




5-2020

Determination of Thermodynamic Properties of Non-Protein Amino Acids and Characterization of Multimers of Carbamazepine

Gwendylan Turner

Follow this and additional works at: <https://scholarworks.wm.edu/honorsthesis>

 Part of the [Analytical Chemistry Commons](#), [Medicinal-Pharmaceutical Chemistry Commons](#), and the [Other Chemistry Commons](#)

Recommended Citation

Turner, Gwendylan, "Determination of Thermodynamic Properties of Non-Protein Amino Acids and Characterization of Multimers of Carbamazepine" (2020). *Undergraduate Honors Theses*. Paper 1471. <https://scholarworks.wm.edu/honorsthesis/1471>

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Determination of Thermodynamic Properties of Non-Protein Amino Acids and Characterization of
Multimers of Carbamazepine

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Science in Department of Chemistry from
The College of William and Mary

by

Gwendylan A. Turner

Accepted for _____ Honors _____

(Honors, High Honors, Highest Honors)



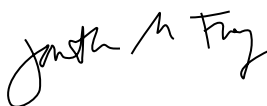
John C. Poutsma, Advisor



Robert D. Pike, Chemistry



Rachel E. O'Brien



Jonathan Frey

Williamsburg, VA

May 5, 2020

Abstract

This study has two seemingly unrelated parts that come together remarkably in displaying the comprehensive interplay between chemical structure and properties as well as the variety of analytical applications of mass spectrometry.

The first part of this study describes the determination of thermodynamic properties of several non-protein amino acids using the extended kinetic method. This is a continuation of work started in the Poutsma lab in Spring of 2017. The non-protein amino acids (NPA) studied here hold notable relevance in their unique ability to be mis-incorporated into peptide chains, as shown by the Hartman group at Virginia Commonwealth University.^[1] By understanding the effects of methylation on the NPA's inherent thermochemical properties, such as proton affinity and $\Delta_{\text{acid}}H$, we acquire insight into how these species may alter the behavior of the peptide chains in which they are incorporated. We found the experimental $\Delta_{\text{acid}}H$ of α -methylserine, L-penicillamine, and 3-methylthreonine to be 1379 ± 23 , 1380 ± 18 , and 1378 ± 23 kJ/mol respectively. Within bounds of reasonable uncertainty, these values agree with computational predictions done at the B3LYP/6-311++G**//B3LYP/6-31+G* level of theory.

The second part of this study examines the gas-phase tetramer of carbamazepine (CBZ), an active pharmaceutical ingredient in anticonvulsants.^[2] Highly polymorphic, CBZ is well-suited for studying the fundamentals of the self-assembly process in organic crystals, and more information on base-level assembly is required for effective predictive models of organic crystallization.^[2] Because typically only one polymorph of a drug is approved by the Food and Drug Administration as a pharmaceutical active ingredient, polymorphism is an important phenomena in the pharmaceutical industry. In this study, we used High-Field Asymmetric Ion Mobility Spectrometry and traditional mass spectrometry to characterize the tetramer of CBZ

and evaluate its relative stability. We confirmed that an intensity anomaly existed in both the protonated and sodiated forms in which the tetramer is larger than the trimer or the pentamer; the tetramer is a magic number cluster. This agrees with STM data taken of CBZ monolayers by our colleagues at Notre Dame.

Table of Contents

Chapter I: Introduction and Guide to Thesis	
1.1: Introduction and Application	1
1.2 Gas-phase Thermochemistry	2
1.3 Field Asymmetric Ion Mobility Spectroscopy	6
1.4 Computational Studies	11
1.5 Non-Protein Amino Acids	12
1.6 Carbamazepine and Polymorphism	13
Chapter 2: Procedures	16
2.1 Experimentally Determined Gas-Phase Acidities	16
2.2 Computational Predictions of Gas-Phase Properties	18
2.3 Clusters of Carbamazepine	20
Chapter 3: Results and Discussion	21
3.1 Experimentally Determined Quantities	
3.1.1 α -methylserine Results	21
3.1.2 3-Methylthreonine Results	24
3.1.3 L-Penicillamine Results	28
3.2 Computational Predictions	31
3.2.1 Serine	32
3.2.2 α -methylserine	33
3.2.3 3-methylthreonine	35
3.2.4 Cysteine	37
3.2.5 α -methylcysteine	39

3.2.6 L-penicillamine	41
3.3 Summarized Non-Protein Amino Acid Summarized Results	43
3.4 Carbamazepine Cluster Study	45
3.3.1 Tetramer Characterization	46
3.4.3 Ion Mobility Studies of Carbamazepine Clusters	51
<u>Chapter 4: Conclusion</u>	56

Figures and Tables

Figure 1.1: A Schematic of the Kinetic Method and Associated Rate Constants ^[13]	4
Figure 1.2: Competitive Dissociation of the Heterodimer ^[14]	4
Figure 1.3: A Variety of IMS Methods ^[19]	8
Figure 1.4: Illustration of Thermo Fisher's Asymmetric DF ^[22]	8
Figure 1.5: Illustrated Filtering Ability of FAIMS ^[21]	9
Figure 1.6: FAIMS Differentiation between Leucine and Isoleucine ^[26]	10
Figure 1.7: Sampling Conformers on a Potential Energy Surface ^[31]	12
Figure 1.8: Structures of Non-Protein Amino Acids of Interest	13
Figure 1.9: Structures of Carbamazepine and Dibenzazepine	14
Table 2.1: Reference Acids Used for Each Compound	16
Figure 3.1: Kinetic Method Plot 1: α -methylserine	21
Figure 3.2: Kinetic Method Plot 2: α -methylserine	22
Figure 3.3: Effective Temperature plot: α -methylserine	23
Figure 3.4 Orthogonal Distance Regression: α -methylserine	24
Figure 3.5: Kinetic Method Plot 1: 3-methylthreonine	25
Figure 3.6: Kinetic Method Plot 2: 3-methylthreonine	25
Figure 3.7: Effective Temperature plot: 3-methylthreonine	26
Figure 3.8 Orthogonal Distance Regression: 3-methylthreonine	27
Figure 3.9: Kinetic Method Plot 1: L-penicillamine	28
Figure 3.10: Kinetic Method Plot 2: L-penicillamine	29
Figure 3.11: Effective Temperature plot: L-penicillamine	30
Figure 3.12: Orthogonal Distance Regression: L-penicillamine	31

Figure 3.13: Illustrated Motifs of Amino Acid Hydrogen Bonding ³⁷	32
Figure 3.14: Low-Energy Conformer of Neutral Serine	32
Figure 3.15: Low-Energy Conformers of Deprotonated Serine	33
Figure 3.16: Lowest Energy Conformers of neutral α -methylserine	34
Figure 3.17: Lowest Energy Conformers of deprotonated α -methylserine	34
Figure 1.18: Lowest Energy Conformers of neutral 3-methylthreonine	35
Figure 2.19: Lowest Energy Conformers of deprotonated 3-methylthreonine	36
Figure 3.20: Lowest Energy Conformers of neutral cysteine	37
Figure 3.21: Lowest Energy Conformer of deprotonated cysteine	38
Figure 3.22: Lowest Energy Conformer of carboxylic-acid deprotonated cysteine	38
Figure 3.23: Lowest Energy Conformers of neutral α -methylcysteine	39
Figure 3.24: Lowest Energy Conformer of deprotonated α -methylcysteine	39
Figure 3.25: Lowest Energy Conformer carboxylic-acid deprotonated α -methylcysteine	40
Figure 3.26: Lowest Energy Conformers of neutral L-penicillamine	41
Figure 3.27: Lowest Energy Conformer of deprotonated L-penicillamine	42
Figure 3.28: Lowest Energy Conformer of carboxylic-acid deprotonated L-penicillamine	42
Figure 3.29: Studied Non-Protein Amino Acids	43
Table 3.1: A Summary of Computational and Experimental Quantities	44
Table 3.2: Literature Values for Glycine, Alanine, and Valine	44
Figure 3.30: Structure of Dibenzazepine	45
Figure 3.31: Structure of Carbamazepine	45
Figure 3.32: Raw Mass Spectrum of 10.00 mmol Carbamazepine	46
Figure 3.33: Magic Number Plot of 10.00 mmol Carbamazepine	47

Figure 3.34: Raw Mass Spectrum of 1.00 mmol Carbamazepine	47
Figure 3.35: Magic Number Plot of 1.00 mmol Carbamazepine	48
Figure 3.36: Raw Mass Spectrum of 0.10 mmol Carbamazepine	48
Figure 3.37: Magic Number Plot of 0.10 mmol Carbamazepine	49
Figure 3.38: Raw Mass Spectrum of 0.01 mmol Carbamazepine	49
Figure 3.39: Magic Number Plot of 0.01 mmol Carbamazepine	50
Figure 3.40: Compensation Voltage Scan of Protonated Carbamazepine Monomer	52
Figure 3.41: Compensation Voltage Scan of Protonated Carbamazepine Dimer	52
Figure 3.42: Compensation Voltage Scan of Protonated Carbamazepine Trimer	52
Figure 3.43: Compensation Voltage Scan of Sodiated Carbamazepine Trimer	52
Figure 3.44: Energy Minimized Model of Carbamazepine Trimers ^[40]	53
Figure 3.45: Compensation Voltage Scan of Protonated Carbamazepine tetramer	54
Figure 3.46: Compensation Voltage Scan of Sodiated Carbamazepine tetramer	54

Gratitude and Acknowledgements

I would like to give my first and biggest thanks to Dr. J.C. Poutsma for his enduring support throughout my years in the Poutsma Lab. I would also like to thank the Department of Chemistry at The College of William and Mary. They have been integral to my academic experience and are largely responsible for my success in the field as an undergraduate.

I am grateful for all my labmates in Ionlab, especially William Moeller and Alexis Brender A. Brandis, who contributed greatly to the lab environment, both intellectually and socially. I would like to thank Henry Cardwell who helped immensely with this project.

Finally, I greatly appreciate the National Science Foundation for funding this project.

Chapter I: Introduction and Guide to Thesis

1.1: Introduction and Application

Attempting to quantify and understand the processes that characterize the meaning of life, biochemistry investigates the intersection between life and matter. It is the perpetual quest to determine how non-living molecules arrange support for beings as simple as bacteria or as complex as the assembly of organs reading this thesis. Naturally, the easiest place to start is somewhere simple, and hundreds of years have been dedicated to characterizing these starting points. One such point is the study of amino acids, the substructures of proteins and peptides. There are 20 protein amino acids regularly coded for in biological processes, but hundreds of non-protein amino acids (NPAAs) are also useful in biological operations such as in metabolism and defense.^{[3],[4]}

The Poutsma lab has a history of studying non-protein amino acids and evaluating the differences in thermochemical properties that arise as a result of relatively small structural changes such as side chain-shortening^[5], alteration of ring size^[6], and systematic atom substitutions^[7]. Through these studies, we attempt to reveal the subtle relationship between structure, energetics, and reactivity.

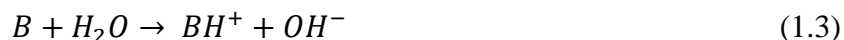
In a related project, we are interested in the intrinsic stabilities of gas-phase polymorphic molecules and their connection to their varying crystalline forms^{[2], [8]}. This is especially important in the pharmaceutical industry because identical compounds with different crystalline structures have a variable bioavailability, solubility, and reactivity^[2]. Often times, only one polymorph of an active pharmaceutical ingredient is approved for consumption. Thus, understanding what governs the assembly of crystalline structures and what properties lead to polymorphic activity is a relevant, active field of research.

1.2 Gas-phase Thermochemistry

Proton transfer reactions continue to be important in a fundamental understanding of chemistry and biochemistry alike;^[9] knowledge of the acidity and basicity of species is integral in fully grasping bonding behavior, molecular geometries, possible reaction mechanisms, and a variety of other important molecular events. Studying proton transfer reactions in the gas phase is advantageous in comparing inherent properties of similar molecules due to the lack of solvation effects and other intermolecular interactions.^[10] This lack of solvation, however, does change the scale of acidity and basicity. Traditional solvent-based acidity and basicity are scaled on a pK_a and pK_b scale, respectively. These scales are determined by the following reactions and equations:^[11]



$$pK_a = \log (K_a) = \log \frac{[A^-][H_3O^+]}{[AH]} \quad (1.2)$$



$$pK_b = \log (K_b) = \log \frac{[BH^+][OH^-]}{[B]} \quad (1.4)$$

Here, Equation 1.1 and Equation 1.3 show the definition of a traditional acid and base where water is the solvent. The traditional solvent-based acid and base scales revolve around an equilibrium constant, or a ratio of product species to that of parent molecule. While it may be straightforward to measure the concentration of H^+ and derive the quantities of interest in solution, this is not always the case in the gas phase. For this reason, acidity and basicity are instead scaled around the energy required for dissociation of the parent molecule. With this scale, the Bronsted-Lowry definition of proton donating and accepting are preserved, as can be seen in the associated chemical reactions in equations 1.5, 1.7, 1.9, and 1.10. ^{[10], [11]}

In the gas phase, the *proton affinity* (PA) is defined as the negative enthalpy of the protonation reaction described in Equation 1.5. In other words, it is equivalent to the heat of deprotonation of species BH^+ . Similarly, the *gas-phase basicity* (GB) is the negative Gibbs free energy change associated with the reaction (1.5).^[10]



$$\Delta G_{rxn} = -GB; \Delta H_{rxn} = -PA[B] \quad (1.6)$$

The *gas-phase acidity* (GA) of a molecule, AH, is the Gibbs free energy change ($\Delta_{acid}G$) of the following reaction (Equation 1.7) at 298K. The enthalpy of this reaction, $\Delta_{acid}H$, is thus the proton affinity of the associated anion, A^- .^{[10],[11]}



$$\Delta_{acid}G = GA; \Delta_{acid}H = PA[A^-] \quad (1.8)$$

When trying to experimentally determine these gas-phase properties, studying the above hypothetical reactions (1.5, 1.7) is difficult and often impractical. Thankfully, the quantities in question can still be extracted by looking at *relative acidity* and *relative basicity*, which refers to the Gibbs energy change of the reactions described in equations 1.9 and 1.10 respectively.^[10] When using this relationship to extract the mentioned quantities, there will be a reference compound with a known acidity or basicity, and the equilibrium constant, K_{eq} , is extracted. ΔG can then be derived through the standard thermochemical manipulation $\Delta G = -RT \ln(K_{eq})$. In the following equations, A_i and B_i refer to a molecule of interest, while A_r and B_r refer to their respective reference compounds. This method of determining gas-phase thermochemical properties of molecules is called the equilibrium method.



A requirement of the equilibrium method is the ability to determine the equilibrium constant which involves accurate measurement of the concentrations (partial pressures) of the neutral molecules A_i and A_r . When attempting to measure thermochemical properties for non-volatile neutrals, such as amino acids and peptides, other methods must be used.

An elegant method for evaluating these properties for non-volatile species is the kinetic method, created by Cooks and Kruger.^[12] The kinetic method studies thermochemical properties of molecules in the gas phase using rates of competitive dissociation of cluster ions, as illustrated in Figure 1.1.^[13] This method is accomplished using a mass spectrometer, where cluster ions can be formed directly using a variety of ion sources and then can easily be mass selected and dissociated. A mass spectrometer conveniently simplifies sample preparation, as the mass filtering of the heterodimer filters out impurities. Since this thesis deals mostly with acidity, we will examine the kinetic method in this context.

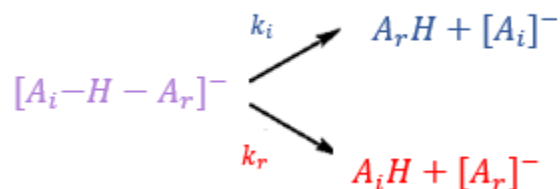


Figure 1.1: A Schematic of the Kinetic Method and Associated Rate Constants^[13]

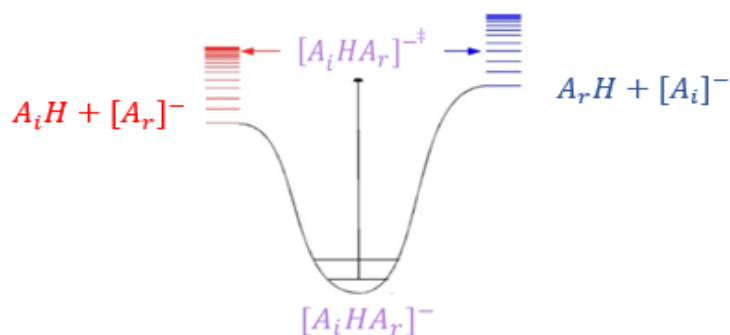


Figure 1.2: Competitive Dissociation of the Heterodimer^[14]

In these figures, it is important to note that $[A]^-$ represents a deprotonated species, which would have an m/z value of $M-1$ in the mass spectrum, where M is the molecular weight of the neutral amino acid or reference acid. Meanwhile, AH is a neutral species which takes the bridging proton of the cluster after dissociation. As expected for a neutral molecule, this species does not appear in the mass spectrum.

The simplest version of the kinetic method assumes that the rates of the dissociation, k_i and k_r (Figure 1.1), are solely dependent on the relative enthalpy of each dissociation channel, visually seen as the difference in the enthalpy of the dissociation asymptotes on either side of the energy well pictured in Figure 1.2. Using this logic, the populations of the resulting ions should follow a Boltzmann distribution using the relevant enthalpy. This is exemplified in equation 1.11:^[13]

$$\ln \left(\frac{k_2}{k_1} \right) \approx \ln \left(\frac{[A_r]^-}{[A_i]^-} \right) \approx \frac{\Delta\Delta H}{RT_{eff}} \quad (1.11)$$

In this equation, $\Delta\Delta H$ represents the difference in $\Delta_{acid}H$ between the reference and compound of interest. However, this method ignores differences in the entropy requirements for the competitive channels and reverse activation energies that may be relevant.^[13] In reality, $\Delta\Delta G$, or the difference in Gibbs free energy between the two dissociation channels, should be in the numerator of the kinetic method equation, as shown in Equation 1.12. Examining the Gibbs free energy equation^[11] in Equation 1.13, we can start to devise methods to separate the enthalpic and entropic components of dissociation:

$$\ln \left(\frac{k_2}{k_1} \right) \approx \ln \left(\frac{[A_r]^-}{[A_i]^-} \right) \approx \frac{\Delta\Delta G}{RT_{eff}} \quad (1.12)$$

$$\Delta G = \Delta H - T\Delta S \quad (1.13)$$

$$\ln \left(\frac{k_z}{k_i} \right) \approx \ln \left(\frac{[A_r]^-}{[A_i]^-} \right) \approx \frac{\Delta\Delta H}{RT_{eff}} + \frac{\Delta\Delta S}{T_{eff}} \quad (1.14)$$

In Equation 1.13, the spontaneity of the reaction, (ΔG) is affected by the enthalpy of the reaction (ΔH) as well as the product of the temperature (T) and the entropy (ΔS). The assumptions of the simple kinetic method give a good approximation for the relative enthalpy difference until entropic effects begin to greatly affect preference of one direction over another.^{[15],[16],[17]} Accounting for entropy is done effectively using the extended kinetic method, largely developed by Fenselau, Wesdemiotis, and others.^{[16],[17]} This method utilizes different collision energies to isolate the associated entropy by exploiting the effective temperature dependence in the entropy term.^{[13],[15]} In this method, the effective temperature for each collision energy is derived from the slope of the linear fits^[15] as discussed in the procedure (Section 2.2).

It is important to note here that the entropy derived from this type of analysis is the difference in *apparent* transition state entropy of the two dissociation processes. It is not equivalent to the difference in thermodynamic protonation entropy for the two species.^{[17],[18]} In addition, there is also statistical uncertainty associated with the derivation of each effective temperature, which leads to uncertainties in derived enthalpy and entropy values. Despite the uncertainty in the approximations, the extended kinetic method is still effective in determining the quantities of interest within a reasonable uncertainty using the proper statistical analysis.^{[16],[18]}

1.3 Field Asymmetric Ion Mobility Spectrometry

The ability of mass spectrometry to conveniently combine with separation techniques expands the diversity and depth of properties available for examination. Ion mobility is one such property that can be analyzed by incorporating ion mobility spectrometry (IMS) within the sample introduction of ESI-MS. Ion mobility is useful in characterizing different conformations

of ions of the same mass to charge ratio because ions with identical structure but different mobilities suggest a measurable change in shape between the two.

Ion Mobility is the gas-phase analog to electrophoresis. Though the dependencies are not completely analogous, ion mobility is still a type of *electric mobility*, which is the ability of a charged particle to move through a medium.^[19] Experimentally, an ion's mobility, K , is nothing more than the proportionality factor between field strength, E , and drift velocity, v , of an ion through a medium:^{[19], [20]}

$$K = \frac{v}{E} \quad (1.15)$$

Ion mobility depends on several things, such as the mass of the ion, the mass of the surrounding media, temperature of the surroundings, charge of the ion, and collisional cross section.

Ion mobility spectrometry (IMS) refers to the principles, methods, and instrumentation involved in characterizing ion mobility.^[19] Ion mobility spectrometry has a variety of different methods and specialized instrumentation. The realm of different IMS techniques as well as where Thermo Fisher Scientific FAIMS falls in this space is given in figure 1.3.^[19]

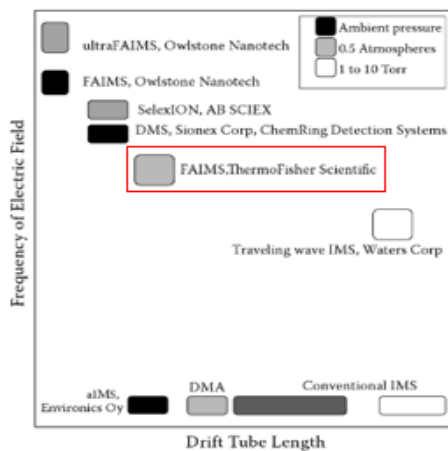


Figure 1.3: A variety of IMS methods. For context, the instrumentation used for this experiment was ThermoFisher Scientific FAIMS^[19]

In conventional IMS, the instrument consists of a drift tube with a uniform electric field, and the amount of time an ion takes to traverse a drift tube is measured and used to calculate ion mobility.^{[20],[21]} This is very similar to the concept of a time-of-flight mass spectrometer, and if conventional drift-cell IMS is the time-of-flight analog, Field Asymmetric Ion Mobility Spectrometry (FAIMS) is the quadrupole analog. Through the application of an alternating electric field, also known as a dispersion field (DF), such as one shown in Figure 1.4, an asymmetric field filters the ions allowed into the mass spectrometer by moving ions perpendicular to their trajectory towards the source.^[20] This filtering ability is shown in Figure 1.5.

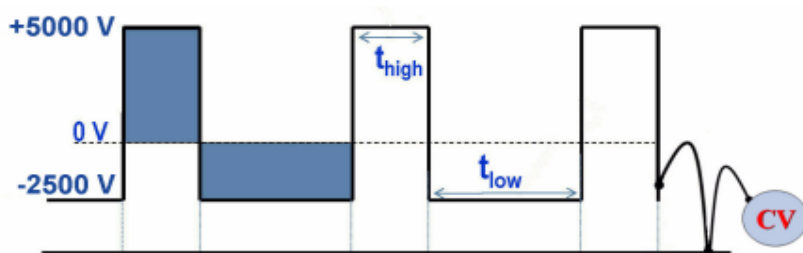


Figure 1.4: Illustration of Thermo Fisher's Asymmetric Alternating Current as a function of time ^[22]

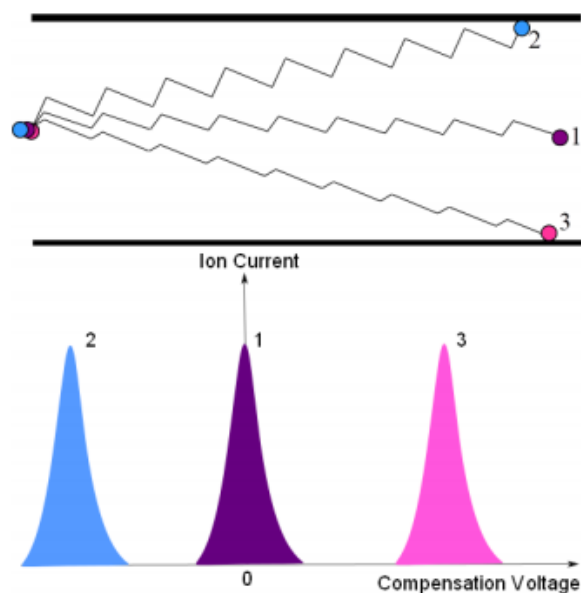


Figure 1.5: Illustrated Filtering Ability of FAIMS and the relationship to compensation voltage ^[21]

Though this dispersion field does not change, a compensation voltage holds the power of altering the drift correction to allow ions of different mobilities to pass through the filter. In the illustration (Figure 1.5), ion 2 requires a compensation voltage opposite in sign to ion 3 because it hits the plate in the opposite direction.

Within FAIMS, there are both low and high electric fields utilized. With low electric fields, the mobility is heavily dependent on the mass of the particle, and it thus provides little separation for close m/z values.^{[21],[23]} For this reason, we use High-Field Asymmetric Ion Mobility Spectrometry. In this type of spectrometry, the velocity is a function of the difference between high- and low-field mobility of the ion.^[21] This makes ion mobility more difficult to predict because instead of relying mostly on mathematically absolute quantities, a variety of factors come into play such as: ion declustering with drift gas and thermal heating of drift gas at high field strengths, orientation within the field, field-induced structural changes, and collision-induced structural change.^{[21],[23]}

Regardless of the complexity of these factors, they prove useful in enhancing the effect of cross section on mobility and thus improving separation capabilities. High-Field Asymmetric Ion Mobility Spectrometry is sensitive enough to distinguish even small structural differences beyond cross section, such as between enantiomer complexes^{[24],[25]} and structural isomers, such as leucine (Leu) and isoleucine (Ile), as shown by Purves et al in Figure 1.6.^[26]

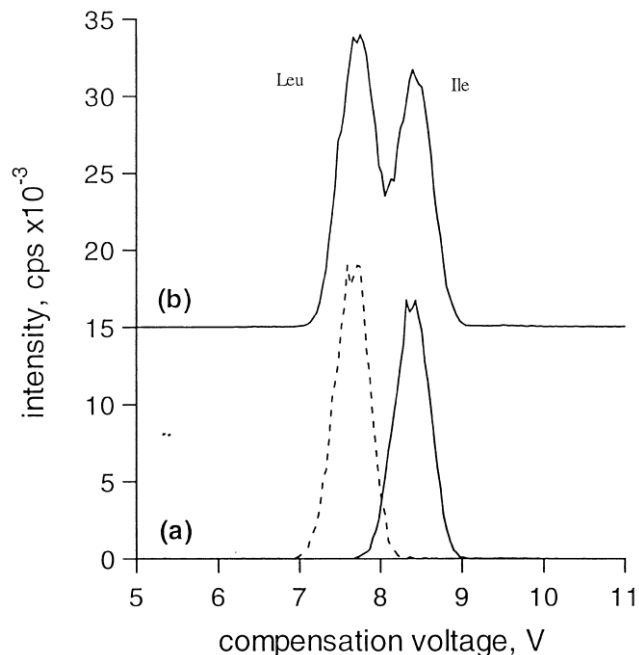


Figure 1.6: FAIMS peak identification using single-component standard solutions: (a) IS–CV spectra of 5 μ M Leu (dashed trace) and 5 μ M Ile (solid trace). (b) IS–CV spectrum of 5 μ M mixture of Leu and Ile.

Mobility changes between high and low fields may increase separation, but these changes also come with unfortunate decreases in signal, highly in part to the factors differential ion mobility relies on. As mentioned, collision-induced structural changes can occur at high-field strengths as a result of collisions between the ion and drift gas. With weakly-bound ion clusters, these collisions often have enough energy to cause fragmentation.

Due to the separation FAIMS provides, it can be useful in characterizing unique ion cluster arrangements. In particular, distinguishing multiple mobilities of stable gas-phase ion arrangements can lend a perspective to the assembly of different crystalline structures built around the same multimer. This is an interesting perspective to consider in the study of the multimers of CBZ.

1.4 Computational Studies

Computational predictions are only as accurate as the assumptions under which they operate. The prediction methods used in this thesis utilize many theories and areas of computational chemistry that have been proven useful in determining properties of similar molecules.^{[27],[28],[29]} These theories and areas include but are not limited to: molecular mechanics, Hartree-Fock theory and density functional methods. Using these methods, we aim to find the energy of the global minimum structure of a molecule of interest on its potential energy surface so that we can derive useful energetic information. For example, from the quantitative comparison of energetics between the lowest energy conformers of an amino acid and its deprotonated form, we can derive a prediction for the gas phase acidity.

Because of its empirical nature, molecular mechanics is an efficient calculation that serves as an adequate starting point in involved calculations. Molecular mechanics is a branch of computational chemistry that attempts to calculate molecular energies through the implementation of a molecular force field. This force field consists of energy penalties dependent on strain and sterics and is based on experimentally compiled data. The MMX forcefield used in this study was built mostly for organic species and includes, among other factors, a correction term for the anharmonicity of bonds.^[30] It is important to note that this method of energy calculation is based on the heat of formation of the molecule and thus is dependent on the numbers and types of atoms within the molecule. For this reason, we cannot make comparisons between neutral, protonated, and deprotonated species using strictly this method.^[30] However, we can compare the energies of different conformations of a given molecule. By allowing a molecule to move within an established force field, we can start to generate a list of low energy structures. These initial structures are used as starting points for more sophisticated calculations.

Generating this list of conformers is like sampling a potential energy surface at a variety of places. In two dimensions, this would be something like Figure 1.7^[31]

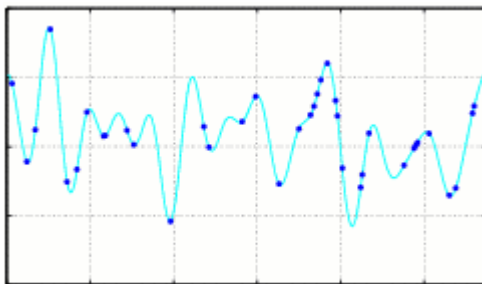


Figure 1.7: A visual representation of sampling various conformers on a potential energy surface^[31]

After generating a list of these sampled conformers, we then use more sophisticated methods to find optimized geometries with the hope that enough structures have been sampled to have at least one structure reach the global minimum. Since the wavefunctions of a multibody molecule cannot be solved analytically, other methods have been developed to address this. For our experiments, we use increasingly complex calculations involving a mix of Hartree-Fock and density functional theory.

1.5 Non-Protein Amino Acids

In choosing which molecules to study using experimental and computational methods, we look for a variety of properties: availability, relationship to structures with established thermochemical properties, and relevance to current research. In collaboration with the Hartman group at Virginia Commonwealth University, the thermodynamic properties of the following four non-protein amino acids were studied:

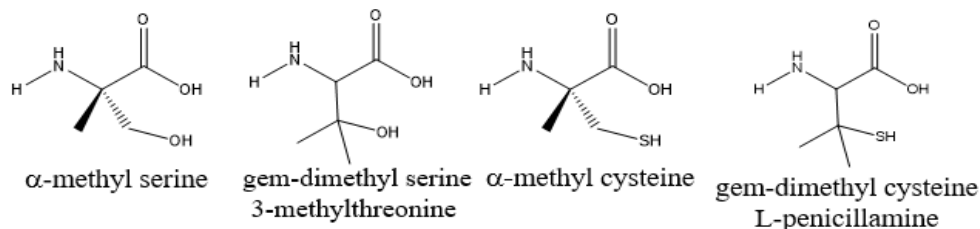


Figure 1.8: Non-Protein Amino Acids of Interest

These compounds were chosen because of their ability to be misincorporated into peptide chains using a promiscuous valine-t-RNA synthetase.^[1] Knowing how these slight modifications affect gas-phase properties, such as proton affinity and gas-phase acidity, may give some insight into how these species may alter the behavior of the peptide chains in which they are incorporated.

The proton affinities of the non-protein amino acids studied here were determined in a past study by Danielle Long and other members of Poutsma lab. In this connected study, the experimentally determined and computationally predicted proton affinities were found to be in excellent agreement with each other, and it was found that methylated non-protein amino acids had a higher proton affinity than their respective protein amino acid counterparts.^[32] In this thesis, the gas-phase acidities of the four NPAs are presented. Results for the NPA acidity study are presented in Chapter 3.

1.6 Carbamazepine and Polymorphism

The major over-arching goal of the research in the Poutsma lab is to uncover the interplay between molecular structure, energetics and reactivity. While the first study in this thesis centers on the connection between structure and thermochemical properties of biological molecules, the second study focuses on the relationship between the structure of small gas-phase clusters of carbamazepine and its crystalline assemblies. Carbamazepine (CBZ) is especially interesting because it forms multiple solid-state crystalline structures, or polymorphs. The scrutiny of CBZ is well paired with its analog, dibenzazepine because it isolates the role of the amide group in clustering activity.

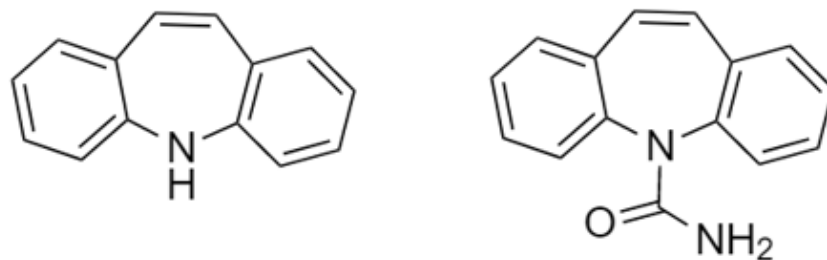


Figure 1.9: Structure of Dibenazepine (left) and Carbamazepine (right)

In chemistry, polymorphism refers to the ability of a solid material to exist in more than one form or crystal structure.^[33] This definition is extended to organic crystalline structures, which are of obvious interest in the pharmaceutical industry. Even though a compound may be pure in structure, different polymorphs exhibit variable solubility constants and thus variable bioavailability.^{[2],[34]} In many cases, only one polymorph of a compound is effective as an active pharmaceutical ingredient, and there is currently a limited amount of predicting power in the self-assembling nature of organic crystalline structures. For this reason, studies at the fundamental level of self-assembly are required.^[2]

Electrospray-mass spectrometry (ESI-MS) of gas-phase clusters serves as a good companion technique to microscopic and diffraction studies of this phenomena in the solid state. Although there is no imaging associated with ESI-MS, it has historically been integral in identifying metastable clusters in the gas phase through a variety of studies.^{[2],[35]} These clusters have characteristically high relative intensities when compared to neighboring peaks, suggesting an increased level of stability. Studying clusters in the gas phase is advantageous because it provides insight into intrinsic assembly interaction, improving our fundamental understanding of self-assembling molecules and providing new, helpful information to the field of organic-crystal self-assembly prediction.

Carbamazepine (CBZ) is an active pharmaceutical ingredient that has shown to be highly polymorphic, with five known anhydrous polymorphs.^{[2],[34]} This tendency makes this molecule a special interest to the study of the mentioned fundamental interactions, and our collaborators at the University of Notre Dame took an interest in this molecule and its analog, dibenzazepine (DBZ) as part of a larger study on crystal polymorphism.^[2]

They studied CBZ's clustering behavior with scanning tunneling microscopy and mass spectrometry. In their STM studies, they found that solution deposition of CBZ resulted in an overwhelming presence of networked tetramers in the monolayer on an Au(111) surface, while the DBZ monolayer possessed no apparent order. The stability of the CBZ tetramer was supported by the uncharacteristically large intensity of the sodiated tetramer in their mass spectrum, but they were unable to acquire data on the protonated tetramer.^[2] In section 2.3 , experiments on the protonated and sodiated cluster ions of CBZ are described.

Chapter 2: Procedures

2.1 Experimentally Determined Gas-Phase Acidities

All compounds were purchased in their pure form and stored appropriately until sample preparation. Stock solutions of the compounds of interest and reference acids were prepared in pure methanol with a 10^{-3} M concentration. For each sample, the stock solutions of the compound of interest and reference acid were mixed in a 3:2 ratio to promote heterodimer formation. Each solution containing α -methylserine or 3-methylthreonine was then spiked with 10% aqueous ammonia for deprotonation. Solutions for L-penicillamine and α -methylcysteine were spiked with only 5% ammonium hydroxide due to the compounds' tendency to make disulfide bonds. Unfortunately, α -methylcysteine was unstable in the analysis solution so it was excluded in the gas-phase acidity study. The list of reference compounds used for experimentation is as seen in

Table 1.

Reference Acids Used for Each Compound				
Reference	$\Delta_{\text{acid}}\text{H}$ kJ/mol	<i>α-methylserine</i>	<i>L-penicillamine</i>	<i>3-methylthreonine</i>
2-nitrobenzoic acid	1388		X	
4-hydroxybenzophenone	1390	X	X	X
3-nitrophenol	1392	X	X	X
3-trifluoromethylbenzoic acid	1399	X	X	X
3-fluorobenzoic acid	1406		X	
4-fluorobenzoic acid	1410	X		X

Table 2.1: Reference Acids Used for Each Compound. An 'X' denotes that the reference acid was successfully used for the determination of the GA of the corresponding non-protein amino acid.

Prepared samples were then injected using an external syringe pump at a rate of 60 $\mu\text{L}/\text{hour}$ into an external electrospray ionization source on a Finnigan TSQ Quantum Ultra quadrupole mass spectrometer. Upon formation of the heterodimer, the mass spectrometer focusing conditions were optimized to maximize the signal for the heterodimer before MS/MS analysis. In this analysis, the heterodimer was isolated using the first quadrupole and then allowed to collide in the second quadrupole with argon gas at collision energies ranging from 3

to 30 eV in intervals of 3eV. Spectra in negative ion mode of the fragments from each collision energy were averaged over one minute (approximately 115 scans), normalized for total ion count, and then repeated on three separate days. Fragmentation spectra of the monomers of both reference acids and compounds of interest were also obtained and averaged, which helped identify secondary fragmentation products. The mass and intensity of each product ion was then exported to Microsoft Excel®. Using the spectrum of each monomer, secondary fragmentation peaks of notable intensity were assigned to either the reference acid or compound of interest. After this assignment, the ratio of summed intensities (reference/amino acid) at each collision energy was averaged across the four trials.

The $\Delta_{\text{acid}}H$ of each reference acid was recorded and then averaged with the other reference acids used for any given compound of interest. This served to normalize the relative nature of the data to a usable scale. The ratio of the reference (A_r) to the interest (A_i) was determined, and the natural logarithm of this ratio was plotted as a function of $\Delta_{\text{acid}}H_r - \Delta_{\text{acid}}H_{\text{avg}}$ to give Kinetic Method (KM) plot 1. For reference, plot 1 represents Figures 3.1, 3.5, and 3.9. From the relationships described in Equations 2.3 we can derive that the slope, m , of the line of best fit through each collision energy is equivalent to:

$$\Delta G = \Delta H - T\Delta S \quad (2.1)$$

$$\ln \ln \left(\frac{k_2}{k_1} \right) = \ln \ln \left(\frac{[A_r]^-}{[A_i]^-} \right) = \frac{\Delta\Delta H}{RT_{\text{eff}}} + \frac{\Delta\Delta S}{T_{\text{eff}}} \quad (2.2)$$

$$m = -\frac{1}{RT_{\text{eff}}} \quad (2.3)$$

From this, we were able to derive the T_{eff} at each collision energy, giving an understanding into entropic effects in each case. We also plotted the negative intercept of the lines of best fit versus their slopes giving KM plot 2. The slope of plot 2 is equal to the relative

$\Delta_{\text{acid}}H$ of the compound of interest. Deviations in linearity of KM Plot 2 at both low and high collision energies are sometimes observed, and these data points at these energies are omitted in later steps as a result. At low energies, dissociation of at least one product is often disproportionately low due to an activation energy between the asymptotes. At high energies, product collection efficiency decreases for smaller, fast moving molecules. The molecules are also sometimes fragmented into species that are outside of the detection range of the instrument, compromising the integrity of the observed ratio.

In theory, all lines in the KM plot 1 should cross at an isothermal point with an x-coordinate representing the relative $\Delta_{\text{acid}}H$ of the compound of interest and the y-coordinate equal to $\Delta S/R$. This point is initially estimated from KM plot 2 using the methods described and is used as a starting point for a Monte-Carlo based orthogonal distance regression analysis developed by Ervin and Armentrout.^{[16],[18]} In this analysis, a crossing point is marked as a constraint for each line of best fit and effective temperatures are determined to best fit the data. The uncertainty for the enthalpy and entropy of the isothermal point is obtained from a Monte Carlo analysis in which random error is introduced to the points within established uncertainties. Generally, there is an uncertainty of approximately 8kJ/mol in the acidity scales of the references, and this uncertainty ultimately serves as the x-axis uncertainty value in the orthogonal distance regression analysis.

2.2 Computational Predictions of Gas-Phase Properties

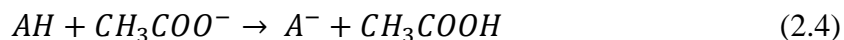
An initial conformation of each the molecule of interest was built in PC Model 9[®], and a list of starting low-energy conformers was generated by random rotation about single bonds and subsequent re-optimization using the MMX force field using GMMX searching routine. All conformations within 40 kJ/mol of the global minimum from the GMMX search were retained.

The resulting PCModel output file was parsed and converted to separate “.com” files that are compatible with the Gaussian09 software.

The resulting files were submitted to Gaussian, and lowest energy conformers were found using increasingly complex computations. Dividing the calculation into separate calculations at increasing levels of complexity is more time-efficient because it both eliminates redundancies in the sampling and starts the most complex calculations at a structure lower in energy than the initial parsed samples. The first-level calculation was done at the rhf/3-21G level of theory, which stands for restricted Hartree-Fock with the associated basis set. This outputs “*.log” files where duplicate energies are dismissed, and this list is submitted to the second level calculation at the B3LYP/3-21G level of theory. At this level, the basis set remains the same, but the theory is a hybrid of density functional theory and Hartree-Fock. This same method is present at the third and fourth level calculations, B3LYP/6-31+G* and B3LYP/6-311++G**. These calculations expand the basis sets to the d-orbital for second row atoms and incorporate polarization and diffuse functions into the calculation.

Ultimately, 298 K enthalpies were obtained at the B3LYP/6-311++G**//B3LYP/6-31+G* level of theory. These steps were repeated with the deprotonated species of each amino acid. For L-penicillamine and α -methylcysteine, the energy for deprotonation at the SH side group was also investigated and compared to the carboxylic acid deprotonation energies. Derived free-energy values were then used to determine a Boltzmann weighted population at 298K and ultimately a 298K enthalpy for each amino acid. The difference in Boltzmann-weighted enthalpy between the neutral and deprotonated amino acid was determined and combined with the 298K enthalpy of a proton (4.184×1.48 kJ/mol) to give a raw $\Delta_{\text{acid}}H$. The B3LYP/6-311++G** method gives a prediction for $\Delta_{\text{acid}}H$ of acetic acid that is ~ 10 kJ/mol too low. We therefore use reaction

2.2.1 to give an isodesmic prediction for the acidity of the amino acids using the experimental acidity of 1456.0 kJ/mol for acetic acid:^[36]



2.3 Clusters of Carbamazepine

For this experiment, solutions of carbamazepine in concentrations of 0.01 mM, 0.1 mM, 1.0mM, and 10 mM were prepared in methanol with 1% formic acid. Plastic storage vessels were used to minimize unintentional ion-exchange with the glass, which forms sodiated ions. These samples were then injected using an external syringe pump at a rate of 10 μ L/min into an external electrospray ionization source on a Finnigan TSQ Quantum Ultra quadrupole mass spectrometer with a capillary temperature of 120 $^\circ$ C. Mass spectra for each concentration with collision-induced dissociation of clustered peaks were taken, and this was repeated with trace amounts of sodium chloride added for the sake of comparison with the sodiated peaks observed by our Notre Dame colleagues. These experiments were repeated with dibenzazepine.

For FAIMS experiments, the ESI source was replaced with the FAIMS source, and the samples were injected with the same syringe pump at 60 μ L/hour. The compensation voltage was optimized for each of the studied masses, and a compensation voltage scan spectrum was obtained for each mass.

Chapter 3: Results and Discussion

3.1 Experimentally Determined Quantities

3.1.1 α -methylserine Results

In each sample, stock solutions of α -methylserine and a reference acid were mixed in a 3:2 ratio and spiked with 10% ammonium hydroxide. The reference acids for this analysis were 4-hydroxybenzophenone, 3-nitrophenol, 3-trifluoromethylbenzoic acid, and 4-fluorobenzoic acid.

Using an external syringe pump at a rate of 60 $\mu\text{L}/\text{hour}$, prepared samples were injected into an external electrospray ionization source on a Finnigan TSQ Quantum Ultra quadrupole mass spectrometer operating in negative ion mode. The heterodimer was isolated using the first quadrupole and then allowed to collide in the second quadrupole with argon gas at collision energies ranging from 3 to 30 eV in intervals of 3 eV. Spectra of the resulting fragments from each collision energy were averaged over a minute and normalized for total ion count. This process was repeated over three separate collection periods. Fragmentation spectra of α -methylserine and each reference acid and compounds of interest were also obtained and averaged, which helped identify secondary fragmentation products. After assigning peaks to their relevant parent ion, the ratio of summed intensities (reference/analyte) at each collision energy was averaged across the four trials.

Upon data collection, we plotted the natural log of the ratios between the reference and α -methylserine ($\ln([Ar]/[Ai])$) as a function of relative acidity ($\Delta_{\text{acid}}H_{\text{avg}} - \Delta_{\text{acid}}H_r$) of the reference (Kinetic Method plot 1) at each collision energy, as seen in Figure 3.1.

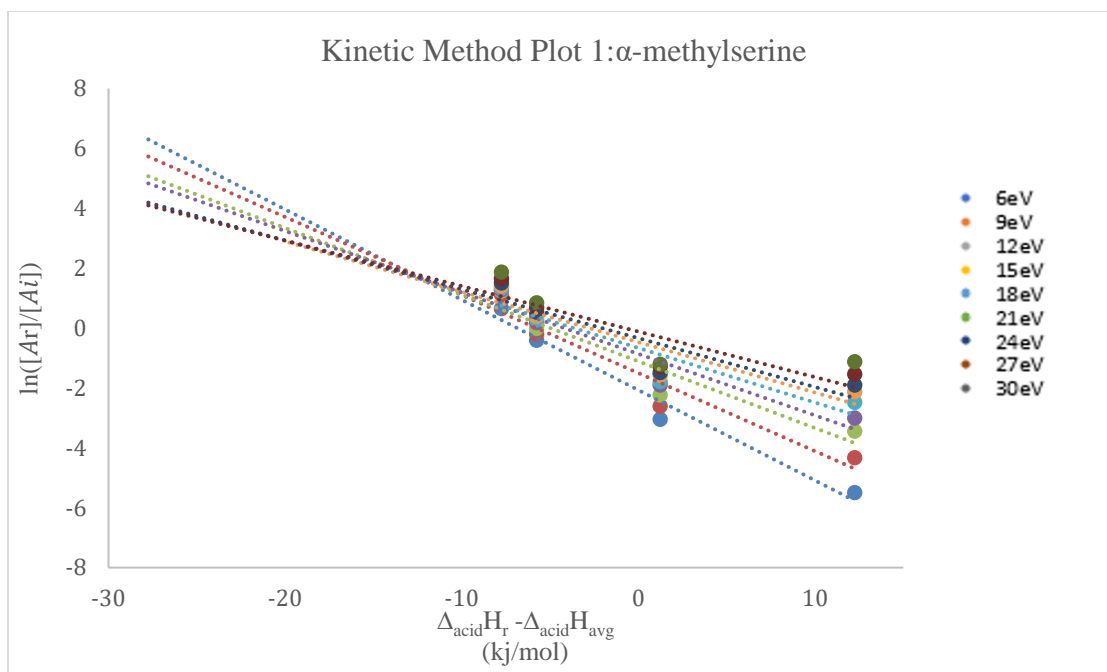


Figure 3.1: Kinetic Method Plot 1: α -methylserine

From the preliminary KM plot 1 shown in Figure 3.1, we calculated the slope and intercept of each line and plotted the negative intercepts as a function of their slopes to give KM Plot 2.

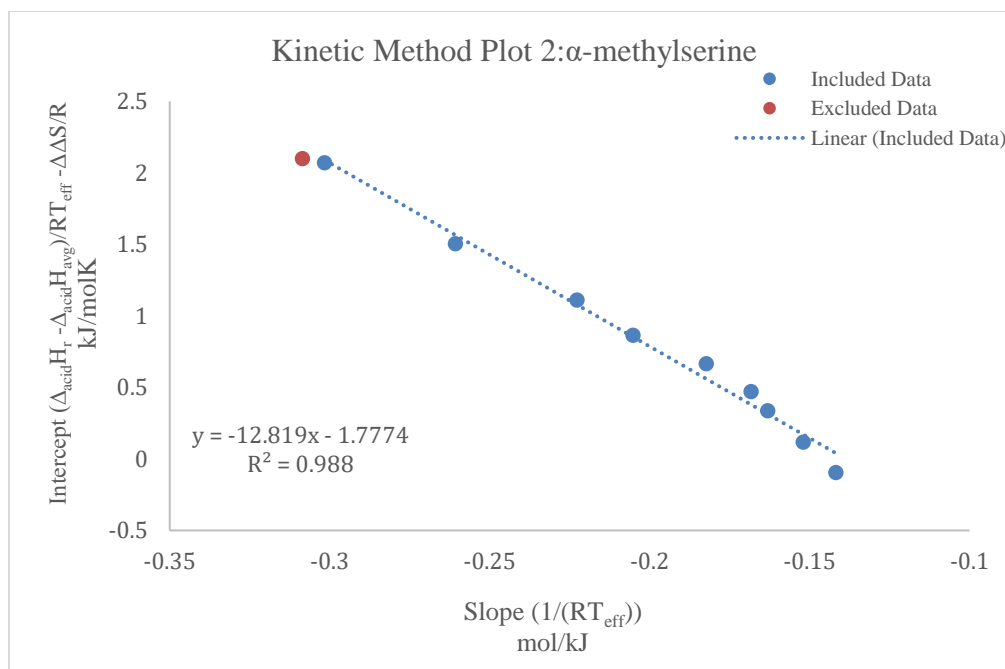


Figure 3.2: KM Plot 2 of α -methylserine study

From the slope of each collision energy line in Figure 3.1, we derived an effective temperature for each collision energy using Equation 2.3. We plotted the effective temperature as a function of collision energy to address deviations in linearity and exclude appropriate collision energies in our analysis as shown in Figure 3.3. For α -methylserine, the data taken at the collision energy of 3eV was excluded from further analysis due to deviation from linearity.

When excluding data appropriately in KM plot 2 (Figure 3.2), the slope gives a relative acidity of -12.8 kJ/mol, and the intercept gives a $\frac{\Delta\Delta S}{R}$ value of 1.77. When adjusting these values using the average reference acidity and appropriate value for the gas constant, R, the preliminary prediction gives a value of 1385 kJmol⁻¹ for $\Delta_{\text{acid}}H$ and -14.8 Jmol⁻¹K⁻¹ for the apparent entropy.

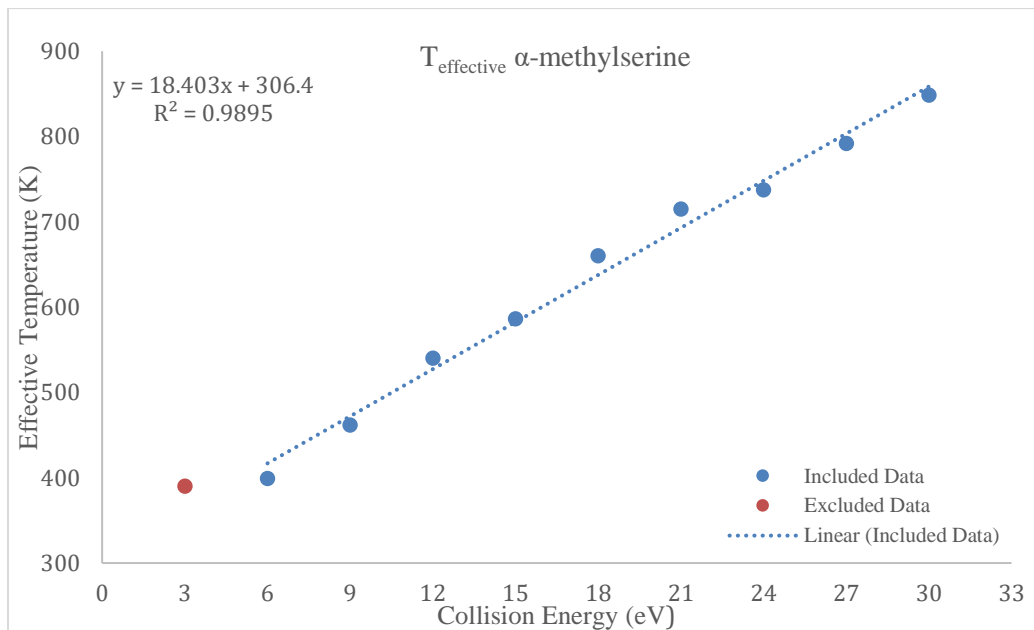


Figure 3.3: Effective Temperature Plot of α -methylserine study

Preliminary predictions of the slope and intercept were used as starting points for the Orthogonal Distance Regression (ODR) analysis. The mentioned collision energies and all reference acids within the analysis were included in the final orthogonal-distance-regression analysis because no reference acid showed excessive deviation from expected trends, as seen in

the preliminary plot 1 (Figure 3.1). The ODR analysis gave a $\Delta_{\text{acid}}H$ value of $1379 \pm 23 \text{ kJmol}^{-1}$ and an apparent entropy value of $-26 \pm 41 \text{ Jmol}^{-1}\text{K}^{-1}$. Figure 3.4 shows the original data with the ODR-derived effective temperature lines and the final isothermal point used to extract ΔH and ΔS .

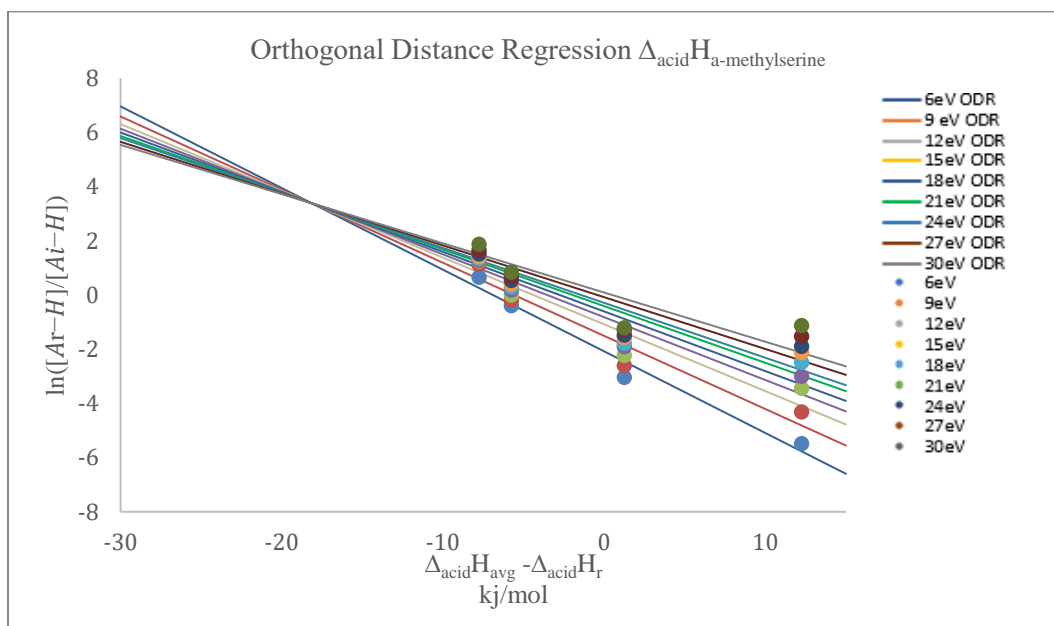


Figure 3.4: Orthogonal Distance Regression of α -methylserine

The experimental $\Delta_{\text{acid}}H$ for α -methylserine is significantly lower than that of serine ($\Delta_{\text{acid}}H = 1393 \text{ kJ/mol}$), its protein amino acid analog. This means that the methylated amino acid is more acidic in the gas phase.

3.1.2 3-Methylthreonine Results

Samples were prepared, injected, and analyzed as discussed in section 3.1.1. The references for this analysis were: 4-hydroxybenzophenone, 3-nitrophenol, 3-trifluoromethylbenzoic acid, and 4-fluorobenzoic acid.

The preliminary KM plot 1 was generated using the same methods as for α -methylserine and is shown in figure 3.5.

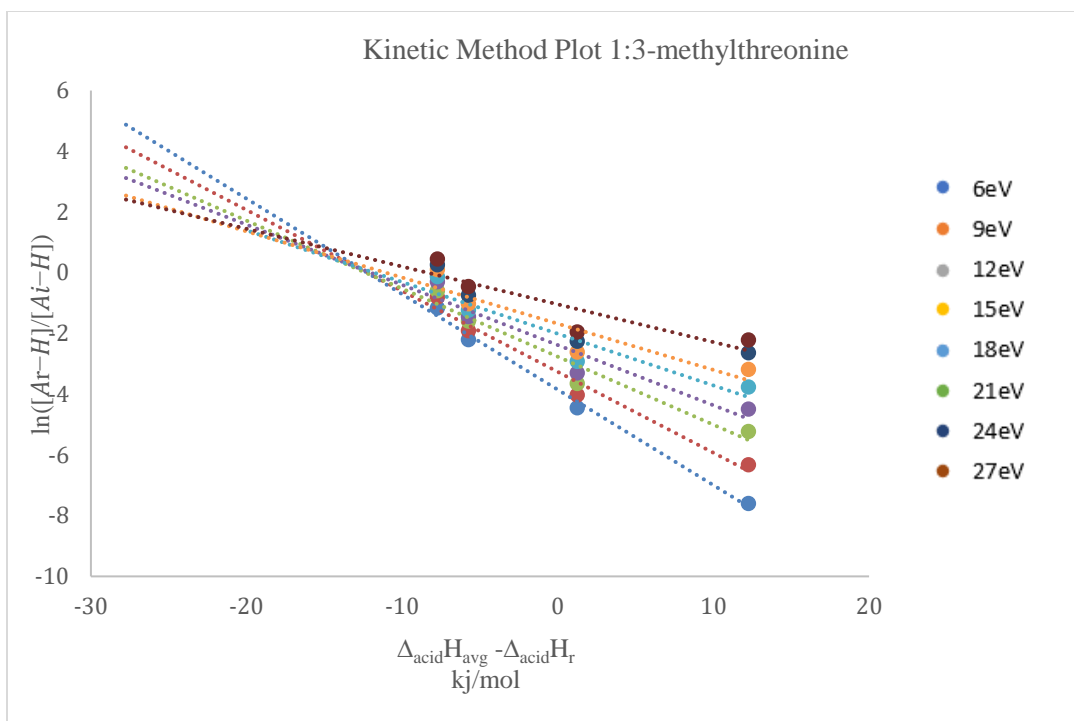


Figure 3.5: Preliminary KM plot 1 for 3-methylthreonine study

From the preliminary KM plot 1 shown in Figure 3.5, we calculated the slope and intercept of each line and plotted the intercept as a function of the slope to give KM plot 2 shown in Figure 3.6.

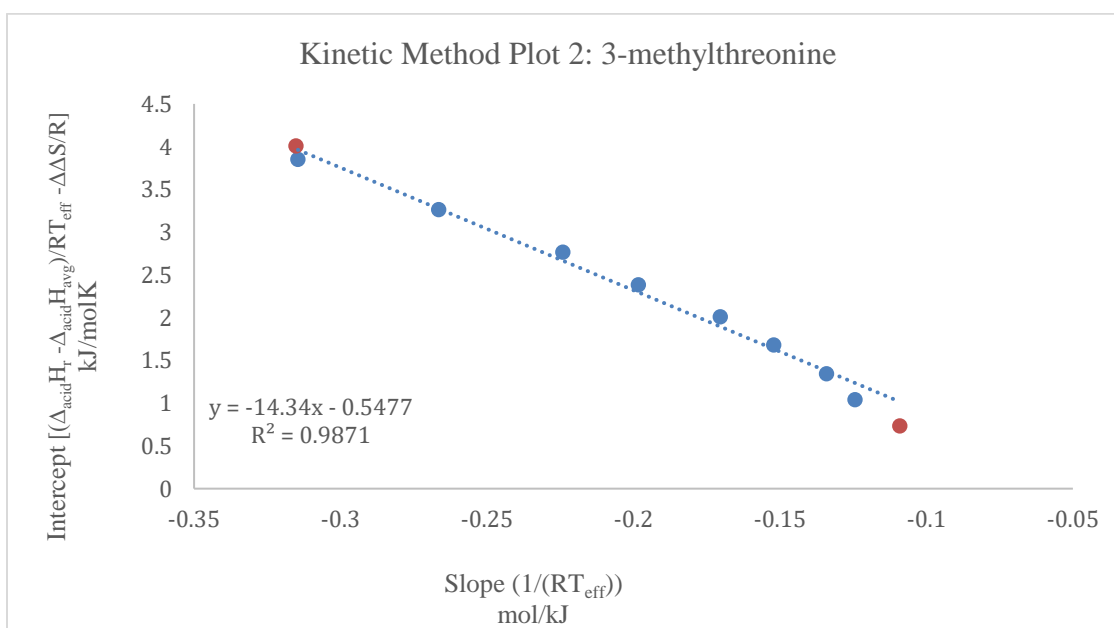


Figure 3.6: KM Plot 2 of 3-methylthreonine study

After deriving the effective temperature for each collision energy using Equation 2.3, we examined deviations from linearity and excluded appropriate collision energies in our analysis. For 3-methylthreonine, the data taken at the collision energies of 3eV and 30 eV were excluded due to the obvious deviation from linearity in the effective temperature plot (Figure 3.7). All reference acids within the collection were included in the final orthogonal-distance-regression analysis because no reference acid showed excessive deviation from expected trends, as seen in the preliminary plot (Figure 3.8).

When excluding data appropriately in the slope vs intercept plot (Figure 3.6) and applying the proper mathematical operations, the preliminary prediction gives a value of 1383 kJmol^{-1} for $\Delta_{\text{acid}}H$ and $-4.6 \text{ Jmol}^{-1}\text{K}^{-1}$ for $\Delta\Delta S$.

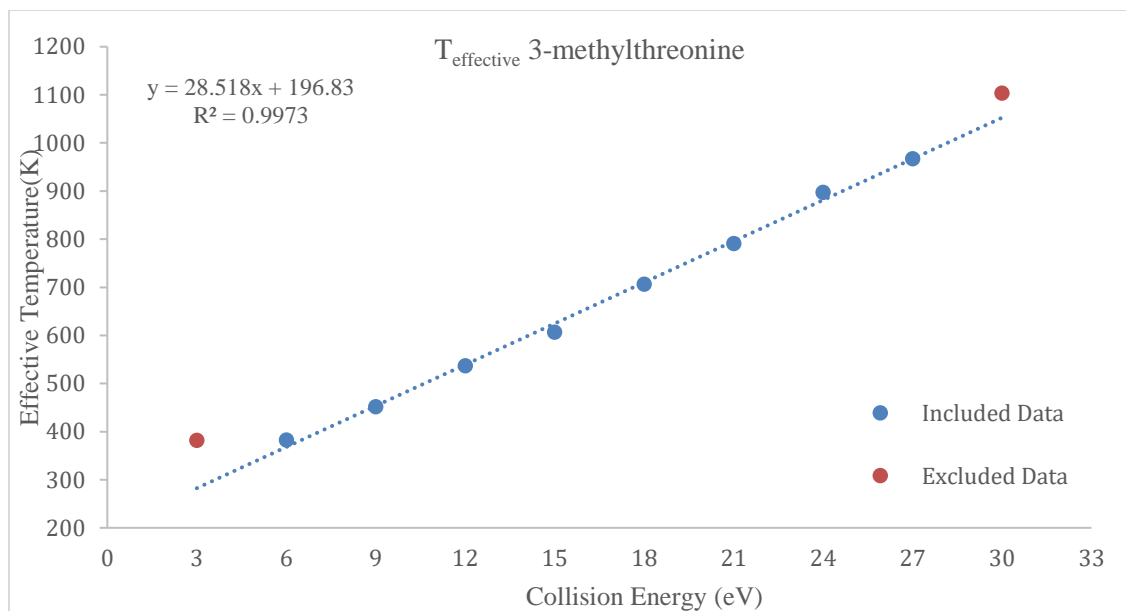


Figure 3.7: Effective Temperature plot for 3-methylthreonine study

Preliminary predictions of ΔH and ΔS were used as starting points for the Orthogonal Distance Regression (ODR) analysis. All reference acids within the analysis were included in the final orthogonal-distance-regression analysis because no reference acid showed excessive deviation from expected trends, as seen in the preliminary plot (Figure 3.5). The ODR analysis

gave a $\Delta_{\text{acid}}H$ value of $1378 \pm 23 \text{ kJmol}^{-1}$ and an apparent entropy value of $-14 \pm 37 \text{ Jmol}^{-1}\text{K}^{-1}$.

Figure 3.7 shows the original data with the ODR-derived effective temperature lines and the final isothermal point used to extract ΔH and ΔS .

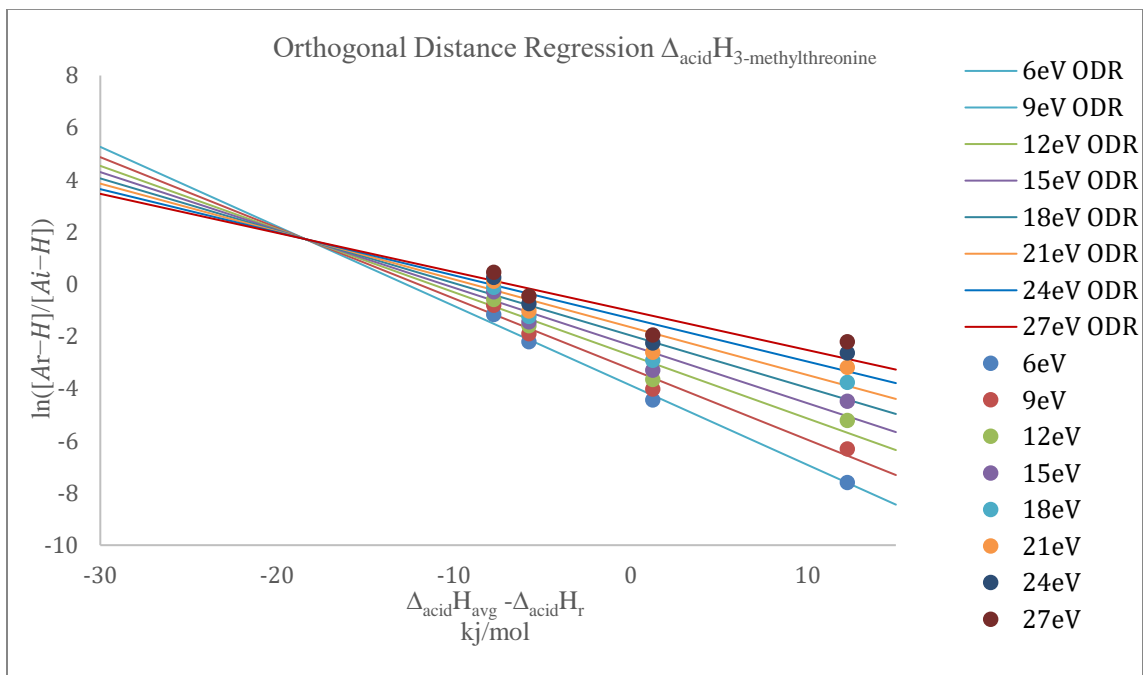


Figure 3.8: Orthogonal Distance Regression Analysis for 3-methylthreonine

The experimental $\Delta_{\text{acid}}H$ for 3-methylthreonine (gem-dimethylserine) is also significantly lower than that of serine (1393 kJ/mol), its analog. This means that the methylated amino acid acts more acidic than the protein amino acid in the gas phase. The increase in acidic behavior across both analogs implies that methylation lowers the gas-phase acidity independently of carbon position in serine.

3.1.3 L-Penicillamine Results

The samples containing L-penicillamine were prepared in a similar manner to both 3-methylthreonine and α -methylserine. Instead of containing 10% ammonium hydroxide, however, samples contained only 5% in order to prevent the formation of disulfide bonds. For this analysis, the following reference acids were used: 2-nitrobenzoic acid, 4-hydroxybenzophenone, 3-nitrophenol, 3-trifluoromethylbenzoic acid, and 3-fluorobenzoic acid.

The preliminary KM plot 1 was generated using the same methods as in section 3.1.1 and is shown in Figure 3.9.

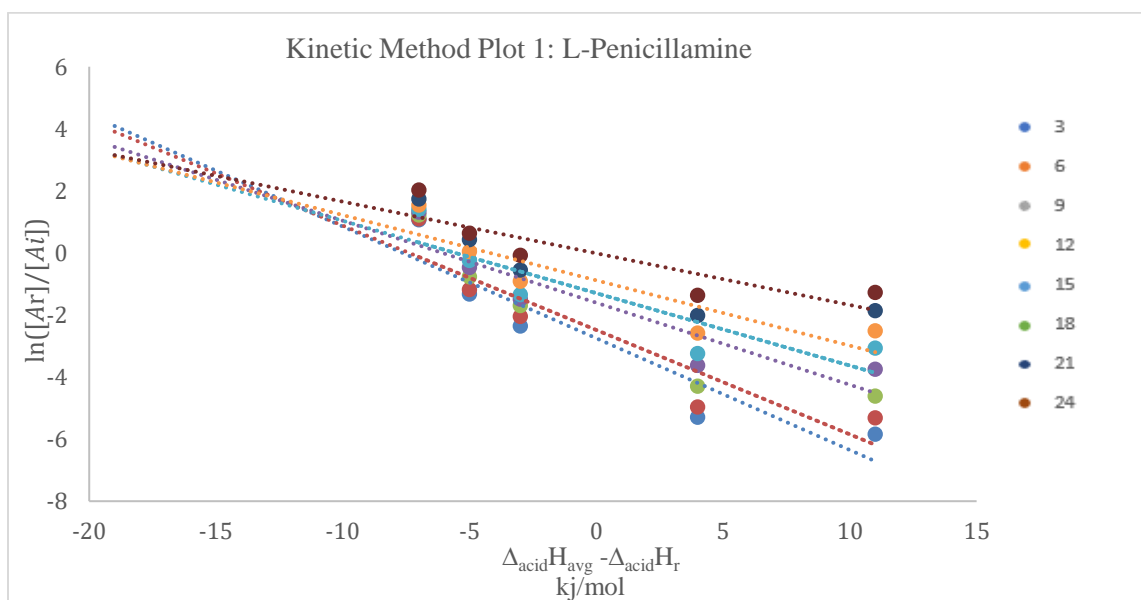


Figure 3.9: Preliminary KM plot 1 for L-penicillamine study

From the preliminary KM plot 1 shown in Figure 3.1.9, we calculated the slope and intercept of each line and plotted the intercept as a function of the slope to give KM plot 2 in Figure 3.10

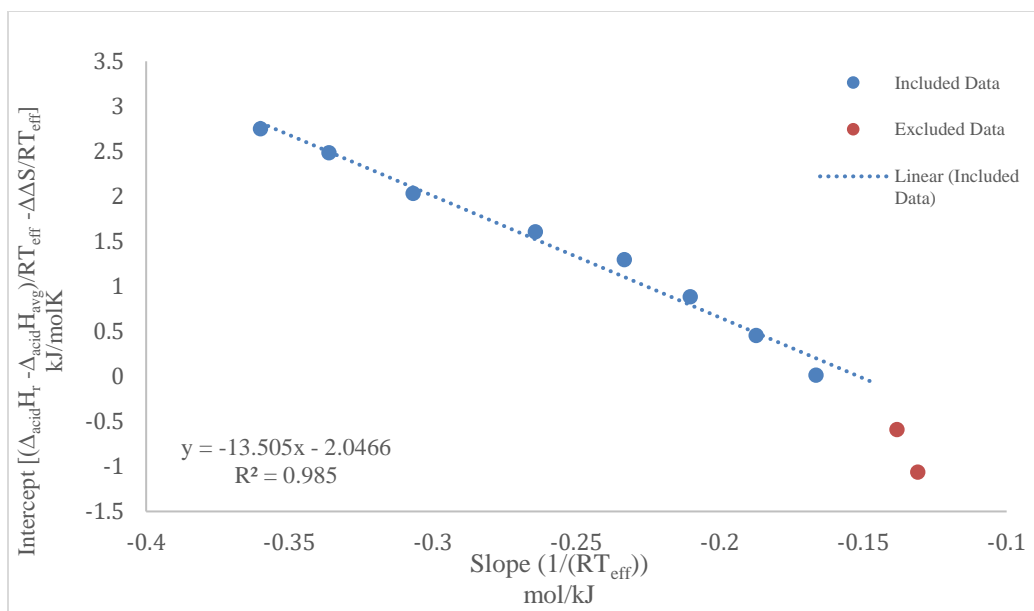


Figure 3.10: KM plot 2 of L-penicillamine study

After deriving the effective temperature for each collision energy using Equation 2.3, we examined deviations from linearity and excluded appropriate collision energies in our analysis. For L-penicillamine, the data taken at the collision energies of 27eV and 30eV was excluded due to deviation from linearity in both plot 2 (Figure 3.10) and effective temperature plot (Figure 3.11).

This is likely due specifically to the formation of sulfide anion at high collision energies, which was not accounted for in the selected mass range (35Da-350Da). In the collision-induced dissociation of deprotonated L-penicillamine, the sulfur anion became a statistically significant to the ion count at a collision energy of 25eV. Unfortunately, due to the facility closure, we were unable to re-collect the data with a lower mass range on the quadrupole. These experiments will be performed before publication.

When excluding data appropriately in plot 2 (Figure 3.10) and applying the proper mathematical operations, the preliminary prediction gives a value of 1381 kJmol⁻¹ for Δ_{acid}H and 17 Jmol⁻¹K⁻¹ for the apparent entropy.

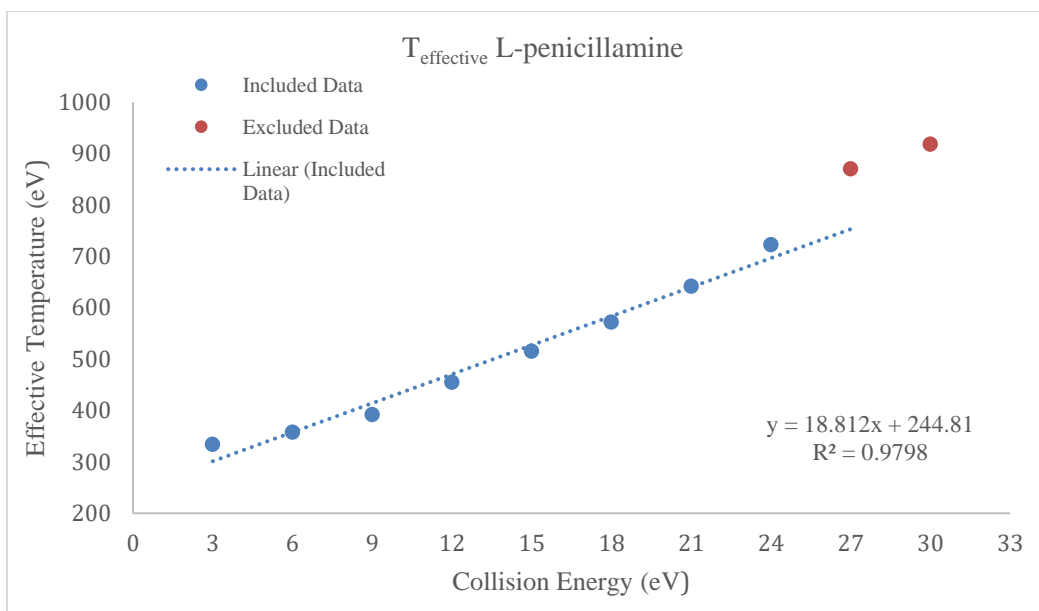


Figure 3.11: Effective Temperature of L-penicillamine collection

Preliminary predictions of DH and DS were used as starting points, and all reference acids within the analysis were included in the final ODR analysis because no reference acid showed excessive deviation from expected trends, as seen in the preliminary plot 2 (Figure 3.9). The ODR analysis gave a $\Delta_{\text{acid}}H$ value of $1380 \pm 18 \text{ kJmol}^{-1}$ and an apparent entropy value of $-24 \pm 22 \text{ Jmol}^{-1}\text{K}^{-1}$. Figure 3.12 shows the original data with the ODR-derived effective temperature lines and the final isothermal point used to extract ΔH and ΔS .

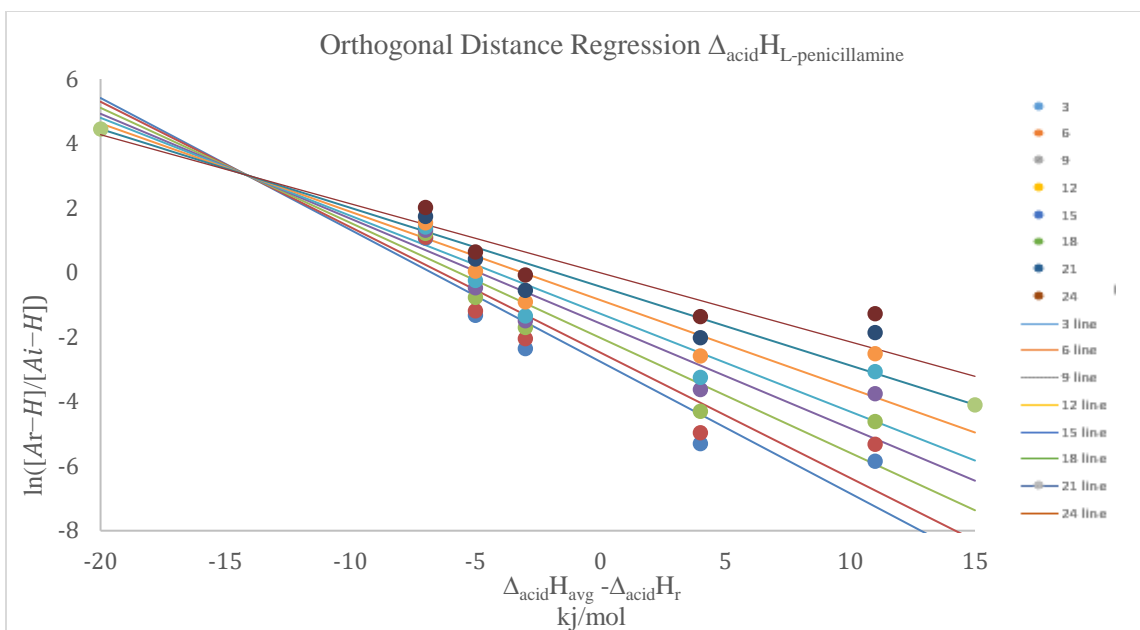


Figure 3.12 Orthogonal Distance Regression Analysis of L-penicillamine data

The experimental $\Delta_{\text{acid}}H$ for L-penicillamine (gem-dimethylcysteine) is significantly lower than that of cysteine (1396 kJ/mol), its analog. This means that the methylated amino acid acts more acidic in the gas phase.

3.3 Computational Predictions

As shown in Figure 3.13, there are three major hydrogen bonding motifs found in amino acids with aliphatic side chains.^{[37],[38]} Motif type *I* has a *syn*-arrangement of the carboxylic acid group and involves weak hydrogen bonding interactions between the hydrogen atoms on the amine group and the carbonyl oxygen atom. This arrangement is lowest calculated free energy and enthalpy conformer for the aliphatic amino acids glycine, alanine, valine, leucine, and isoleucine.^{[37],[38]} The lowest energy conformer of proline, motif type *II* also has a *syn* arrangement of the carboxylic acid group and involves similar hydrogen bonding of the amino hydrogen atoms to the carboxyl OH oxygen atom. Motif *III* has an *anti*-arrangement of the carboxylic acid group and involves a strong hydrogen bond between the carboxyl hydrogen atom and the amino nitrogen.

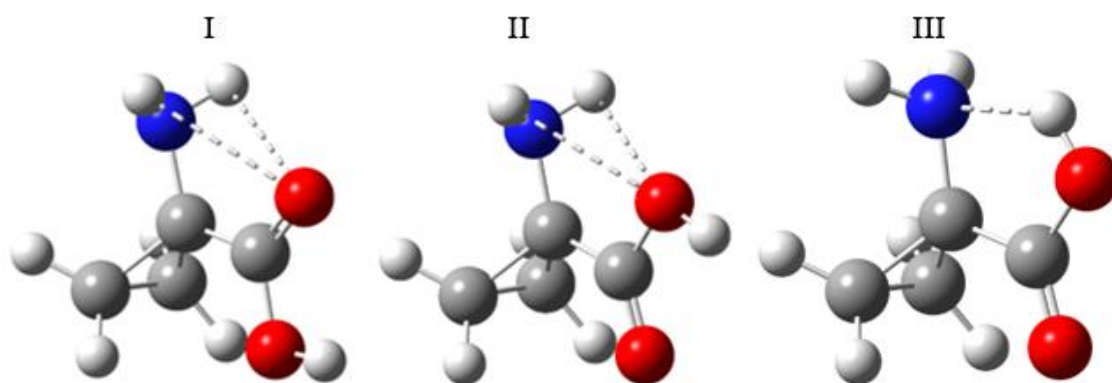


Figure 3.13: Illustrated Motifs of Amino Acid Hydrogen Bonding

3.3.1 Serine

In order to compare the molecules of interest, 3-methylthreonine and α -methylserine, to their protein-amino acid analog, the lowest energy conformers of neutral and deprotonated serine were obtained at the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory. The resulting lowest conformers are reproduced and discussed below. In these images, dark grey atoms correspond to carbon atoms, white to hydrogen atoms, red to oxygen atoms, and blue to nitrogen atoms. Solid lines represent covalent bonds, while dotted lines depict relevant hydrogen bonding. Though it is difficult to comprehensively compare all bonding modes without all conformers, these figures can still provide helpful insight for comparison in following sections.

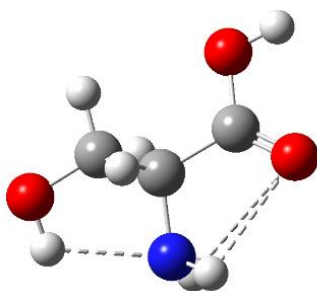


Figure 3.14: : The lowest conformers of neutral serine representing 54.0% of the 298K Boltzmann-weighted population.

The conformation of the lowest-energy structure of neutral serine follows motif I as described by Bouchoux^[37] in which the primary hydrogen bonding interaction involves the

amine hydrogen atoms with the carbonyl nitrogen. In serine, this motif is strengthened due to the fact the side chain hydroxyl group can also form a ‘type III’ hydrogen bond with the nitrogen allowing for additional stability.

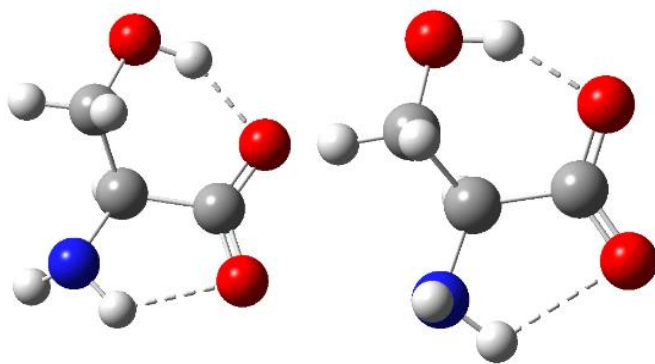


Figure 3.15: : The two lowest conformers of deprotonated serine representing 54.0% (left) and 43.5 % (right) of the 298K Boltzmann-weighted population.

For deprotonated serine, two low-energy conformers contribute to most of the Boltzmann population at 298 K. The main difference between these two conformers is a result of the rotation of the amide bond with both structures having the same hydrogen bonding arrangements (type I and an additional OH---O hydrogen bond. The high percentage of both conformers in the population suggests little steric hindrance in rotating the amide bond.

The computational results gave an isodesmic prediction of 1392 kJ/mol, which is in excellent agreement with the literature value, 1393 kJ/mol.

3.2.2 α -methylserine Results

For neutral and deprotonated α -methylserine, the lowest energy conformers were obtained at the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory. The resulting lowest conformers are reproduced and discussed below.

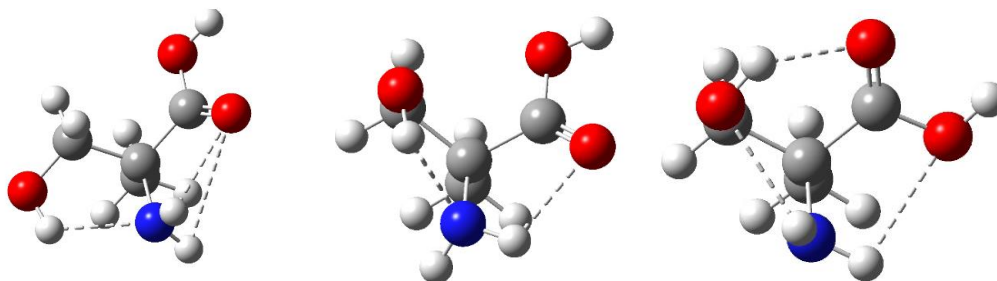


Figure 3.16: The three lowest energy conformers of neutral α -methylserine corresponding to 32.7%, 27.7%, and 25.4% of their 298K Boltzmann-weighted populations from left to right. .

In all structures, slight rotations of the bonds attaching the alcohol, amine, and carbonyl groups allow multiple instances of hydrogen bonding. The lowest energy conformer (left) is the most consistent with the stabilizing hydrogen bonding seen in the lowest conformer for neutral serine (Figure 3.14) with type I and type III motifs. The middle conformer has a similar hydrogen bonding arrangement with slight rotations of the single bonds. In the conformer shown on the right, the type I motif has been replaced with a type II NH---OH arrangement that is further stabilized by an OH---O=C H bond. These three conformations are within 0.7 kJ/mol in enthalpy at 298K, resulting in an almost equal mixture at room temperature.

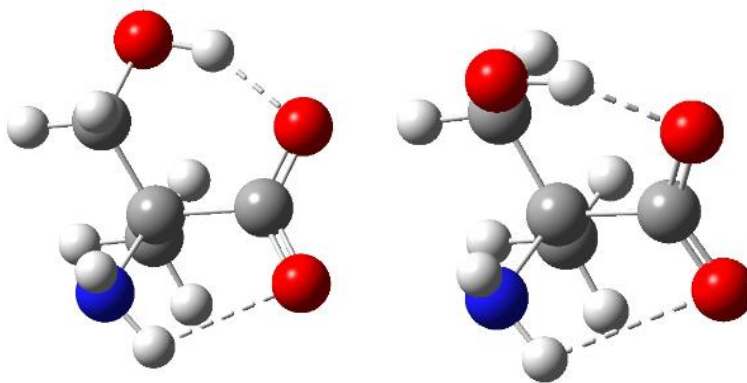


Figure 3.17: The two lowest conformers of deprotonated α -methylserine representing 81% (left) and 18.7% (right) of their 298K Boltzmann-weighted population.

For deprotonated α -methylserine, two low energy conformers were found that contributed to 99.7% of the 298K Boltzmann weighted population with a 3.6 kJ/mol enthalpy difference at 298K. The main difference between these conformers stems from a rotation of the

hydroxyl group. Both lowest structures of deprotonated α -methylserine involve hydrogen bonding with either side of the carboxylate, which is consistent with both displayed conformers of deprotonated serine. The position of the amine group mimics the second lowest energy conformer of deprotonated serine (right, Figure 3.15). The higher energy of the left arrangement of deprotonated α -methylserine suggests that it is not energetically favorable due to steric interactions associated with the added methyl group. This likely contributes to the preference of one rotational position over the other.

The computational results gave an isodesmic prediction of 1392 kJ/mol, which is in moderate agreement with the experimentally determined value within bounds of reasonable uncertainty.

3.2.3 3-Methylthreonine Results

For neutral and deprotonated 3-methylthreonine, the lowest conformers were obtained at the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory.

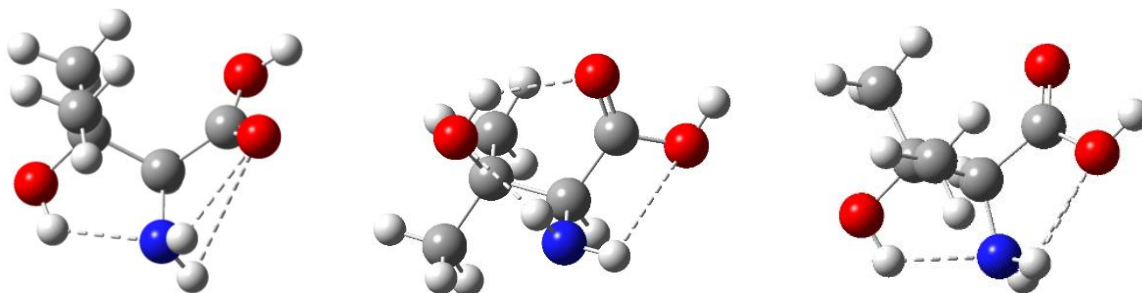


Figure 3.18: Lowest Energy Conformers of neutral 3-methylthreonine representing 72.0%, 9.6%, and 9.3% of the 298K Boltzmann-weighted population from left to right.

For neutral 3-methyl threonine, three structures were found that contributed to 90.9% of the population. The lowest energy conformer emulates the lowest energy conformer of neutral serine, and a type I / type III H-bonding arrangement and was found to be dominant. Two other conformers slightly higher in energy were located within 5kJ/mol. The middle structure in Figure

3.18 has a different H-bonding scheme with the amine hydrogen atoms forming hydrogen bonds with both hydroxyl oxygen atoms while the structure on the right displays a type II/type II motif. In comparison to α -methylserine, the additional geminal methyl groups decreased the stability of other hydrogen bonding modes by increasing the steric hindrance of closely-bound groups.

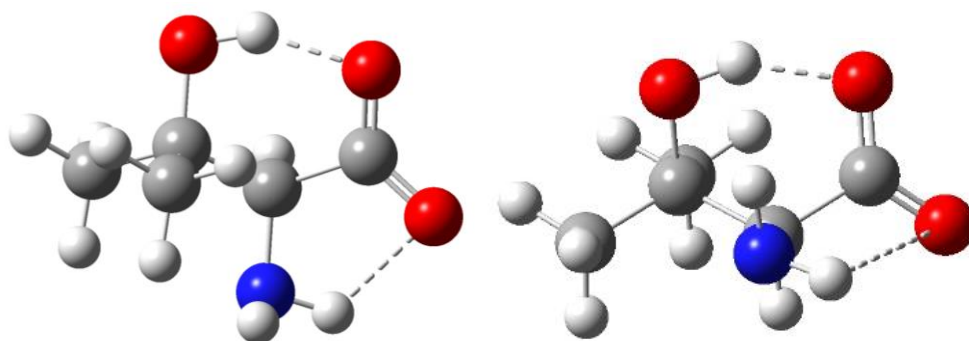


Figure 3.19: Lowest Energy Conformers of deprotonated 3-methylthreonine representing 88.3% and 10.0% of the 298K Boltzmann-weighted population from left to right..

Two low energy conformers were located within 6kJ/mol of the lowest calculated energy for deprotonated 3-methylthreonine. The dissimilarity between these two conformers is rooted in the position of the amine group. The lowest energy structure of deprotonated 3-methylthreonine is consistent with the form of the second-lowest energy conformer of deprotonated serine (right, Figure 3.15). The absence of the left arrangement (left, Figure 3.15) of deprotonated serine suggests that it is not energetically favorable due to steric interactions associated with the addition of the geminal methyl group. This contributes to the preference of one rotational position over the other.

The computational results gave an isodesmic prediction of 1390 kJ/mol, which is in moderate agreement with the experimentally determined value within bounds of reasonable uncertainty.

3.2.4 Cysteine

In order to compare the molecules of interest, L-penicillamine and α -methylcysteine, to their protein-amino acid analog the lowest conformers of neutral and deprotonated cysteine were obtained at the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory. Though it is difficult to comprehensively compare all bonding modes without access to all conformers of the reference compounds, these figures can still provide helpful insight for comparison in following sections.

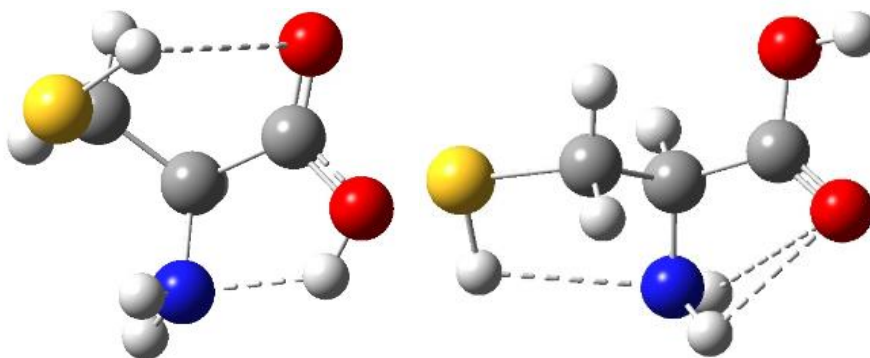


Figure 3.20: Lowest Energy Conformers of neutral cysteine representing 35.8% and 25.5% of the 298K Boltzmann-weighted population from left to right.

The conformation of the lowest energy conformer (left) generally follows motif III as described by Bouchoux.^[37] In this motif, the carboxylic acid group has an anti-arrangement and strong hydrogen-bonding between the carboxyl hydrogen atom and the amino nitrogen. In cysteine, an additional stabilization is offered by means of a hydrogen bond between the thiol proton and the carbonyl oxygen atom. The second-lowest energy structure is not pictured but is very similar to the first and accounts for 35.6% of the population. The third lowest-energy structure of neutral cysteine (right, Figure 3.20) follows motif II as described by Bouchoux.^{[37],[38]} In cysteine, this motif is strengthened because the side chain thiol group can also form a ‘type III’ hydrogen bond with the nitrogen, which allows for additional stability.

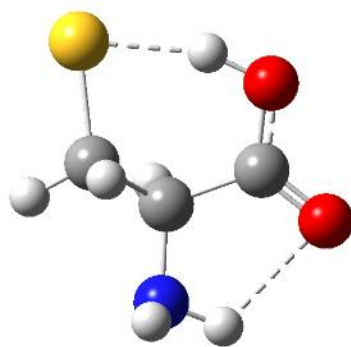


Figure 3.21: Lowest Energy Conformer of deprotonated cysteine, representing 55.8% of the 298K Boltzmann-weighted population.

The lowest energy structure of deprotonated cysteine is characterized by hydrogen-bonding on either end of the carboxylic acid, much like the lowest conformer of the neutral. This structure looks similar in shape to the lowest conformer of deprotonated serine. However, in this structure, the sulfur is the site of deprotonation, making it unique among the 20 protein amino acids. Deprotonation at the sulfur atom allows for the strong type I hydrogen bonding arrangement to be maintained.

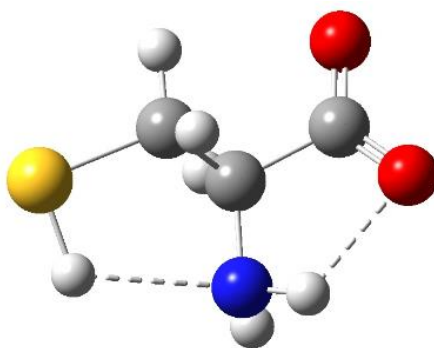


Figure 3.22: Lowest Energy Conformer of carboxylic acid-deprotonated cysteine representing 2.1% of the 298K Boltzmann-weighted population.

The lowest energy structure of carboxylic acid-deprotonated cysteine is characterized by hydrogen-bonding on either end of the amino group. This is vastly different from the favored interaction on either side of the carboxylate but is reminiscent in shape to the third lowest neutral conformer (Figure 3.20, right).

The computational results gave an isodesmic prediction of 1396 kJ/mol, which is in exact agreement with the literature value, 1396 kJ/mol.

3.2.5 α -methylcysteine

For neutral and deprotonated α -methylcysteine, the lowest conformers were obtained at the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory.

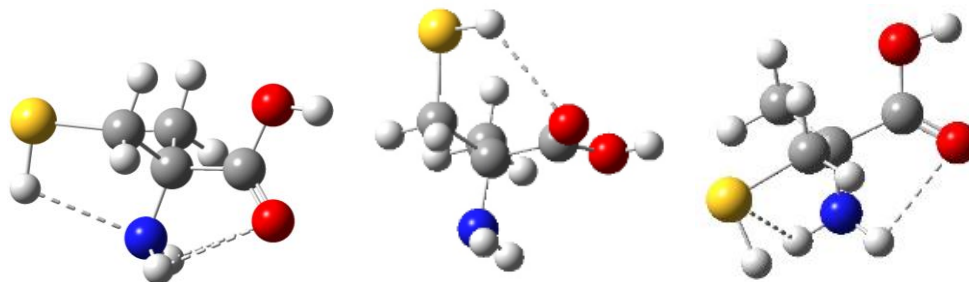


Figure 3.23: Lowest Energy Conformer of neutral α -methylcysteine neutral α -methylcysteine representing 45.7%, 28.6%, and 15.6% of the 298K Boltzmann-weighted population from left to right

The lowest energy structure of α -methylcysteine follows motif I with the added stability of the thiol-amine hydrogen bond. None of the pictured structures demonstrate hydrogen bonding on either side of the carboxylic acid, which was a characteristic of 71.6% of its protein amino acid homolog's neutral form. This is likely due to the steric strain added in folding the end of the molecule over the new methyl group.

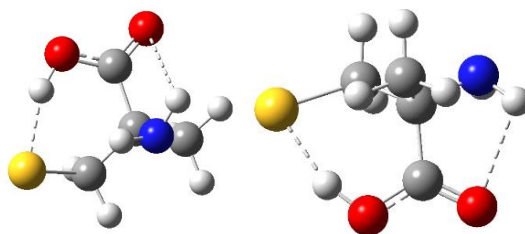


Figure 3.24: Lowest Energy Conformers of deprotonated α -methylcysteine neutral representing 73.4%, and 26.3% of the 298K Boltzmann-weighted population from left to right

Two low energy conformers were obtained within 4 kJ/mol of the lowest calculated energy. Both conformers show hydrogen bonding on either side of the carboxylic acid group, a

tendency familiar with cysteine. Compared to the lowest energy conformer of the neutral molecule, there appears to be far less steric strain associated with the methyl-sulfur interaction on the lowest energy conformer of the deprotonated molecule. Instead of having three modes of hydrogen bonding as seen in the neutral lowest energy conformer, the amine of the deprotonated lowest energy species has only one mode of hydrogen bonding, allowing a more symmetric, compact arrangement of atoms.

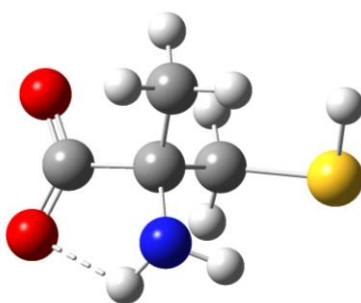


Figure 3.25: The lowest-energy conformer of a-methylcysteine. This structure had a 20kJ/mol energy difference when compared to the S-deprotonated structure in Figure 3.24

The lowest energy conformer of the carboxylate form of the anion accounted for less than a hundredth of a percent of the 298K Boltzmann population. Despite starting calculations with structures containing the S-H component, any conformer in which the sulfur and oxygen atoms came near to each other resulted in proton transfer during optimization. This structure, therefore, is an upper limit to the difference in acidity between the two sites. In cysteine the two deprotonation sites were much closer in energy as evidenced by hydrogen-deuterium exchange experiments.^[39] In this case the difference is probably larger. Regardless, hydrogen-deuterium exchange would be useful in further analyzing the energetics and mechanisms of deprotonation on either site.

The computational results gave an isodesmic prediction of 1387 kJ/mol for the sulfur deprotonated site and 1407 kJ/mol for the carboxylic acid-deprotonated site. Though we were not

able to obtain experimental data on this compound, the consistency in calculation and experimentation for L-penicillamine (section 3.2.6) imply that this method of calculation is sound for similar molecules.

3.2.6 L-penicillamine Results

For neutral and deprotonated L-penicillamine, the lowest conformers were obtained at the B3LYP/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory.

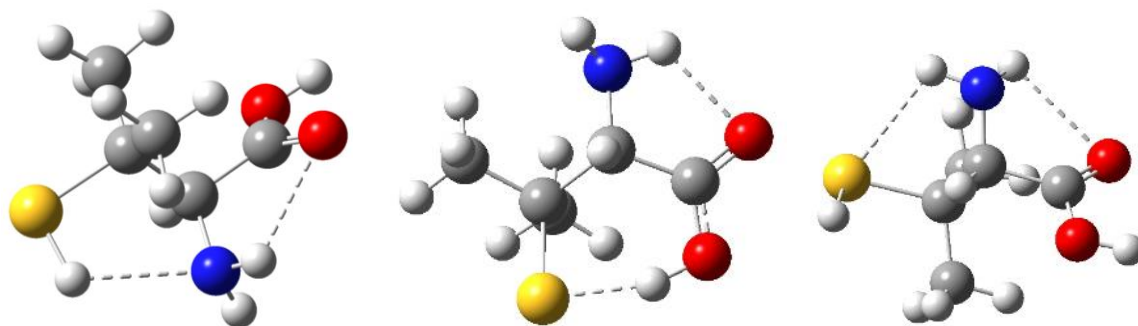


Figure 3.26: Lowest Energy Conformer of neutral L-penicillamine representing 44.0%, 19.6%, and 10.0% of the 298K Boltzmann-weighted population from left to right

Three low energy conformers of neutral L-penicillamine were found within 3.5kJ/mol of the lowest calculated energy. Like α -methylcysteine, the lowest energy structure of L-penicillamine follows motif I with the added stability of the thiol-amine hydrogen bond. Unlike α -methylcysteine, a structure that follows motif III and is very similar to the lowest energy conformer of neutral cysteine (left, Figure 3.20) accounts for 19.6% of the Boltzmann-weighted population. This conformer sees the same added stability with the thiol-carboxylic acid hydrogen bond.

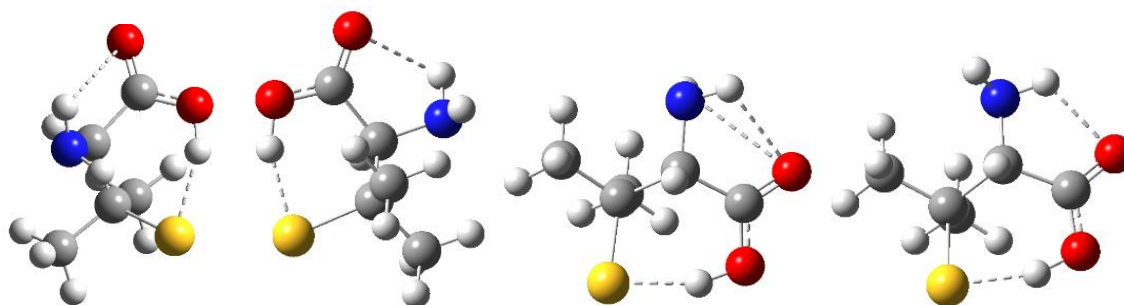


Figure 3.27: Lowest Energy Conformer of deprotonated L-penicillamine representing 52.9%, 15.1%, 15.0%, and 13.3% of the 298K Boltzmann-weighted population from left to right

Three low energy conformers of deprotonated were found within 5kJ/mol of the lowest calculated energy. All of the low energy conformers feature hydrogen bonding on either side of the carboxylic acid, each a thiol-OH interaction. These conformers are very similar in shape to the lowest energy conformer of deprotonated cysteine (left, Figure 3.20) and α -methylcysteine.

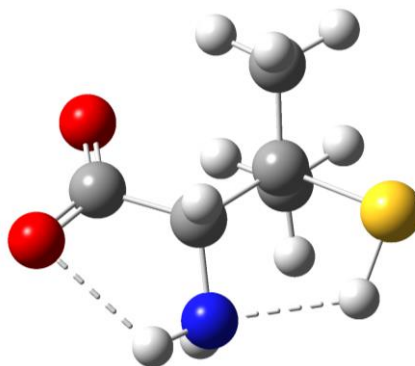


Figure 3.28 The lowest-energy conformer of carboxylic acid-deprotonated L-penicillamine. This structure had a 25 kJ/mol energy penalty when compared to the S-deprotonated structures in Figure 3.27

As with α -methyl serine, the lowest-energy conformer of carboxylic acid-deprotonated structure has the sulfur abnormally far away from the ionized carboxylate and had an energy difference of 25kJ/mol, accounting for only a fraction of a percent of the 298K Boltzmann distributed population. Again, hydrogen-deuterium exchange would be useful in further analyzing the energetics and mechanisms of deprotonation on either site. The computational results gave an isodesmic prediction of 1383 kJ/mol for the sulfur deprotonated site and 1408

kJ/mol for the carboxylic acid-deprotonated site. The experimental results are in excellent agreement with the sulfur-deprotonated site.

Overall, computational predictions and analysis of structures support the notion that methylation similarities in deprotonated structures of protein-amino acids and their non-protein amino acid analogs support the notion that the increase in acidity due to methylation most likely a factor of destabilization of the neutral molecule rather than significant stabilization of the deprotonated species. A small exception of this may manifest itself in the steric strain of methylation providing a preference in the rotational position for serine homologs. This could also have the opposite effect and thus increase the calculated isodesmic acidity.

3.3 Non-Protein Amino Acid Summarized Results

Using the methods described, we aimed to computationally predict and experimentally determine the $\Delta_{\text{acid}}H$ of the non-protein amino acids pictured in Figure X. Of these, α -methylserine and 3-methyl-threonine are analogs of the protein amino acid, serine while α -methylcysteine and L-penicillamine are analogs of cysteine. Though we were able to make computational predictions for the acidities all the compounds of interest, we were not able to obtain experimental data on α -methylcysteine due to its tendency to make disulfide bonds, even in only slightly basic solution. The quantities determined are summarized in Table 3.1, along with the evaluated $\Delta_{\text{acid}}H$ of their homologs for reference.

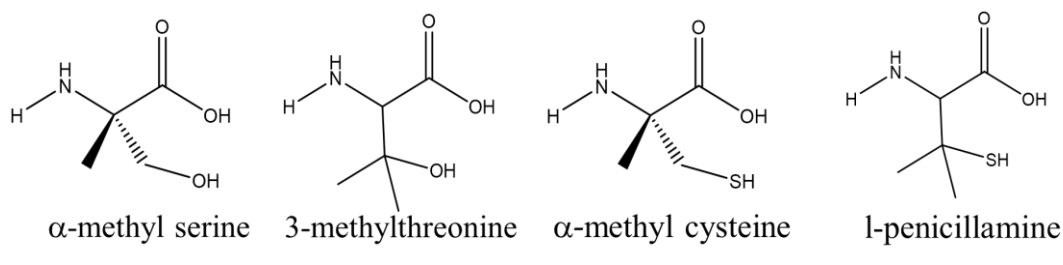


Figure 3.29: Non-Protein Amino Acids of Interest Studied

Experimental and Computed $\Delta_{\text{acid}}\text{H}$ for Non-Protein Amino Acids			
	$\Delta_{\text{acid}}\text{H}$ (predicted) (COO ⁻) (kJ/mol)	$\Delta_{\text{acid}}\text{H}$ (predicted) S ⁻ (kJ/mol)	$\Delta_{\text{acid}}\text{H}$ (experimental) (kJ/mol)
α -methylserine	1392		1379 \pm 23
3-methylthreonine (gem-dimethylserine)	1390		1378 \pm 23
α -methylcysteine	1407	1387	n/a
L-penicillamine (gem-dimethylcysteine)	1408	1383	1380 \pm 18
		$\Delta_{\text{acid}}\text{H}$ (evaluated) (kJ/mol)	
serine ^[28]		1393	
cysteine ^[28]		1396	

Table 3.1: A summary of the computational and experimental quantities extracted for the compounds of interest

As seen above, both the computational and experimental values of the methylated-amino acids have a lower $\Delta_{\text{acid}}\text{H}$ than their un-methylated counterparts, meaning they are more acidic in the gas-phase. The experimental value for L-penicillamine is extremely consistent with the computed value for the S-deprotonation site. Though the values for the serine homologs have the computed value within 95% confidence bounds, the experimental values are centered around a much lower value than expected. Large entropy values lead to extrapolation of the isothermal point during ODR analysis, which can lead to large uncertainties shown here. Methylation of both sites lower the gas phase acidity within the experimental data, and this implies that methylation lowers the gas phase acidity at least somewhat independently of position.

Literature Values for Glycine, Alanine, and Valine	
	$\Delta_{\text{acid}}\text{H}$ (literature) ^[28] (kJ/mol)
Glycine	1433
Alanine (α-methylglycine)	1425
Valine (gem-dimethylalanine)	1420

Table 3.2: Literature Values for protein amino acids and their analogs

Looking at established acidities of amino acids and their analogs, we can examine the general trend in acidity as a function of methylation and determine whether it agrees with the relationship displayed in the results of this thesis. Based on literature values, it appears that the alpha-methylation of glycine lowers the $\Delta_{\text{acid}}H$ by ~ 8 kJ/mol, as seen from the acidity value of alanine. Adding a geminal dimethyl group, as seen from alanine to valine, also lowers the $\Delta_{\text{acid}}H$ by a ~ 5 kJ/mol. This general trend agrees with the data presented for cysteine, serine, and their analogs and supports the notion that methylation lowers acidity regardless of the position.

3.3 Carbamazepine Cluster Study

The major over-arching goal of the research in the Poutsma lab is to uncover the connections between molecular structure, energetics and reactivity. One way to explore these connections in a very relevant way is through the study of polymorphic active pharmaceutical ingredients. The Poutsma lab has experience in characterizing active pharmaceutical ingredients, such as isatin, in the gas-phase.^[8] This thesis examines clusters of the compound carbamazepine and its analog, dibenzazepine to isolate the role of the amide group in clustering.

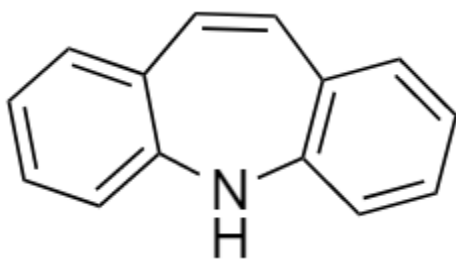


Figure 3.30: Structure of Dibenzazepine

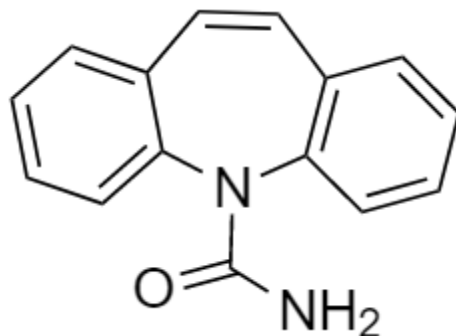


Figure 3.31: Structure of Carbamazepine

3.3.1 Tetramer Characterization:

Solutions of carbamazepine (Figure 3.3.2) in concentrations of 0.01 mM, 0.1 mM, 1.0mM, and 10 mM were prepared in methanol with 1% formic acid. Using an external syringe pump at a rate of 60 μ L/hr, the samples were injected into an external electrospray ionization source on a Finnigan TSQ Quantum Ultra quadrupole mass spectrometer with a capillary temperature of 120° C. The masses and intensities were exported, and the intensity of both the sodiated and the protonated clusters were extracted and normalized to the protonated monomer. Magic number clusters are clusters composed of a number of molecules, n, characterized by an unusually large intensity compared to clusters with either n-1 or n+1 molecule(s). These clusters are easily distinguishable in magic number plots, where the natural log of these normalized ion intensities is plotted as a function of the number of molecules in the associated cluster. Below, the raw mass spectra and the corresponding magic number plot are shown at each concentration.

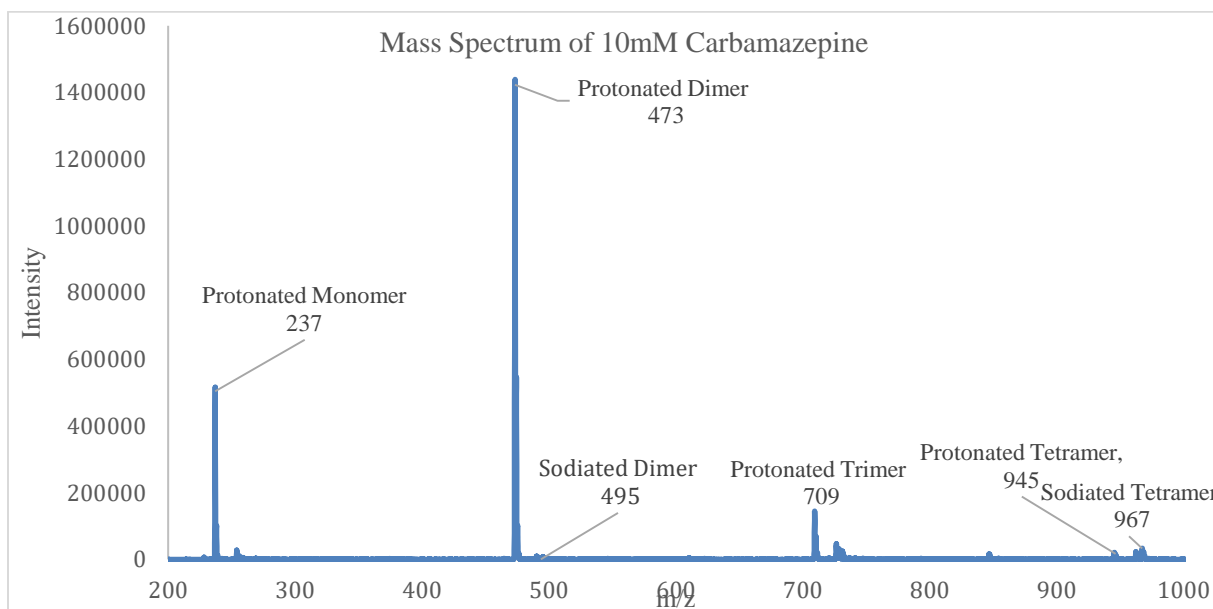


Figure 3.32: Raw Mass Spectrum of 10.00 MM CBZ solution with 1% formic acid'

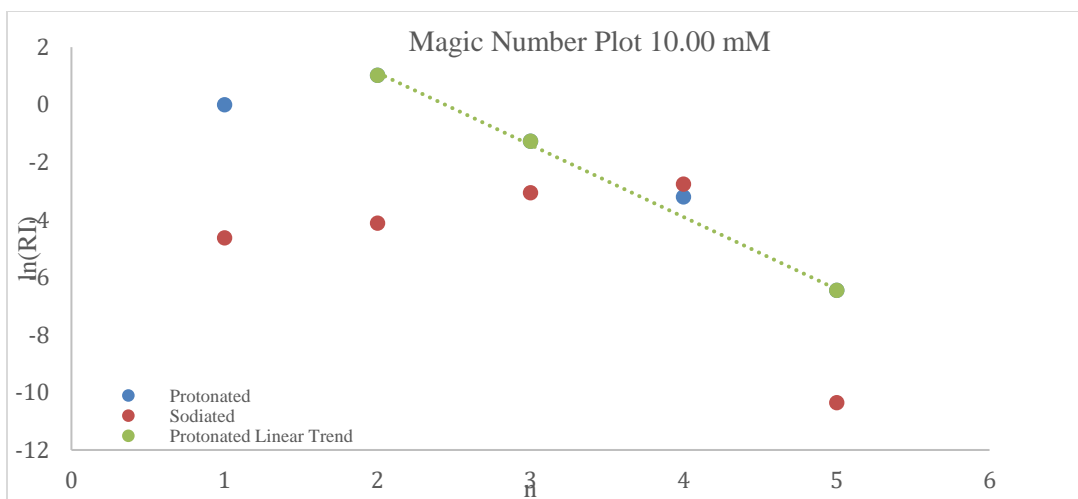


Figure 3.33: Magic Number plot of Carbamazepine at 10 mM concentration. Here, the intensities of each peak have been normalized to the protonated monomer for a relative intensity (RI).

At 10 mM/L the homodimer (n=2) is the protonated species with the highest intensity. However, a tetramer anomaly presents itself as positive deviation in an otherwise negative linear trend. In the sodiated species, the tetramer anomaly is especially magic, representing the sodiated cluster with the highest normalized intensity.

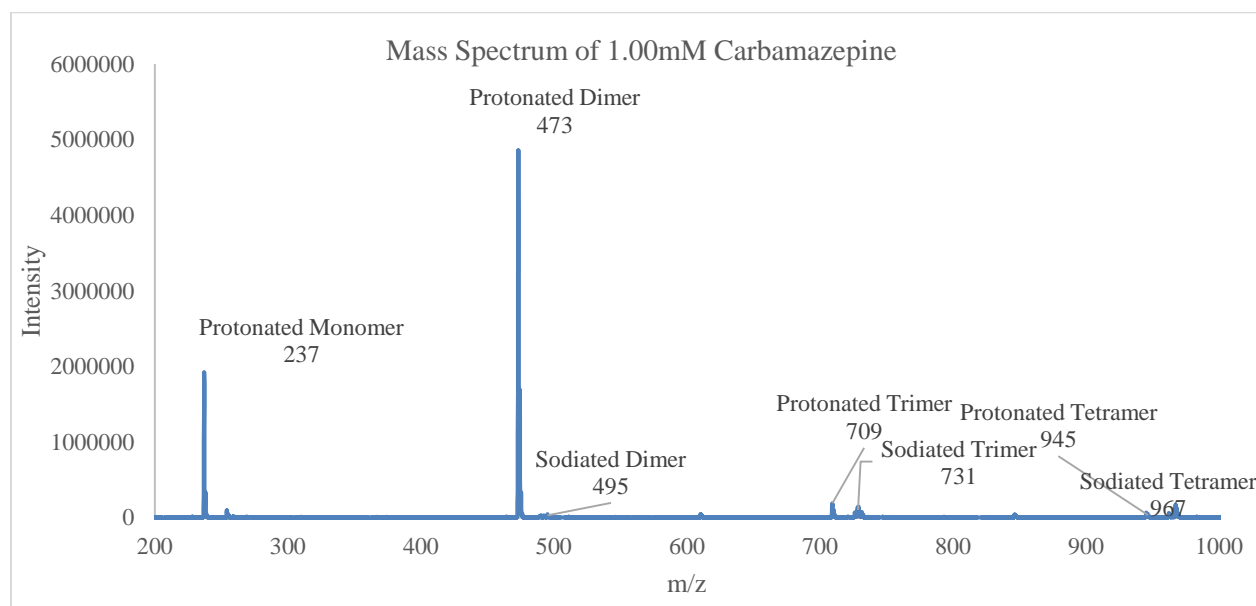


Figure 3.34: Raw Mass Spectrum of 1.00 mM CBZ solution with 1% formic acid

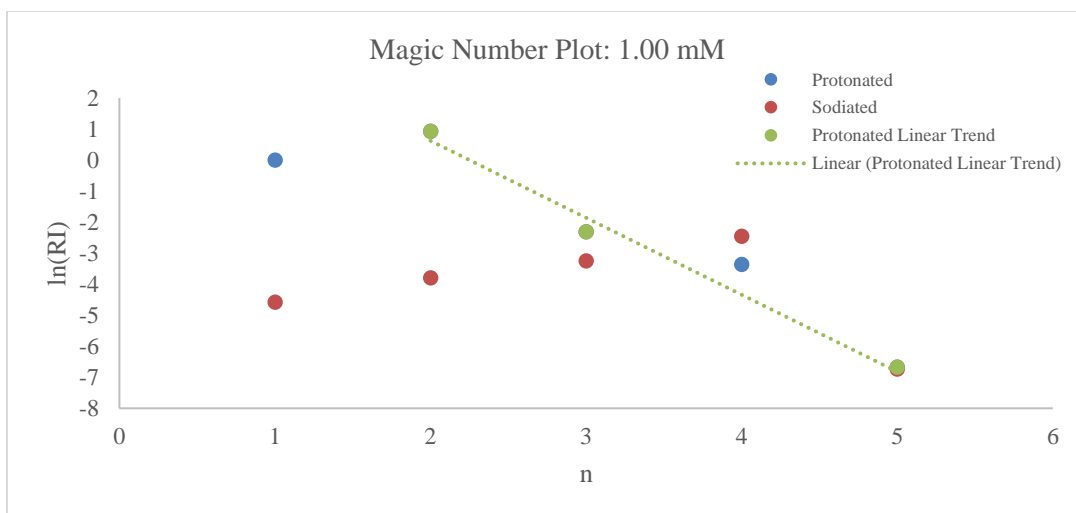


Figure 3.35 Magic Number plot of Carbamazepine at 1.00 mM concentration. Here, the intensities of each peak have been normalized to the protonated monomer for a relative intensity (RI).

Similar results are seen in the 1 mM solution in which, a tetramer anomaly again presents itself as positive deviation in a negative linear trend starting from the homodimer (n=2) for the protonated species. In the sodiated species, the tetramer anomaly is still especially magic, representing the sodiated cluster with the highest normalized intensity.

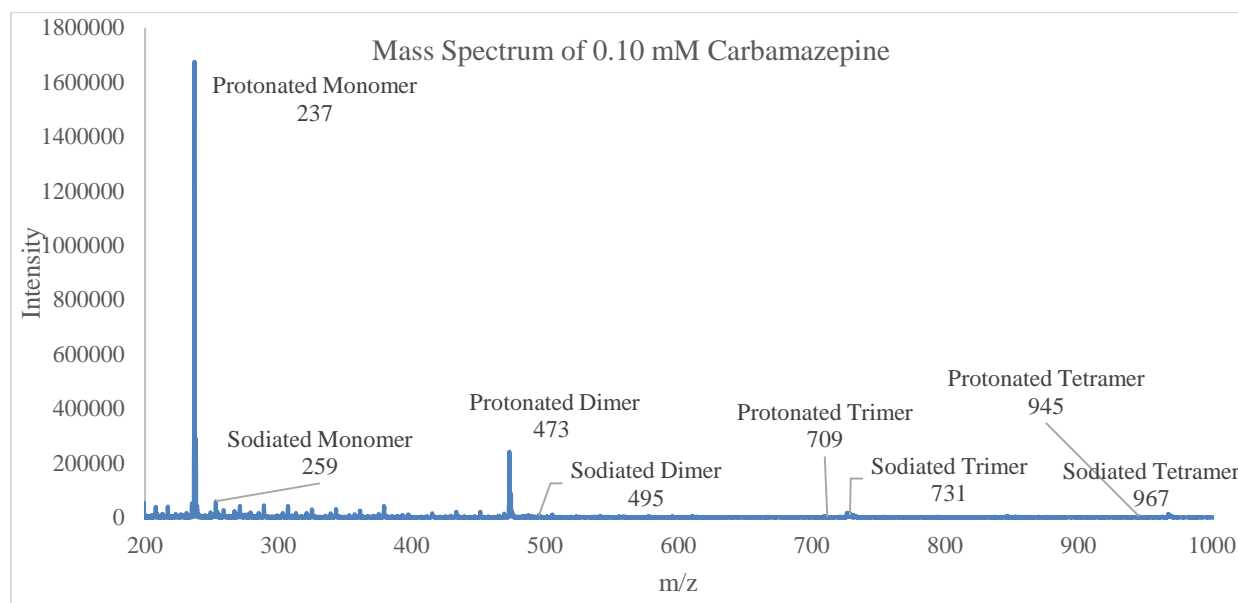


Figure 3.36: Raw Mass Spectrum of 0.10 MM CBZ solution with 1% formic acid

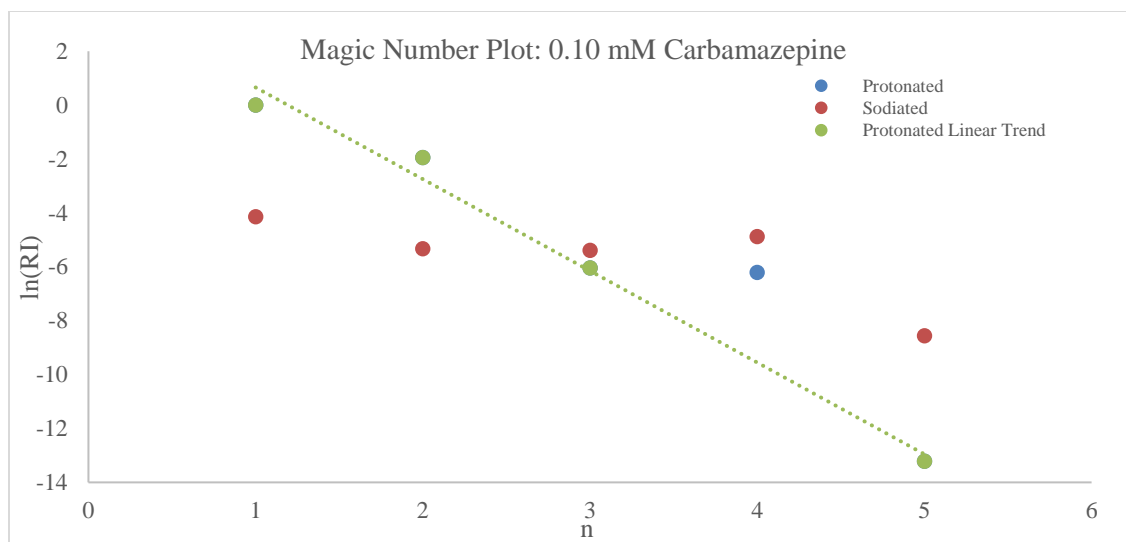


Figure 3.37: Magic Number plot of Carbamazepine at 1.00 mM concentration. Here, the intensities of each peak have been normalized to the protonated monomer for a relative intensity (RI). At this concentration, the deviation from linearity of the protonated tetramer is the most positive.

At a concentration of 0.10 mM, the homodimer is no longer the most intense peak of the spectrum. Here, the protonated tetramer anomaly appears more obvious than in the more concentrated samples and presents itself as discussed above. In the sodiated species, the tetramer anomalously is still magic, but has lost claim to the most intense peak of the sodiated clusters.

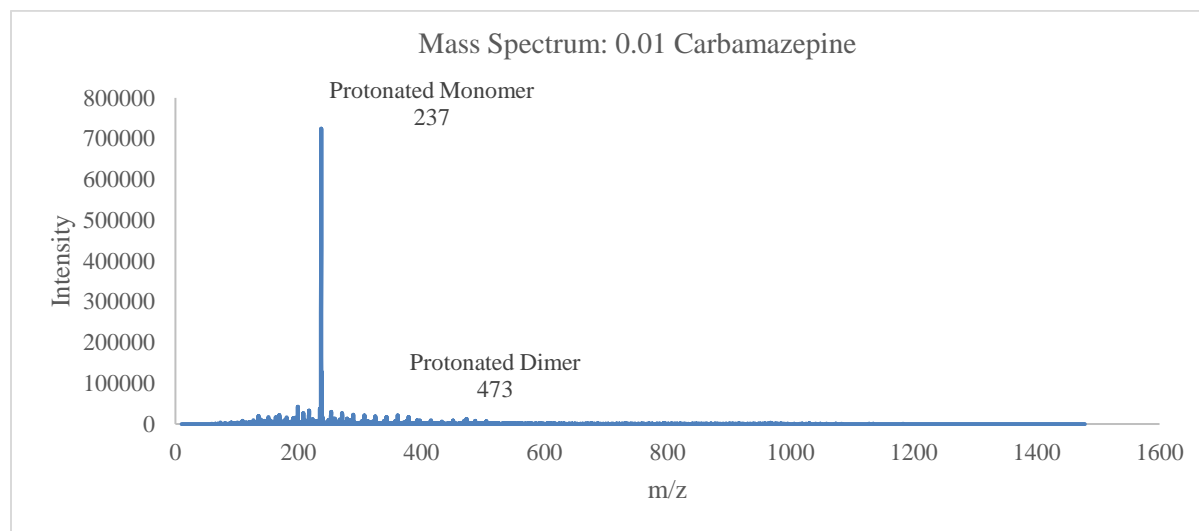


Figure 3.38 Raw Mass Spectrum of 0.01 mM CBZ solution with 1% formic acid

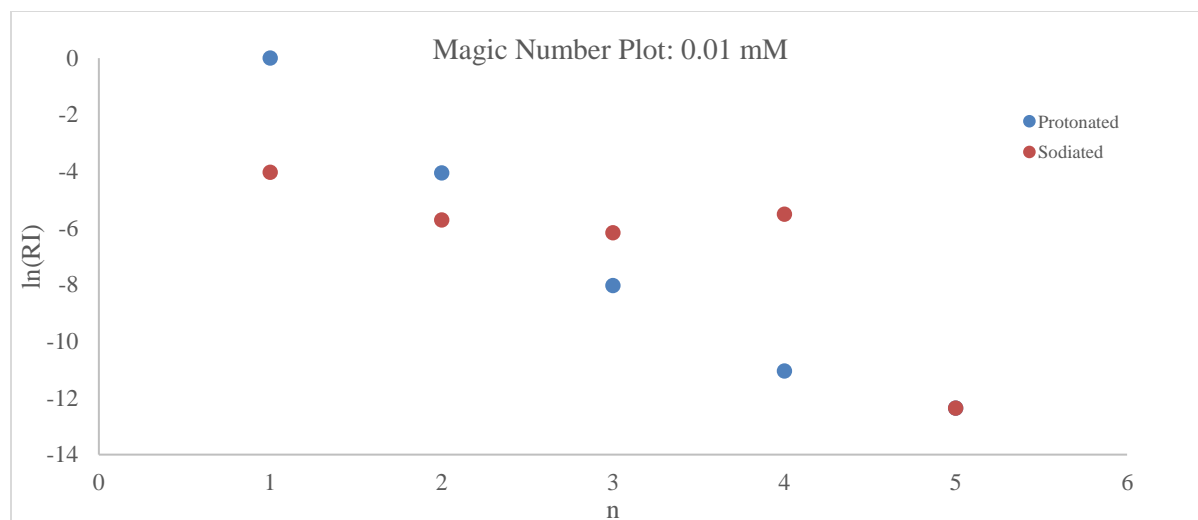


Figure 3.39: Magic Number plot of Carbamazepine at 0.01 mM concentration. Here, the intensities of each peak have been normalized to the protonated monomer for a relative intensity (RI). At this concentration, the protonated tetramer begins to follow the linear trend and lose anomalous properties. However, the sodiated species does not see the same decrease in activity.

At a concentration of 0.01 mM, the protonated tetramer is no longer an anomaly, following the linear decreasing trend of the protonated species. In the sodiated species, the sodiated tetramer is still magic, appearing more intense than both the sodiated trimer and pentamer.

The raw mass spectra and the magic number plots support the existence of increased stability of the sodiated and protonated tetramer. In the protonated species, this anomaly presents itself as positive deviation in an otherwise negative linear trend. In the sodiated species, the tetramer is the sodiated cluster with the highest intensity relative to the protonated monomer. Our mass spectra supported the existence of an anomalous tetramer of carbamazepine in concentrations down to 0.10 mM, both in the sodiated and protonated form. The protonated tetramer appears the most anomalous at a concentration of 0.10 mM. The stability of this tetramer is further supported by replicate measurements of the sodiated clusters taken by the Kandel group at Notre Dame and corroborates the results of their STM imaging, which features a uniform network of tetramers on the surface of Au(111) foil after solution deposition. The sodiated species shows anomalous activity across all concentrations. Meanwhile, this trend is consistent in all concentrations down to 0.10 mM for the protonated tetramer. At this

concentration, the protonated tetramer demonstrates the greatest positive deviation from linearity.

In addition to these spectra, we also took spectra of carbamazepine solutions spiked with sodium chloride. Though we were unable to retrieve the spectra from the lab computer due to lab closure, the base peak within these spectra is the sodiated tetramer, emphasizing its anomalous quality. In these spectra, the highest intensity peaks are all sodiated clusters, implying that sodium further stabilizes the clusters. Because these results are similar to Notre Dame's mass spectra, which were prepared without sodium chloride, this highlights the amount of adventitious sodium in Notre Dame's facility instrument.

These experiments were repeated with dibenzazepine, which exhibited significantly reduced clustering behavior in the mass spectrum. Likely due to the absence of the amide group that allows increased hydrogen bonding, the lack of clustering agrees with the STM data collected at Notre Dame where the monolayer of dibenzazepine displayed little to no networked-clustering behavior.

3.3.2 Ion Mobility Studies of Carbamazepine Clusters

Along with stability, ion mobility spectrometry is of interest to the studies of polymorphic structures. Being able to separate a given multimer mass-to-charge ratio based on shape gives insight into how clusters with the same multimer building block may assemble into different crystalline structures.

For FAIMS experiments, the ESI source was replaced with the FAIMS source, and the samples were injected with an external syringe pump at 60 $\mu\text{L}/\text{hour}$, and a compensation voltage scan spectrum was obtained for each mass of interest. These scans and their descriptions follow below.

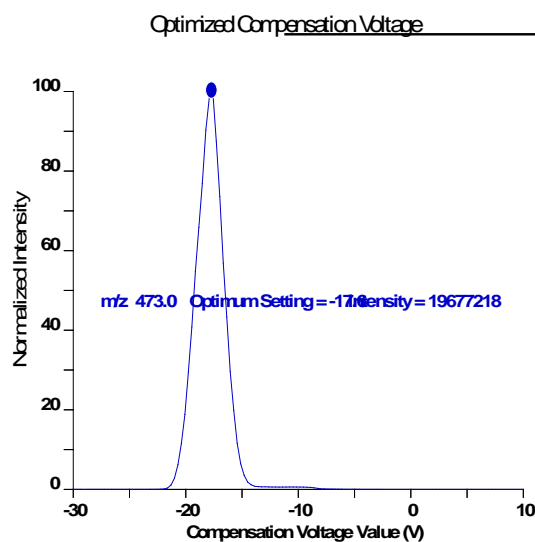
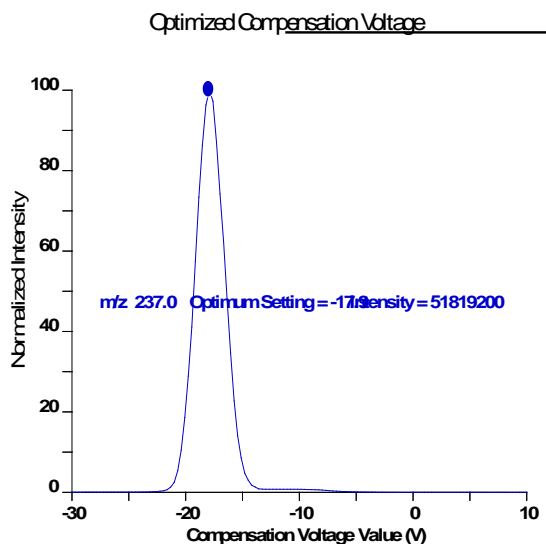


Figure 3.40: Compensation Voltage Scan of CBZ monomer Figure 3.41: Compensation Voltage Scan of protonated CBZ dimer

As observed in Figure 3.40 and 3.41, the compensation voltage scan results in one peak only for the protonated monomer and dimer. This supports an ensemble of ions with one mobility for each cluster.

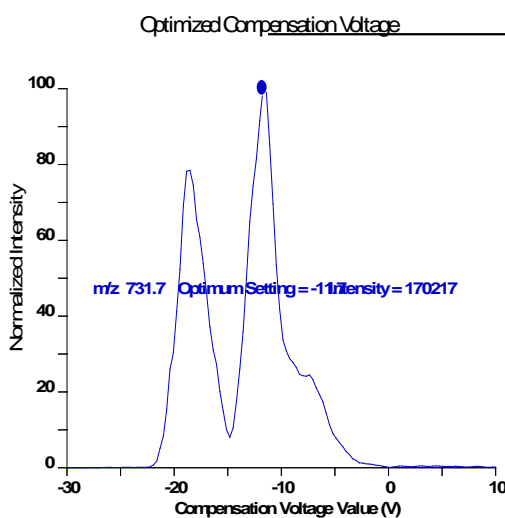
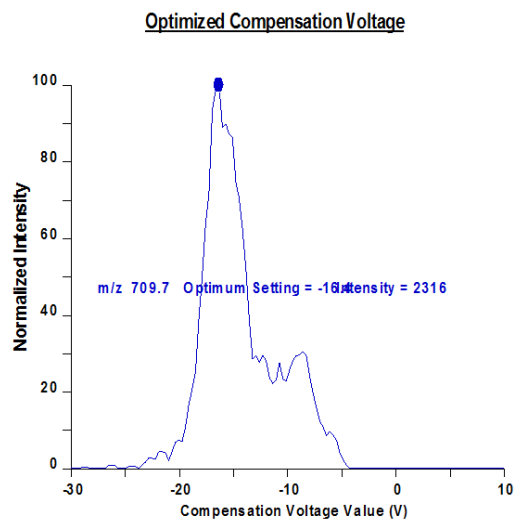


Figure 3.42: Compensation Voltage Scan of CBZ trimer(H)

Figure 3.43: Compensation Voltage Scan of CBZ trimer(Na)

The intensity of the protonated trimer ($m/z=709$) was not high enough to make any conclusive judgement about the ion mobility of the species, but the sodiated trimer did have a

significant intensity. The compensation voltage (CV) scan was characterized by two distinct peaks, supporting an ensemble of ions with two mobilities. Collision induced dissociation of ions at either compensation voltage yielded significantly different results, which are unfortunately inaccessible due to university shutdown. Thus, it is unclear whether this was due to a contaminant or to a unique arrangement of carbamazepine molecules at either voltage. This is a subject for further study and would be best paired with a scanning tunneling microscopy study of carbamazepine trimers with methods similar to those of Sykes^[40].

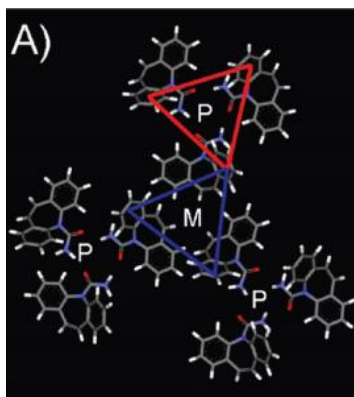


Figure 3.44. Energy minimized model of three CBZ trimers on an Au(111) surface. The red triangle designates a CBZ trimer, while the blue triangle designates a DBZ trimer. The chiral configuration of each trimer is denoted by a white letter, either “P” or “M”

In this study, they observed two primary methods of stabilization in CBZ and DBZ trimers, as shown in Figure 3.44. In this study, the Sykes group found that, in two dimensions, CBZ formed trimers mostly in conformation P while DBZ was characterized of trimers in conformation M when both were prepared via solution deposition onto an Au(111) surface. This was attributed to the role of the amide group in increased symmetric hydrogen bonding interactions within CBZ, and their semiempirical geometry optimization of a free CBZ trimer removed from the surface showed that the hydrogen bonding motif was retained in the optimized, isolated arrangement.^[40] Though outside the scope of this experiment, it would be interesting to determine whether multiple arrangements of the CBZ trimer with different

hydrogen bonding motifs exist in three dimensions within our study's conditions. It is currently unclear how sodium affects the configuration of the trimer, and the mobility determinations of this ion cannot necessarily be transferred over to the proton-bound or neutral species. Regardless, further study would benefit the understanding of favorable CBZ interactions.

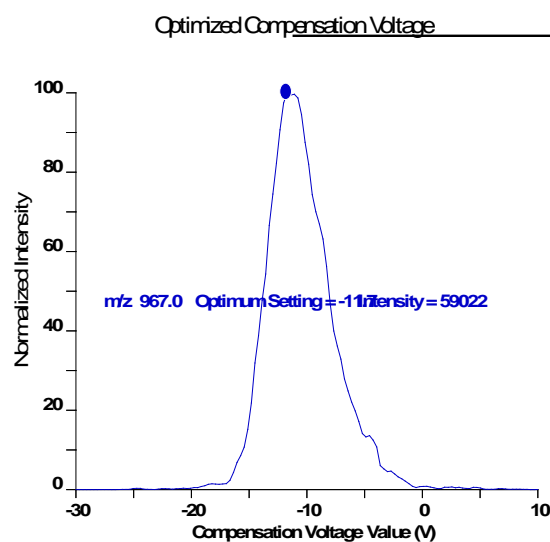
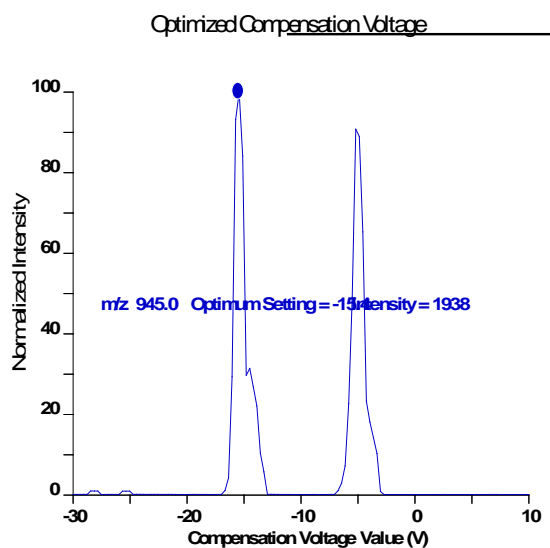


Figure 3.45: Compensation Voltage Scan of CBZ tetramer(H) Figure 3.46: Compensation Voltage Scan of CBZ tetramer(Na)

Like the protonated trimer, the protonated tetramer ($m/z=946$) was not intense enough to make any conclusive judgement about the ion mobility of the species. The sodiated tetramer did have a significant intensity, and the CV scan supported an ensemble of ions with only one ion mobility. As with the trimer, it is unclear how sodium affects the configuration of the proton-bound tetramer, and the mobility determinations of this ion cannot be transferred over to the proton-bound species.

Unfortunately, an acceptable ion count was not achieved to study the ion mobility of the protonated tetramer. This is most likely due to the tendency of FAIMS to break apart weakly-bound multimers before introduction into the mass spectrometer, which decreases the signal disproportionately for trimers and tetramers. The compensation voltage scan of the sodiated

tetramer ion ($m/z=967$) has only one visible peak (Figure 3.46), but a larger intensity is required to make any conclusive statements about the existence of other mobilities. Nevertheless, the connection to the self-assembling crystalline structure is limited since the properties of the sodiated cluster may not transfer to the protonated cluster or neutral molecule. The mobilities of protonated monomer ($m/z=237$), dimer ($m/z=473$), and sodiated trimer ($m/z=731$) were also investigated. Though the trimer also presented inconclusive spectra, the scans of the dimer and monomer overwhelmingly support an ensemble of ions with a single mobility, implying one associated geometry.

Chapter 4: Conclusions

We found the experimental $\Delta_{\text{acid}}\text{H}$ of α -methylserine, L-penicillamine, and 3-methylthreonine to be 1379 ± 23 , 1380 ± 18 , and 1378 ± 23 kJ/mol respectively, which agree with computational predictions within bounds of reasonable uncertainty. Both the computational and experimental values of the methylated-amino acids have a lower $\Delta_{\text{acid}}\text{H}$ than their protein amino acid homologs, meaning they have a higher tendency to act acid as an acid in the gas-phase. In previous studies of these compounds^[32], it was found that the methylated non-protein amino acids had a consistently larger proton affinity than cysteine and serine, meaning these compounds also have a higher tendency to act as a base in the gas phase. As a whole, this implies that methylation increases reactivity of protonation and deprotonation sites within the studied species.

A more comprehensive study of how these alterations may affect peptide behavior comprises an interesting continuation of this study. This could include similar calculations and experimentation done on these base molecules with modified termini or incorporated into peptides. Additionally, the proton transfer mechanisms of the sulfur-containing analogs could be examined through hydrogen-deuterium exchange, similar to the methods of Poutsma and Kass.

We also used High-Field Asymmetric Ion Mobility Spectrometry (FAIMS) and traditional mass spectrometry to characterize the tetramer of carbamazepine and evaluate its relative stability. Our results confirmed that an anomaly existed in both the sodiated and protonated form, which implies a magic number tetramer structure that agrees with the data from our colleagues at Notre Dame.^[2] Subsequently, we found that the compensation voltage scan seemed to support an ensemble of ions with one mobility in the sodiated tetramer, but higher intensities would be required to make any conclusive statements.

With additional resources, an interesting development of this project would revisit the mobility of the protonated and sodiated trimer to draw conclusive evidence about its mobility and

investigate the possible existence of two trimer structures with bonding interactions in consistent with interactions in one or both two-dimensional chiralities. This could be compared to the results of established studies on the trimer formation as studied by Sykes^[40] and possibly help isolate intrinsic and external roles in self-assembly. Applications outside the scope of this thesis could incorporate the findings on the tetramer into existing models for self-assembly.

References:

1. Iqbal, E. S.; Dods, K. K.; Hartman, M. C. T. Ribosomal incorporation of backbone modified amino acids via an editing-deficient aminoacyl-tRNA synthetase. *Org. Biomol. Chem.* **2018**, *16*, 1073-1078.
2. Silski-Devlin, A. M.; Petersen, J. P.; Liu, J.; Turner, G. A.; Poutsma, J. C.; Kandel, S. A. Hydrogen-Bonded Tetramers of Carbamazepine. *J. Phys. Chem. C* **2020**, *124*, 5213-5219.
3. Bell E. A. Non-Protein Amino Acids in Plants. *Encycl. Plant. Phys.* 1975, 403.
4. Rubenstein, E. Biological Effects of and Clinical Disorders Caused by Nonprotein Amino Acids. *Medicine* **2000**, *79*, 80
5. Schroeder, O. E.; Andriole, E. J.; Carver, K. L.; Colyer, K. E.; Poutsma, J. C. Proton Affinity of Lysine Homologues from the Extended Kinetic Method. *J. Phys. Chem. A* **2004**, *108*, 326-332.
6. Raulfs, M. D. M.; Breci, L.; Bernier, M.; Hamdy, O. M.; Janiga, A.; Wysocki, V.; Poutsma, J. C. Investigations of the Mechanism of the “Proline Effect” in Tandem Mass Spectrometry Experiments: The “Pipelicolic Acid Effect”. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1705-1715.
7. . Smith, Z. M.; Steinmetz, V.; Martens, J.; Oomens, J.; Poutsma, J. C. Infrared Multiple Photon Dissociation Spectroscopy of Cationized Canavanine: Side-Chain Substitution Influences Gas-Phase Zwitterion Formation†. *Int J Mass Spectrom* **2018**, *429*, 158-173
8. Silski, A. M.; Brown, R. D.; Petersen, J. P.; Coman, J. M.; Turner, D. A.; Smith, Z. M.; Corcelli, S. A.; Poutsma, J. C.; Kandel, S. A. C–H···O Hydrogen Bonding in Pentamers of Isatin. *J. Phys. Chem. C* **2017**, *121*, 21520-21526.
9. Aoi, W.; Marunaka, Y. Importance of pH Homeostasis in Metabolic Health and Diseases: Crucial Role of Membrane Proton Transport. *Biomed Res Int* **2014**, *2014*
10. Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. Gas Phase Ion and Neutral Thermochemistry. *Journal of Physical and Chemical Reference Data* **1988**, *17*, 5-22.
11. McMurry, J. E.; Fay, R. C. *General Chemistry: Atoms First*; Aqueous Equilibria: Acids and Bases; Pearson Higher Education: 2013; , pp 471-480.
12. Cooks, R.G., Kruger, T.L. Intrinsic Basicity Determination Using Metastable Ions. *J. Am. Chem.* 1977. *99* (4), pp 1279-1291
13. Cooks, R. G.; Wong, P. S. H. Kinetic Method of Making Thermochemical Determinations: Advances and Applications. *Acc. Chem. Res.* **1998**, *31*, 379-386.
14. Poutsma, J. C. Lab Meeting Presentation: Kinetic Method Overview.
15. Bouchoux, G.; Sablier, M.; Berruyer-Penaud, F. Obtaining thermochemical data by the extended kinetic method. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jms.680>

16. Ervin, K. M. Microcanonical analysis of the kinetic method. The meaning of the “apparent entropy”. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 435-452.
17. Cheng, X.; Wu, Z.; Fenselau, C. Collision energy dependence of proton-bound dimer dissociation: entropy effects, proton affinities, and intramolecular hydrogen-bonding in protonated peptides. *J. Am. Chem. Soc.* **1993**, *115*, 4844-4848.
18. Armentrout, P. B. Entropy measurements and the kinetic method: A statistically meaningful approach. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 371-379.
19. Eiceman, G. A.; Karpas, Z.; Jr, H. H. H. *Ion Mobility Spectrometry, Third Edition*; CRC Press: 2013;
20. Edge, A.; Toutoungi, D. An Introduction to Ion Mobility Spectrometry with ultraFAIMS. *Owlstone Medical* **2010**.
21. Toutoungi, D. ultraFAIMS: A New Dimension in Mass Spectrometry. *Owlstone Medical*
22. Thermo Fisher Scientific FAIMS Operators Manual. **2006**.
23. Zrodnikov, Y.; Davis, C. E. The Highs and Lows of FAIMS: Predictions and Future Trends for High Field Asymmetric Waveform Ion Mobility Spectrometry. *J Nanomed Nanotechnol* **2012**, *3*, 109
24. Zhang, J. D.; Mohibul Kabir, K. M.; Lee, H. E.; Donald, W. A. Chiral recognition of amino acid enantiomers using high-definition differential ion mobility mass spectrometry. *International Journal of Mass Spectrometry* **2018**, *428*, 1-7.
25. Mie, A.; Jörntén-Karlsson, M.; Axelsson, B.; Ray, A.; Reimann, C. T. Enantiomer Separation of Amino Acids by Complexation with Chiral Reference Compounds and High-Field Asymmetric Waveform Ion Mobility Spectrometry: Preliminary Results and Possible Limitations. *Anal. Chem.* **2007**, *79*, 2850-2858.
26. Barnett, D. A.; Ells, B.; Guevremont, R.; Purves, R. W. Separation of leucine and isoleucine by electrospray ionization–high field asymmetric waveform ion mobility spectrometry–mass spectrometry. *J Am Soc Mass Spectrom* **1999**, *10*, 1279-1284.
27. Jones, C. M.; Bernier, M.; Carson, E.; Colyer, K. E.; Metz, R.; Pawlow, A.; Wischow, E. D.; Webb, I.; Andriole, E. J.; Poutsma, J. C. Gas-phase acidities of the 20 protein amino acids. *International Journal of Mass Spectrometry* **2007**, *267*, 54-62.
28. Muetterties, C.; Janiga, A.; Huynh, K. T.; Pisano, M. G.; Tripp, V. T.; Young, D. D.; Poutsma, J. C. Gas-phase acid-base properties of homocysteine, homoserine, 5-mercaptonorvaline, and 5-hydroxynorvaline from the extended kinetic method. *International Journal of Mass Spectrometry* **2014**, *369*, 71-80.
29. Muetterties, C.; Touzani, A. D.; Hardee, I.; Huynh, K. T.; Poutsma, J. C. Gas-phase acid-base properties of 1-aminocycloalkane-1-carboxylic acids from the extended kinetic method. *International Journal of Mass Spectrometry* **2015**, *378*, 59-68.
30. Kramer, C. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons: 2002; .

31. Strohmaier, T. Irregular Sampling.
32. Long, D. Determinations of Proton Affinities of Methylated Cysteine and Serine Homologs, The College of William and Mary, 2019.
33. Bernstein, J. *Polymorphism in Molecular Crystals*; Clarendon Press: 2002;
34. Grzesiak, A. L.; Lang, M.; Kim, K.; Matzger, A. J. Comparison of the four anhydrous polymorphs of carbamazepine and the crystal structure of form I. *J Pharm Sci* **2003**, *92*, 2260-2271.
35. Sweeny, B. C.; McDonald, D. C.; Poutsma, J. C.; Ard, S. G.; Viggiano, A. A.; Shuman, N. S. Redefining the Mechanism of O₂ Etching of Aln– Superatoms: An Early Barrier Controls Reactivity, Analogous to Surface Oxidation. *J. Phys. Chem. Lett.* 2020, *11*, 217-220.
36. Bartmess, J. E., "Negative Ion Energetics Data" in **NIST Chemistry WebBook, NIST Standard Reference Database Number 69**, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899
37. G. Bouchoux, S. Huang, B.S. Inda **Acid-base thermochemistry of gaseous aliphatic α -amino acids**. *Phys. Chem. Chem. Phys.*, *13* (2011), p. 651
38. V. Riffet, G. Frison, G. Bouchoux, Acid-base thermochemistry of gaseous oxygen and sulfur substituted amino acids (Ser-Thr-Cys Met) *Phys. Chem. Chem. Phys.*, *13* (2011), p. 18561
39. Zhixin Tian, ‡; Anna Pawlow, †; John C. Poutsma, *; Steven R. Kass*, ‡ Are Carboxyl Groups the Most Acidic Sites in Amino Acids? Gas-Phase Acidity, H/D Exchange Experiments, and Computations on Cysteine and Its Conjugate Base.
- 40 Iski, E. V.; Johnston, B. F.; Florence, A. J.; Urquhart, A. J.; Sykes, E. C. H. Surface-Mediated Two-Dimensional Growth of the Pharmaceutical Carbamazepine. *ACS Nano* **2010**, *4*, 5061-5068.