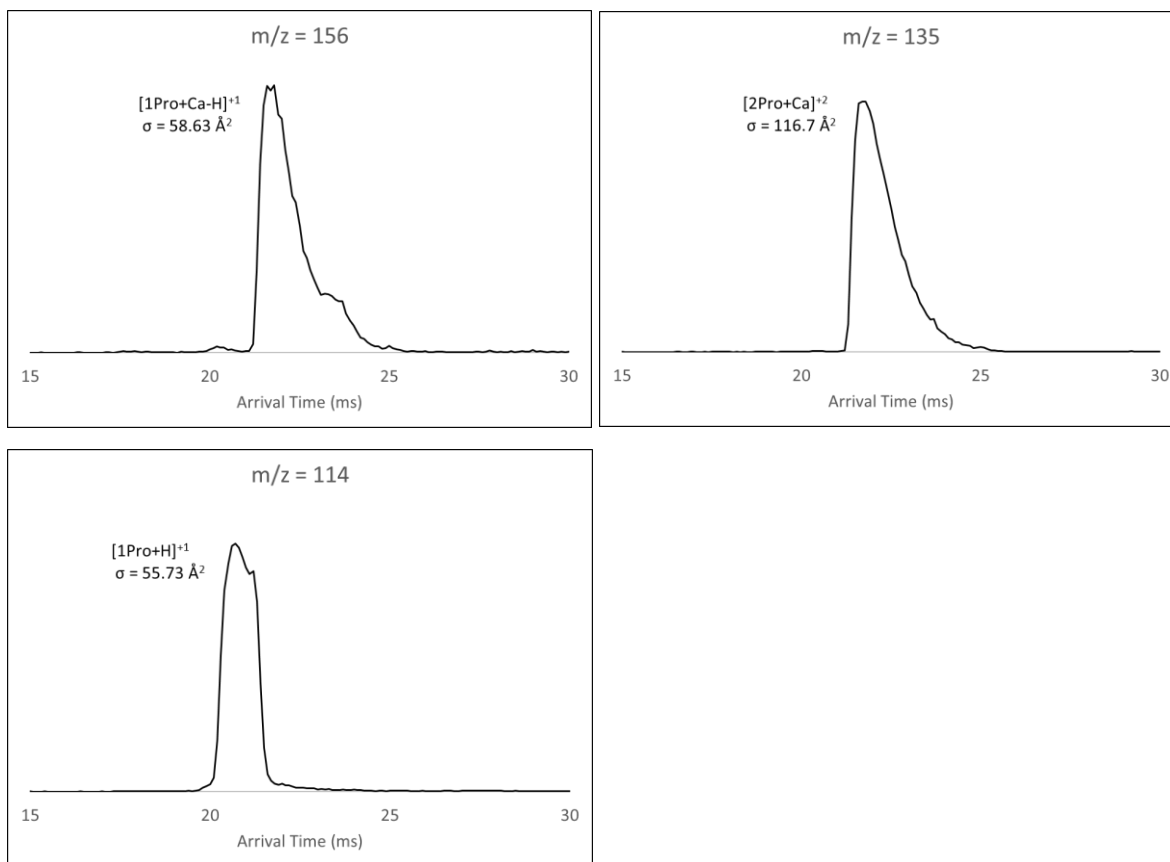


Figure A.16. Arrival time distribution (ATD) plots for oligomers of proline and calcium. Each ATD corresponds to a given mass-to-charge ratio, and each peak is labeled with the assigned oligomer and the experimental CCS. For the highest mass-to-charge ratios, some preliminary peak assignments have been made based on the experimental data, with computational follow-up experiments still ongoing. These preliminary assignments have been marked with an asterisk.

Collision Cross Section Data

The following are tables of various heteroligomers of cysteine and proline with lithium, sodium, potassium and calcium. Many of the theoretical collision cross sections (CCS) presented herein were calculated by myself using structures derived from MD. For several of the proline oligomers, theoretical values were taken from the proline IM-MS study cited earlier in this dissertation.⁶⁵ These values will be marked with an asterisk.

Table A.1. A list of cysteine-lithium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD. More research is ongoing to continuously augment this data.

Cysteine Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Cys+Li}]^{+1}$	56.0 / 63.0	-
$[2\text{Cys+Li}]^{+1}$	79.2 / 89.17	-
$[4\text{Cys+Li}]^{+1}$	131.04	-
$[6\text{Cys+Li}]^{+1}$	160.09	165.4
$[12\text{Cys+2Li}]^{+2}$	253.51	-
$[8\text{Cys+Li}]^{+1}$	196.44	-
$[16\text{Cys+2Li}]^{+2}$	299.65	-

Table A.2. A list of cysteine-sodium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD.

Cysteine Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Cys}+\text{Na}]^{+1}$	54.53 / 61.22	56
$[2\text{Cys}+\text{Na}]^{+1}$	81.1 / 91.5	92
$[4\text{Cys}+\text{Na}]^{+1}$	134.81 / 145.27	132.4
$[8\text{Cys}+2\text{Na}]^{+2}$	200.17	-
$[6\text{Cys}+\text{Na}]^{+1}$	165.75 / 178.83	171.6
$[12\text{Cys}+2\text{Na}]^{+2}$	251.88 / 265.73	-

Table A.3. A list of cysteine-potassium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD.

Cysteine Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Cys}+\text{K}]^{+1}$	61.4	-
$[2\text{Cys}+\text{K}]^{+1}$	85.1 / 98.1	-
$[4\text{Cys}+\text{K}]^{+1}$	140.6	-
$[6\text{Cys}+\text{K}]^{+1}$	149.0 / 187.8	175.3

Table A.4. A list of cysteine-calcium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD.

Cysteine Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Cys}+\text{Ca}-\text{H}]^{+1}$	55.1 / 76	-
$[2\text{Cys}+\text{Ca}-\text{H}]^{+1}$	76.5 / 97.26	-
$[8\text{Cys}+\text{Ca}]^{+2}$	283.5	-
$[16\text{Cys}+2\text{Ca}]^{+4}$	402.2	-
$[4\text{Cys}+\text{Ca}-\text{H}]^{+1}$	141.8	-
$[8\text{Cys}+2\text{Ca}-2\text{H}]^{+2}$	226.6 / 233.4	-
$[12\text{Cys}+2\text{Ca}]^{+4}$	268.6 / 275.42	-
$[6\text{Cys}+\text{Ca}-\text{H}]^{+1}$	187.1	-
$[12\text{Cys}+2\text{Ca}-2\text{H}]^{+2}$	279.3	-

Table A.5. A list of proline-lithium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD. Due to the small size of lithium combined with the weaker binding affinity of proline in comparison to cysteine, we encountered some difficulty in obtaining stable structures from MD. Research is currently ongoing to augment this data.

Proline Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[2\text{Pro}+\text{Li}]^{+1}$	93.48 / 100.8	-
$[4\text{Pro}+\text{Li}]^{+1}$	143.7	-
$[8\text{Pro}+2\text{Li}]^{+2}$	229.9 / 232.9	-
$[6\text{Pro}+\text{Li}]^{+1}$	-	185 *
$[12\text{Pro}+\text{Li}]^{+1}$	-	291 *

Table A.6. A list of proline-sodium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD.

Proline Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Pro}+\text{Na}]^{+1}$	55.8 / 63.7	56.7 *
$[2\text{Pro}+\text{Na}]^{+1}$	89.9 / 99.4	-
$[4\text{Pro}+\text{Na}]^{+1}$	128.9	145.2
$[8\text{Pro}+2\text{Na}]^{+2}$	222.0	-
$[12\text{Pro}+2\text{Na}]^{+2}$	283.82	297 *

Table A.7. A list of proline-potassium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD.

Proline Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Pro}+\text{K}]^{+1}$	64.63	-
$[2\text{Pro}+\text{K}]^{+1}$	83.71	-
$[4\text{Pro}+\text{K}]^{+1}$	154.87	157 *
$[8\text{Pro}+\text{K}]^{+1}$	224.3	230 *
$[12\text{Pro}+2\text{K}]^{+2}$	286.9	296 *

Table A.8. A list of proline-calcium heteroligomers and their experimental and theoretical collision cross sections determined using IM-MS in tandem with MD.

Proline Oligomer	Experimental CCS (\AA^2)	Theoretical CCS (\AA^2)
$[\text{Pro}+\text{Ca}]^{+2}$	58.63 / 63.94	-
$[2\text{Pro}+\text{Ca}]^{+2}$	75.85	-
$[4\text{Pro}+\text{Ca}]^{+2}$	-	144.6
$[8\text{Pro}+2\text{Ca}-2\text{H}]^{+2}$	227.2	
$[16\text{Pro}+2\text{Ca}]^{+4}$	215.6 / 223.8	-
$[6\text{Pro}+\text{Ca}]^{+2}$	-	185.7
$[12\text{Pro}+\text{Ca}]^{+2}$	359.4	-

Potential Energy Diagrams

Included in this section are several example diagrams of potential energy versus collision cross section, which were plotted using the structures derived through simulated annealing. These were used to derive the theoretical cross sections seen in the tables above. For a few of the larger oligomers, representative structures of each oligomer have been included, modeled using the Avogadro software.

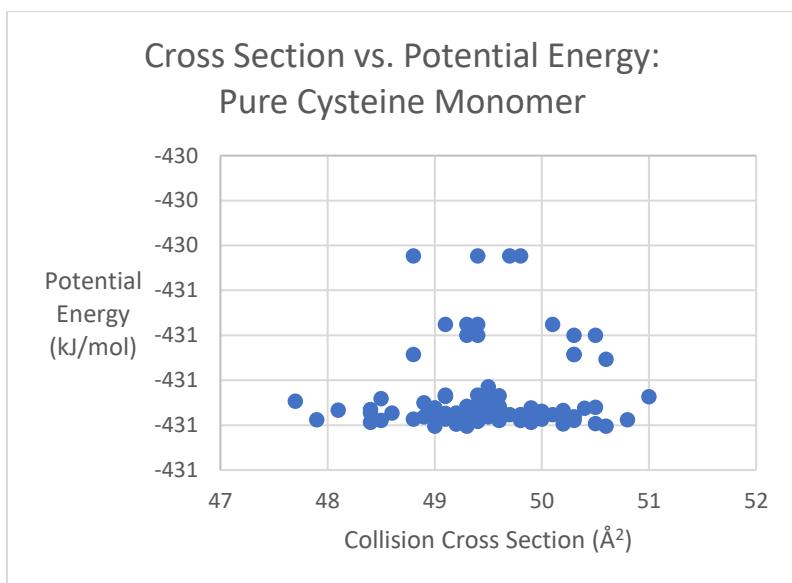


Figure A.17. A scatter plot of potential energy versus CCS for the pure cysteine monomer. The average cross section is 49.5 \AA^2 . The difference in potential energy and cross section are minimal, which makes sense when considering the small size of the cysteine monomer. This particular plot was chosen for comparison with that of the proline monomer (Figure 3.3).

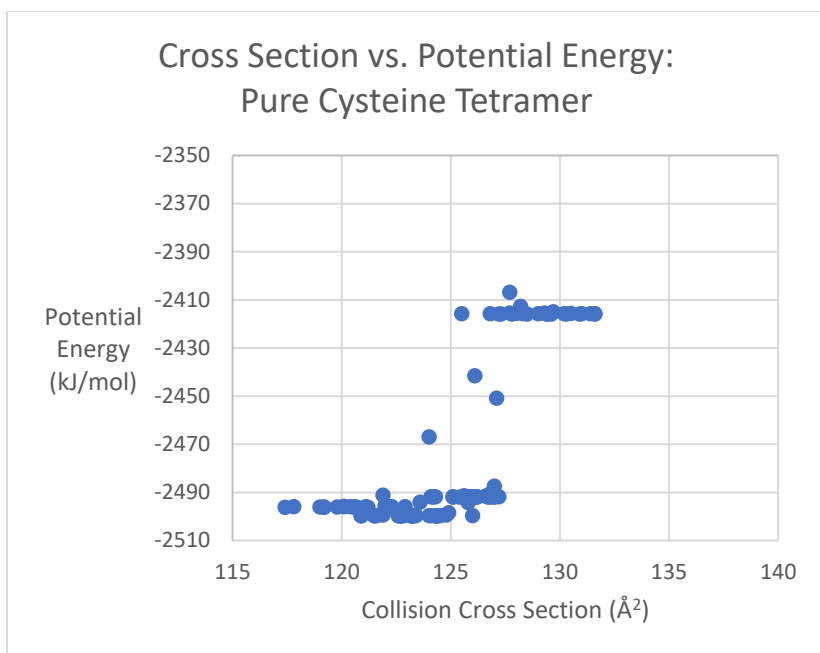


Figure A.17. A scatter plot of potential energy versus CCS for the pure cysteine tetramer. The average cross section is 125.0 \AA^2 . Two distinct regions can be seen in the scatter plot, corresponding to two distinct isomers.



Figure A.18. Computationally modeled structures for the pure cysteine tetramer, a more stable tetrahedral structure (left) and a higher-energy planar structure (right).

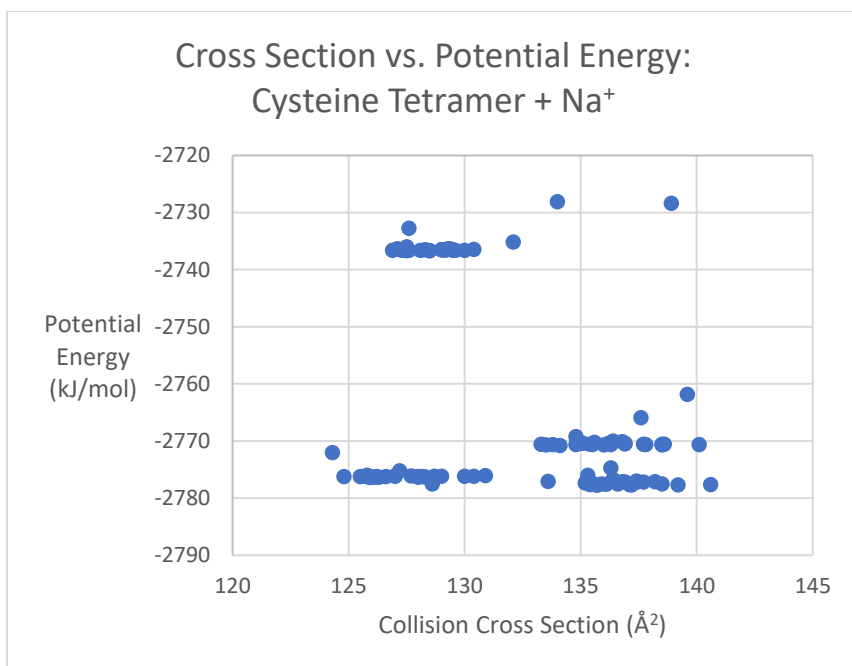


Figure A.19. A scatter plot of potential energy versus CCS for the sodium-complexed cysteine tetramer. The average cross section is 132.4 \AA^2 . The graph is more disordered than for pure cysteine, with a separation based on potential energy, as well as one based on cross section.

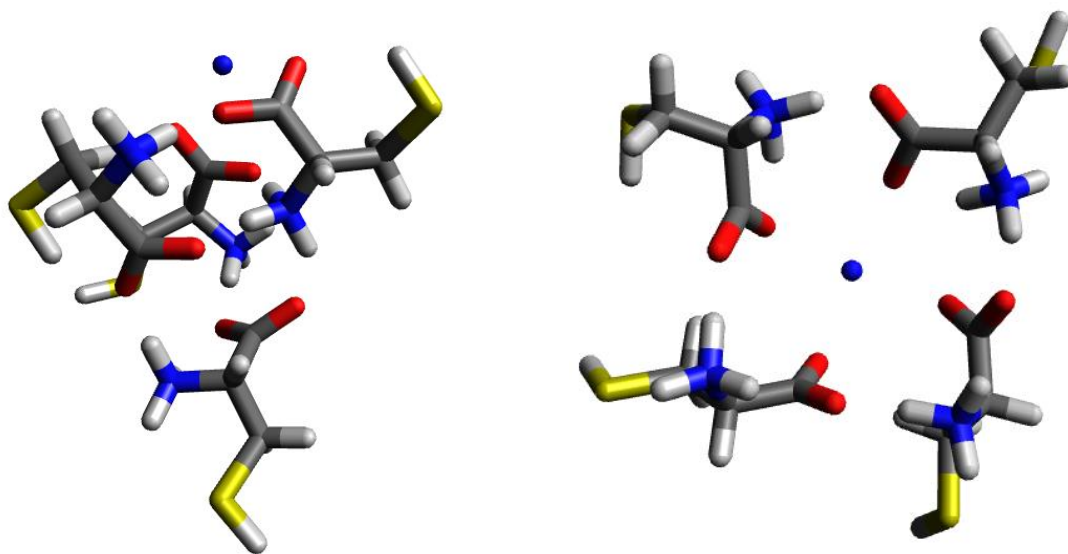


Figure A.20. Computationally modeled structures for the cysteine-sodium tetramer. The lower-energy structure corresponds to the roughly tetrahedral structure (left), while the planar structure with the sodium in the center (right) is higher energy, similar to what was observed for the pure cysteine tetramer.

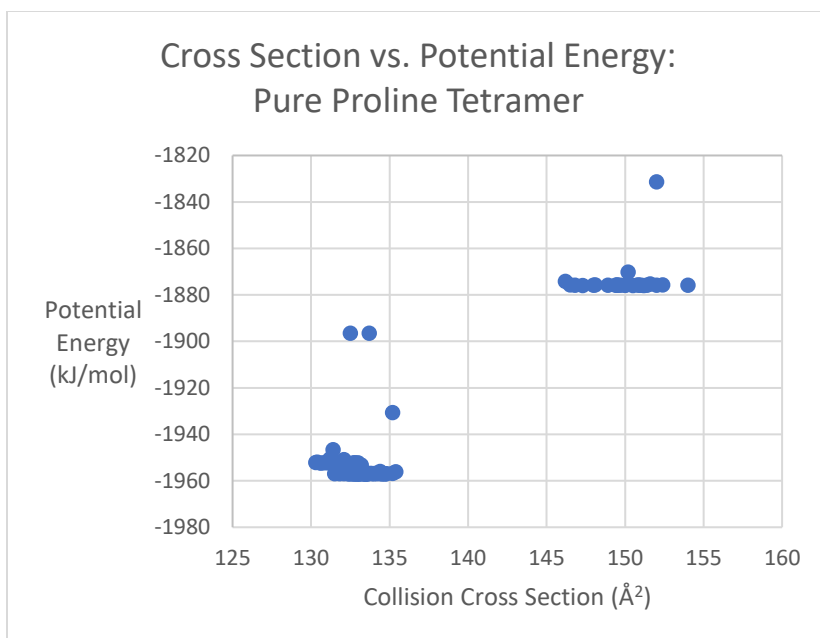


Figure A.21. A scatter plot of potential energy versus CCS for the pure proline tetramer. The average cross section is 137.0 \AA^2 . There are two clear regions in the scatter plot, corresponding to a high-energy, more extended isomer and a low-energy, compact isomer.

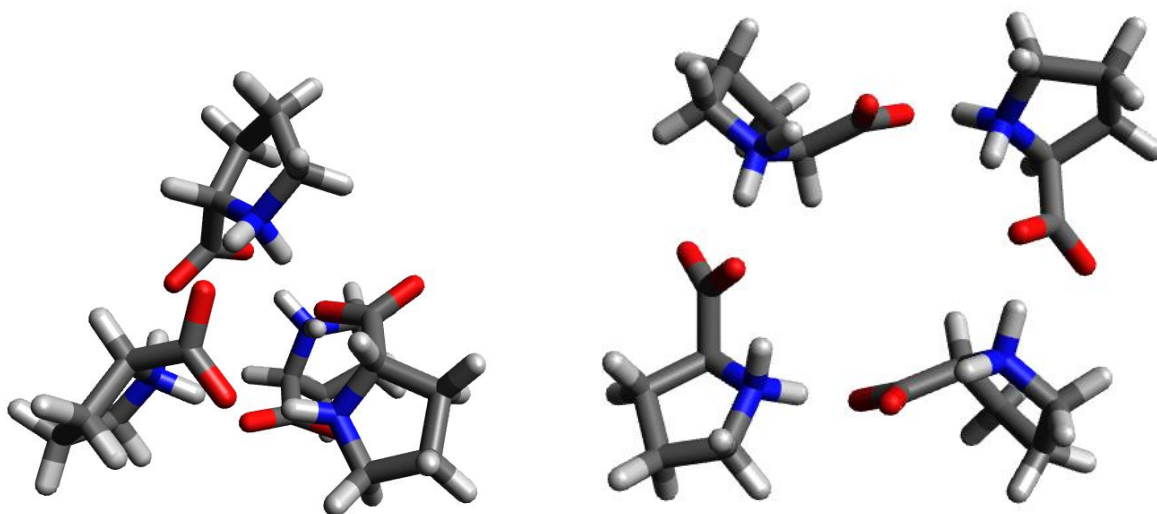


Figure A.22. Computationally modeled structures for the pure proline tetramer. Once again, we see the trend of the lower-energy and more stable tetrahedral structure (left) and the more extended planar ring structure (right), just as was observed for cysteine.

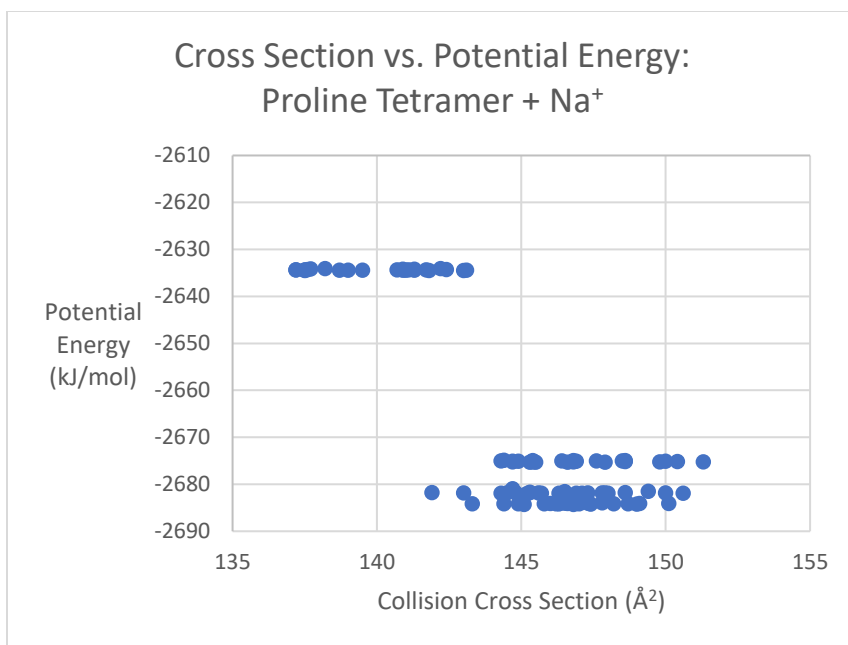


Figure A.23. A scatter plot of potential energy versus CCS for the sodium-complexed proline tetramer. The average cross section is 145.2 \AA^2 .

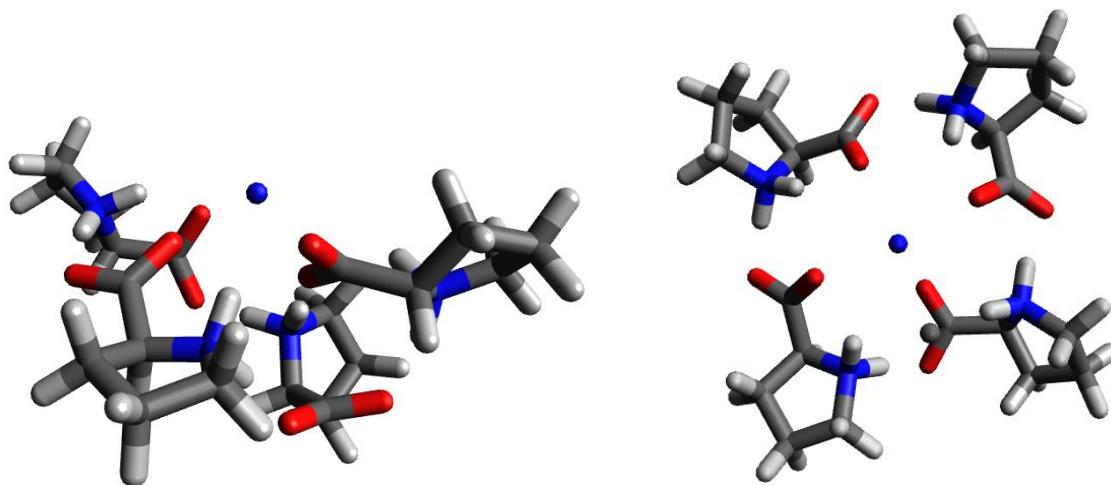


Figure A.24. Computationally modeled structures for the proline-sodium tetramer. Interestingly, proline appears to form a compact “cradle” structure with sodium, in addition to the normal ring structure. This “cradle” structure is even more unstable than the ring structure, which appears to be the preferred structure for proline energy-wise. The proline-sodium heteroligomer may be unable to adopt the stable tetrahedral structure observed for the pure proline tetramer, or perhaps it was simply not sampled in this particular simulated annealing trial.

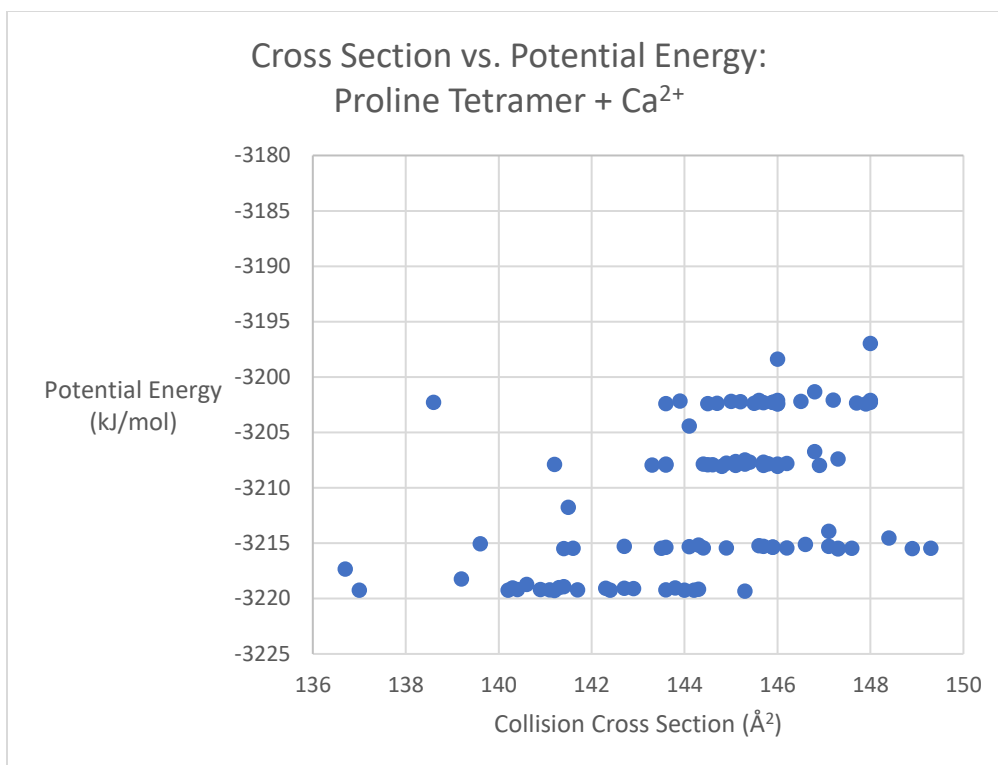


Figure A.25. A scatter plot of potential energy versus CCS for the calcium-complexed proline tetramer. The average cross section is 144.6 Å². Notably, compared to the other sodium heteroligomers, the scatter plot encompasses a number of different isomers, suggesting that when complexed with calcium, the resulting proline structures are of similar stability and no one structure is clearly preferred.

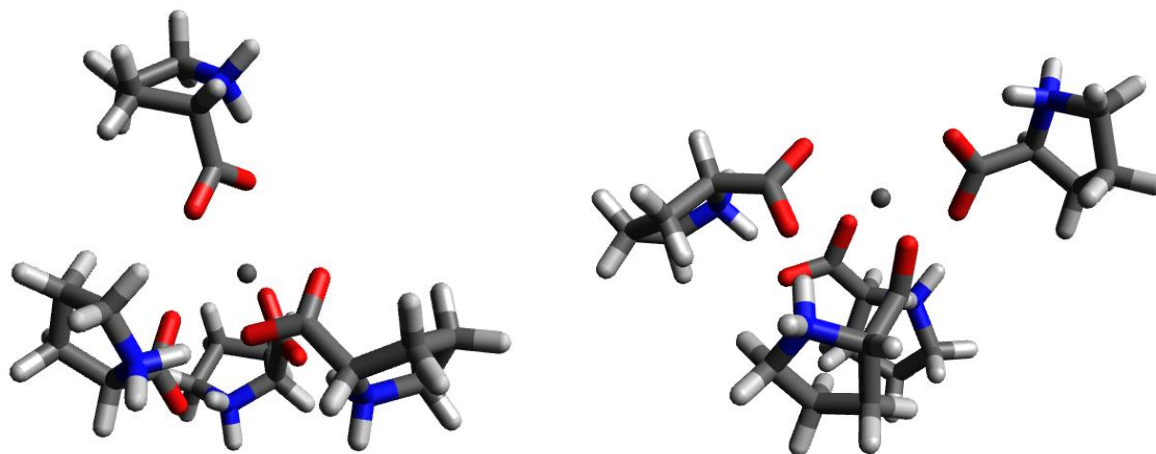


Figure A.26. Computationally modeled structures for the proline-calcium tetramer: a higher energy, roughly tetrahedral structure (left) and a lower energy cradle-like structure (right).

References

-
- ¹ Hardy, J., & Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*, 297 (5580), 353-356. <https://doi.org/10.1126/science.1072994>
- ² Karran, E., Mercken, M., & Strooper, B. D. (2011). The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nature Reviews Drug Discovery*, 10 (9), 698-712. <https://doi.org/10.1038/nrd3505>
- ³ Adler-Abramovich, L., Reches, M., Sedman, V. L., Allen, S., Tendler, S. J., & Gazit, E. (2006). Thermal and chemical stability of diphenylalanine peptide nanotubes: implications for nanotechnological applications. *Langmuir*, 22 (3), 1313-1320. <https://doi.org/10.1021/la052409d>
- ⁴ Busseron, E., Ruff, Y., Moulin, E., & Giuseppone, N. (2013). Supramolecular self-assemblies as functional nanomaterials. *Nanoscale*, 5(16), 7098-7140. <https://doi.org/10.1039/C3NR02176A>
- ⁵ Reches, M., & Gazit, E. (2005). Self-assembly of peptide nanotubes and amyloid-like structures by charged-termini-capped diphenylalanine peptide analogues. *Israel Journal of Chemistry*, 45 (3), 363-371. <https://doi.org/10.1560/5MC0-V3DX-KE0B-YF3J>
- ⁶ Hartmann, T., Bieger, S. C., Brühl, B., Tienari, P. J., Ida, N., Allsop, D, Roberts, G. W., Masters, C. L., Dotti, C. G., Unsicker, K., and Beyreuther, K. (1997). Distinct sites of intracellular production for Alzheimer's disease A β 40/42 amyloid peptides. *Nature Medicine*, 3 (9), 1016-1020. <https://doi.org/10.1038/nm0997-1016>
- ⁷ Murray, M. M., Bernstein, S. L., Nyugen, V. Y., Condrón, M. M., Teplow, D. B., & Bowers, M. T. (2009). Amyloid β protein: A β 40 inhibits A β 42 oligomerization. *Journal of the American Chemical Society*, 131 (18), 6316-6317. <https://doi.org/10.1021/ja8092604>
- ⁸ Bleiholder, C., Do, T. D., Wu, C., Economou, N. J., Bernstein, S. S., Buratto, S. K., Shea, J. E., and Bowers, M. T. (2013). Ion mobility spectrometry reveals the mechanism of amyloid formation of A β (25–35) and its modulation by inhibitors at the molecular level: epigallocatechin gallate and scyllo-inositol. *Journal of the American Chemical Society*, 135 (45), 16926-16937. <https://doi.org/10.1021/ja406197f>
- ⁹ Murray, M. M., Krone, M. G., Bernstein, S. L., Baumketner, A., Condrón, M. M., Lazo, N. D., Teplow, D. B., Wyttenbach, T., Shea, J. E. and Bowers, M.T. (2009). Amyloid β -Protein: Experiment and Theory on the 21– 30 Fragment. *The Journal of Physical Chemistry B*, 113 (17), 6041-6046. <https://doi.org/10.1021/jp808384x>

-
- ¹⁰ Shaham-Niv, S., Adler-Abramovich, L., Schnaider, L., & Gazit, E. (2015). Extension of the generic amyloid hypothesis to nonproteinaceous metabolite assemblies. *Science Advances*, 1 (7), e1500137. <https://doi.org/10.1126/sciadv.1500137>
- ¹¹ Chiti, F., & Dobson, C. M. (2009). Amyloid formation by globular proteins under native conditions. *Nature Chemical Biology*, 5 (1), 15-22. <https://doi.org/10.1038/nchembio.131>
- ¹² Fändrich, M., Fletcher, M. A., & Dobson, C. M. (2001). Amyloid fibrils from muscle myoglobin. *Nature*, 410 (6825), 165-166. <https://doi.org/10.1038/35065514>
- ¹³ Adler-Abramovich, L., Vaks, L., Carny, O., Trudler, D., Magno, A., Caflisch, A., Frenkel, D. and Gazit, E., (2012). Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. *Nature Chemical Biology*, 8 (8), 701-706. <https://doi.org/10.1038/nchembio.1002>
- ¹⁴ Singh, V., Rai, R. K., Arora, A., Sinha, N., & Thakur, A. K. (2014). Therapeutic implication of L-phenylalanine aggregation mechanism and its modulation by D-phenylalanine in phenylketonuria. *Scientific Reports*, 4 (1), 1-8. <https://doi.org/10.1038/srep03875>
- ¹⁵ Shaham-Niv et al., 2015. (see 10)
- ¹⁶ Mason, E. A., & McDaniel, E. W. (1988). *Transport properties of ions in gases* (Vol. 26). New York: Wiley.
- ¹⁷ National Institutes of Health. (n.d.). *Computational Modeling*. <https://www.nibib.nih.gov/science-education/science-topics/computational-modeling>
- ¹⁸ Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural and Molecular Biology*, 9 (9), 646-652. <https://doi.org/10.1038/nsb0902-646>
- ¹⁹ Karplus, M. (2002). Molecular Dynamics Simulations of Biomolecules. *Accounts of Chemical Research*, 35 (6), 321-323. <https://doi.org/10.1021/ar020082r>
- ²⁰ Chemes, L. B., Alonso, L. G., Noval, M. G., & de Prat-Gay, G. (2012). Circular dichroism techniques for the analysis of intrinsically disordered proteins and domains. In *Intrinsically Disordered Protein Analysis* (pp. 387-404). Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-61779-927-3_22
- ²¹ Van Laarhoven, P. J., & Aarts, E. H. (1987). Simulated annealing. In *Simulated Annealing: Theory and Applications* (pp. 7-15). Springer, Dordrecht. https://doi.org/10.1007/978-94-015-7744-1_2
- ²² Kirkpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220 (4598), 671-680. <https://doi.org/10.1126/science.220.4598.671>

-
- ²³ Bertsimas, D., & Tsitsiklis, J. (1993). Simulated annealing. *Statistical Science*, 8 (1), 10-15. <https://doi.org/10.1214/ss/1177011077>
- ²⁴ Ingber, L. (1993). Simulated annealing: Practice versus theory. *Mathematical and Computer Modelling*, 18 (11), 29-57. [https://doi.org/10.1016/0895-7177\(93\)90204-C](https://doi.org/10.1016/0895-7177(93)90204-C)
- ²⁵ Berendsen, H. J., van der Spoel, D., & van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91 (1-3), 43-56. [https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/10.1016/0010-4655(95)00042-E).
- ²⁶ GROMACS Development Team. (n.d.). *About GROMACS*. <https://www.gromacs.org/about.html>
- ²⁷ Robertson, M. J., Tirado-Rives, J., & Jorgensen, W. L. (2015). Improved peptide and protein torsional energetics with the OPLS-AA force field. *Journal of Chemical Theory and Computation*, 11 (7), 3499-3509. <https://doi.org/10.1021/acs.jctc.5b00356>
- ²⁸ Ponder, J. W., & Case, D. A. (2003). Force fields for protein simulations. *Advances in Protein Chemistry*, 66, 27-85. [https://doi.org/10.1016/S0065-3233\(03\)66002-X](https://doi.org/10.1016/S0065-3233(03)66002-X).
- ²⁹ Beachy, M. D., Chasman, D., Murphy, R. B., Halgren, T. A., & Friesner, R. A. (1997). Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. *Journal of the American Chemical Society*, 119 (25), 5908-5920. <https://doi.org/10.1021/ja962310g>
- ³⁰ Bleiholder, C., Wytenbach, T., & Bowers, M. T. (2011). A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). *Method. International Journal of Mass Spectrometry*, 308 (1), 1-10. <https://doi.org/10.1016/j.ijms.2011.06.014>
- ³¹ Bleiholder, C., Contreras, S., Do, T. D., & Bowers, M. T. (2013). A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (II). Model parameterization and definition of empirical shape factors for proteins. *International Journal of Mass Spectrometry*, 345, 89-96. <https://doi.org/10.1016/j.ijms.2012.08.027>
- ³² Anderson, S. E., Bleiholder, C., Brocker, E. R., Stang, P. J., & Bowers, M. T. (2012). A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (III): Application to supramolecular coordination-driven assemblies with complex shapes. *International Journal of Mass Spectrometry*, 330, 78-84. <https://doi.org/10.1016/j.ijms.2012.08.024>
- ³³ Bleiholder, C., Contreras, S., & Bowers, M. T. (2013). A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (IV). Application to polypeptides. *International Journal of Mass spectrometry*, 354, 275-280. <https://doi.org/10.1016/j.ijms.2013.06.011>

-
- ³⁴ Wyttenbach, T., Bleiholder, C., Anderson, S. E., & Bowers, M. T. (2015). A new algorithm to characterise the degree of concaveness of a molecular surface relevant in ion mobility spectrometry. *Molecular Physics*, *113* (15-16), 2344-2349. <https://doi.org/10.1080/00268976.2015.1042935>
- ³⁵ Wyttenbach, T., Kemper, P. R., & Bowers, M. T. (2001). Design of a new electrospray ion mobility mass spectrometer. *International Journal of Mass Spectrometry*, *212* (1-3), 13-23. [https://doi.org/10.1016/S1387-3806\(01\)00517-6](https://doi.org/10.1016/S1387-3806(01)00517-6)
- ³⁶ Wilm, M., & Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. *Analytical Chemistry*, *68* (1), 1-8. <https://doi.org/10.1021/ac9509519>
- ³⁷ Juraschek, R., Dülcks, T., & Karas, M. (1999). Nanoelectrospray—more than just a minimized-flow electrospray ionization source. *Journal of the American Society for Mass Spectrometry*, *10* (4), 300-308. [https://doi.org/10.1016/S1044-0305\(98\)00157-3](https://doi.org/10.1016/S1044-0305(98)00157-3)
- ³⁸ Hamley, I. W. (2012). The amyloid beta peptide: a chemist's perspective. Role in Alzheimer's and fibrillization. *Chemical Reviews*, *112* (10), 5147-5192. <https://doi.org/10.1021/cr3000994>
- ³⁹ Lotharius, J., & Brundin, P. (2002). Pathogenesis of Parkinson's disease: dopamine, vesicles and α -synuclein. *Nature Reviews Neuroscience*, *3* (12), 932-942. <https://doi.org/10.1038/nrn983>
- ⁴⁰ Stefanis, L. (2012). α -Synuclein in Parkinson's disease. *Cold Spring Harbor Perspectives in Medicine*, *2* (2), a009399. <https://doi.org/10.1101/cshperspect.a009399>
- ⁴¹ Glabe, C. G. (2008). Structural classification of toxic amyloid oligomers. *Journal of Biological Chemistry*, *283* (44), 29639-29643. <https://doi.org/10.1074/jbc.R800016200>
- ⁴² Shaham-Niv et al., 2015. (see 10)
- ⁴³ Blau, N., Van Spronsen, F. J., & Levy, H. L. (2010). Phenylketonuria. *The Lancet*, *376* (9750), 1417-1427. [https://doi.org/10.1016/S0140-6736\(10\)60961-0](https://doi.org/10.1016/S0140-6736(10)60961-0)
- ⁴⁴ Hanley, W. B. (2004). Adult phenylketonuria. *The American Journal of Medicine*, *117* (8), 590-595. <https://doi.org/10.1016/j.amjmed.2004.03.042>
- ⁴⁵ Surtees, R., & Blau, N. (2000). The neurochemistry of phenylketonuria. *European Journal of Pediatrics*, *159* (2), S109-S113. <https://doi.org/10.1007/PL00014370>
- ⁴⁶ van Spronsen, F. J., Blau, N., Harding, C., Burlina, A., Longo, N., & Bosch, A. M. (2021). Phenylketonuria. *Nature Reviews Disease Primers*, *7* (1), 1-19. <https://doi.org/10.1038/s41572-021-00267-0>

-
- ⁴⁷ Singh, V., Rai, R. K., Arora, A., Sinha, N., & Thakur, A. K. (2014). Therapeutic implication of L-phenylalanine aggregation mechanism and its modulation by D-phenylalanine in phenylketonuria. *Scientific Reports*, 4 (1), 1-8. <https://doi.org/10.1038/srep03875>
- ⁴⁸ Adler-Abramovich, L., Vaks, L., Carny, O., Trudler, D., Magno, A., Caflisch, A., Frenkel, D. and Gazit, E. (2012). Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. *Nature Chemical Biology*, 8 (8), 701-706. <https://doi.org/10.1038/nchembio.1002>
- ⁴⁹ Do, T. D., & Bowers, M. T. (2015). Diphenylalanine self assembly: novel ion mobility methods showing the essential role of water. *Analytical Chemistry*, 87 (8), 4245-4252. <https://doi.org/10.1021/ac5046774>
- ⁵⁰ Adler-Abramovich, L., Reches, M., Sedman, V. L., Allen, S., Tendler, S. J., & Gazit, E. (2006). Thermal and chemical stability of diphenylalanine peptide nanotubes: implications for nanotechnological applications. *Langmuir*, 22 (3), 1313-1320. <https://doi.org/10.1021/la052409d>
- ⁵¹ Richard, D. M., Dawes, M. A., Mathias, C. W., Acheson, A., Hill-Kapturczak, N., & Dougherty, D. M. (2009). L-tryptophan: basic metabolic functions, behavioral research and therapeutic indications. *International Journal of Tryptophan Research*, 2, IJTR-S2129. <https://doi.org/10.4137/IJTR.S2129>
- ⁵² Pakula, M. M., Maier, T. J., & Vorup-Jensen, T. (2017). Insight on the impacts of free amino acids and their metabolites on the immune system from a perspective of inborn errors of amino acid metabolism. *Expert Opinion on Therapeutic Targets*, 21 (6), 611-626. <https://doi.org/10.1080/14728222.2017.1323879>
- ⁵³ Shaham-Niv, S., Rehak, P., Vuković, L., Adler-Abramovich, L., Král, P., & Gazit, E. (2017). Formation of apoptosis-inducing amyloid fibrils by tryptophan. *Israel Journal of Chemistry*, 57 (7-8), 729-737. <https://doi.org/10.1002/ijch.201600076>
- ⁵⁴ Ahmed, M., Davis, J., Aucoin, D., Sato, T., Ahuja, S., Aimoto, S., Elliott, J.I., Van Nostrand, W.E. and Smith, S.O. (2010). Structural conversion of neurotoxic amyloid- β 1-42 oligomers to fibrils. *Nature Structural & Molecular Biology*, 17 (5), 561-567. <https://doi.org/10.1038/nsmb.1799>
- ⁵⁵ Bernstein, S.L., Dupuis, N.F., Lazo, N.D., Wytttenbach, T., Condrón, M.M., Bitan, G., Teplow, D.B., Shea, J.E., Ruotolo, B.T., Robinson, C.V. and Bowers, M.T. (2009). Amyloid- β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nature Chemistry*, 1 (4), 326-331. <https://doi.org/10.1038/nchem.247>
- ⁵⁶ Haass, C., & Selkoe, D. J. (2007). Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide. *Nature Reviews Molecular Cell Biology*, 8 (2), 101-112. <https://doi.org/10.1038/nrm2101>

-
- ⁵⁷ Baglioni, S., Casamenti, F., Bucciantini, M., Luheshi, L.M., Taddei, N., Chiti, F., Dobson, C.M. and Stefani, M. (2006). Prefibrillar amyloid aggregates could be generic toxins in higher organisms. *Journal of Neuroscience*, 26 (31), 8160-8167. <https://doi.org/10.1523/JNEUROSCI.4809-05.2006>
- ⁵⁸ Do, T. D., De Almeida, N. E., LaPointe, N. E., Chamas, A., Feinstein, S. C., & Bowers, M. T. (2016). Amino acid metaclusters: implications of growth trends on peptide self-assembly and structure. *Analytical Chemistry*, 88 (1), 868-876. <https://doi.org/10.1021/acs.analchem.5b03454>
- ⁵⁹ Dilger, J. M., Glover, M. S., & Clemmer, D. E. (2017). A database of transition-metal-coordinated peptide cross-sections: selective interaction with specific amino acid residues. *Journal of The American Society for Mass Spectrometry*, 28 (7), 1293-1303. <https://doi.org/10.1007/s13361-016-1592-9>
- ⁶⁰ Hoffmann, W., Chang, R., Marianski, M., Seo, J., Bowers, M.T., von Helden, G., and Pagel, K. (2018). Supramolecular Assembly of Proline Mediated by Alkali Metal Cations [Unpublished Manuscript].
- ⁶¹ Dray, A., & Perkins, M. (1993). Bradykinin and inflammatory pain. *Trends in Neurosciences*, 16 (3), 99-104. [https://doi.org/10.1016/0166-2236\(93\)90133-7](https://doi.org/10.1016/0166-2236(93)90133-7)
- ⁶² Wytttenbach, T., von Helden, G., & Bowers, M. T. (1996). Gas-phase conformation of biological molecules: bradykinin. *Journal of the American Chemical Society*, 118 (35), 8355-8364. <https://doi.org/10.1021/ja9535928>
- ⁶³ National Center for Biotechnology Information (2022). *PubChem Compound Summary for CID 5862, Cysteine*. Retrieved November 28, 2022 from <https://pubchem.ncbi.nlm.nih.gov/compound/Cysteine>.
- ⁶⁴ National Center for Biotechnology Information (2022). *PubChem Compound Summary for CID 145742, Proline*. Retrieved November 28, 2022 from <https://pubchem.ncbi.nlm.nih.gov/compound/Proline>.
- ⁶⁵ von Helden et al., 2018. (see 59)