Hydrogen Deuterium Exchange of Lysine Homologs and Construction of a Nanospray Ionization Source

Rita N. Straus
College of William and Mary

Follow this and additional works at: https://scholarworks.wm.edu/honorstheses

Recommended Citation

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.
Hydrogen Deuterium Exchange of Lysine Homologs and Construction of a Nanospray Ionization Source

by

Rita N. Straus

Accepted for Honors

Professor John C. Peutsma, Director

Professor Gary W. Rice

Professor Mark Forsyth

Williamsburg, Virginia
May 3, 2012
# Table of Contents

**Table of Contents** .................................................................................................................i

**Acknowledgements** .............................................................................................................ii

**List of Figures** .......................................................................................................................iii

**List of Tables** .........................................................................................................................iv

**Abstract** .................................................................................................................................v

**Chapter 1: Introduction**

1.1 Mass Spectrometry .............................................................................................................1

1.2 Hydrogen-Deuterium Exchange .......................................................................................4

1.3 Mass Spectrometry Based Proteomics .............................................................................6

**Chapter 2: Experimental and Computational Methods**

2.1 Gas Phase H/D Exchange Methods ..................................................................................9

2.2 Computational Methods ..................................................................................................13

**Chapter 3: H/D Exchange Behavior of Protonated Lysine Homologs**

3.1 Biological Importance of Lysine Homologs .................................................................15

3.2 H/D Exchange Mechanisms .............................................................................................16

3.3 Results ................................................................................................................................19

   3.3.1 Experimental Results .................................................................................................19

   3.3.2 Computational Results ...............................................................................................25

3.2 Discussion .............................................................................................................................27

**Chapter 4: Proteomics**

4.1 Peptide Fragmentation .......................................................................................................29

4.2 Construction of Nanospray Ionization Source ...............................................................33

4.3 Proteomic Analysis of Tryptic Digests ............................................................................36

**References** ...........................................................................................................................38
Acknowledgements

I would like to thank Dr. Poutsma for his guidance throughout this research project and for all of his help and support over the past four years. I would also like to thank all of my friends and lab mates for their optimism and for always making me laugh! I could not have made it without them. Lastly, I am grateful for my wonderful family, especially my parents, who have always supported me no matter what.
List of Figures

Figure 1.1 – Illustration of a quadrupole mass filter (taken from Cooks et. al.) .......... 3

Figure 1.2 – Neutral Structures of Lys, Orn, DABA, and DAPA .................................. 6

Figure 2.1 – Illustration of ESI Quadrupole Ion Trap Mass Spectrometer .............. 10

Figure 3.1 – H/D exchange mechanisms for Glycine Oligomers (taken from Beauchamp et. al.) .................................................................................................................. 17

Figure 3.2 – Sample Spectra of Lysine, Ornithine, DABA, and DAPA at reaction time 250ms ............................................................................................................................... 21

Figure 3.3 – Graphs of normalized intensity vs. time obtained by Kinfit for the lysine homologs ................................................................................................................. 23

Figure 3.4 – Geometry optimized structures of protonated lysine homologs .......... 25

Figure 3.5 – Potential energy surface for the H/D exchange of DAPA ..................... 26

Figure 4.1 – Nomenclature of N-terminal and C-terminal fragments ..................... 30

Figure 4.2 – Steps for mass spectrometry based proteomics ................................. 37
List of Tables

Table 3.1 - Rate coefficients for the H/D exchange of lysine and its homologs
(x10^{-11} cm^3 molecule^{-1} s^{-1}).................................................................................. 20
Table 3.2 – Calculated energy barriers for the H/D exchange mechanisms .......... 26
Table 4.1 – Sample Time Course Gradient............................................................................. 34
Abstract

Hydrogen-Deuterium Exchange Experiments

The first study presented in this thesis describes the hydrogen-deuterium exchange behavior of protonated lysine and its three shorter homologs ornithine, diaminobutyric acid (DABA), and diaminopropionic acid (DAPA). H/D exchange reactions were conducted in an ion trap mass spectrometer with an external electrospray ionization source. All four of the amino acids studied exhibited complete exchange of all six labile hydrogens, but at significantly different rates. DAPA exchanged at the fastest rates, followed by lysine, and then DABA. Ornithine was shown to exchange at substantially slower rates than the other three homologs. Computational modeling of the H/D exchange reactions were performed to help explain these experimental results. Energetic barriers for the proposed exchange mechanisms do not sufficiently account for the large difference in exchange rate coefficients between DAPA and ornithine. Additional modeling must be completed to fully explain this phenomenon.

Construction of Nanospray Ionization Source and Proteomics Experiments

A nanospray ionization source was also constructed for interfacing the ion trap mass spectrometer with high performance liquid chromatography. This instrument was used to separate and analyze peptides from tryptic digestions. The superior sensitivity of nanospray was needed to detect small quantities of peptide from complex mixtures. This instrument will be used to study the protein expressions of the Cluster K bacteriophage. Tandem mass spectrometry, in conjunction with the sequencing software SEQUEST, will be used to identify bacterial and viral proteins from tryptic digest samples.
Chapter 1 – Introduction

1.1 Mass Spectrometry

Mass spectrometry is an analytical technique in which ions are separated by their mass to charge (m/z) ratio. It can be used to identify and quantify compounds, to study thermodynamic properties, and to probe chemical structure. Mass spectrometry has widespread applications to environmental chemistry, forensic analysis, and medicine. For example, it has been used to measure the levels of pesticides in soil, to detect residues of explosive chemicals, and most recently to detect cancer biomarkers.

The first step to performing mass spectrometry involves creating gas phase ions. This has traditionally been done using either electron impact or chemical ionization. Electron impact is an ionization method in which electrons collide with gas phase molecules to create fragment ions. It is referred to as a hard ionization technique because parent molecules are subjected to high-energy collisions. Chemical ionization utilizes lower energy collisions and therefore produces fewer fragment ions. These methods are very useful for studying small volatile organic compounds, but are unsuitable for studying nonvolatile molecules such as proteins and amino acids. It was not until the invention of soft ionization techniques in the 1980’s that large biomolecules could be successfully studied using mass spectrometry. Electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) are the most common of these techniques.

When performing ESI ions are first formed in solution by exposing the molecule of interest to either an acidic or basic solvent. Acidic solutions result in protonation (MH$^+$), while basic solutions result in deprotonation (M-H$^-$). An aerosol spray consisting
of tiny charged droplets is then generated by passing the sample of ions across a high potential. These droplets become increasingly smaller until each droplet is a single analyte ion. Since multiply charged ions can be produced, it is possible to observe higher mass compounds at lower mass to charge ratios. This allows for the study of large proteins.

MALDI is a method of vaporization and ionization that can also be used to produce high molecular weight molecules. The molecule of interest is first dissolved in a solution containing a matrix material that absorbs light of a specific wavelength, and the solvent is then evaporated off so that analyte molecules remain in a solid matrix. A pulsed laser, which emits light corresponding to this wavelength, is used to desorb the molecules from the matrix. Ionization results from a transfer of absorbed energy from the matrix to the analyte.  

Once ions have been created they are separated within the mass analyzer component of the mass spectrometer. The quadrupole mass filter is one of the most common analyzers and consists of two pairs of parallel rods. Positive DC and RF voltages are applied to one pair of the rods, while negative voltages are applied to the other pair. Ions pass through the center of the quadrupole and oscillate in a path that is dependent upon its mass to charge ratio. For a given applied voltage, only molecules of a specific m/z ratio will oscillate in a path that allows them to pass through the end of the quadrupole and reach the detector (see Figure 1.1).  

Triple quadrupole instruments have become popular for conducting tandem mass spectrometry experiments, with the second quadrupole being used as a collision cell.
The quadrupole ion trap mass analyzer also utilizes electric fields to separate charged particles. The three dimensional trap consists of two endcap electrodes and a ring electrode. Ions are oscillated in all three dimensions and trapped in the center of the three electrode system. The RF voltage is then scanned so that increasingly higher m/z ratios become unstable and are ejected to the detector. One advantage of the ion trap is its ability to perform multiple stages of mass spectrometry in one analyzer (MSn)3-5. All experiments presented in this thesis were done in an ion trap mass spectrometer instrument.

Tandem mass spectrometry is often used to study fragmentation or to perform ion molecule reactions. In this lab fragmentation is achieved through collision induced dissociation (CID). CID makes use of an inert gas (usually helium or argon) to fragment molecular and pseudo-molecular gas phase ions. The parent ions are first accelerated in an electric field and then allowed to interact with the collision gas. Dissociation occurs via a transfer of kinetic energy from the neutral gas to the sample ions. Tandem mass spectrometry allows a single m/z ratio to be isolated and then fragmented so that only
peaks associated with the analyte are detected. Ion molecule reactions are another application of tandem mass spectrometry and involve studying the reaction of a gas phase ion with a neutral molecule. This can be done in a tandem mass spectrometer by first isolating the ion of interest, and then allowing it to interact with reagent molecules. These types of experiments can be used to study reaction thermodynamics and kinetics, as well as to investigate gas phase structure.6

1.2 Hydrogen-Deuterium Exchange

Hydrogen-deuterium (H/D) exchange is a chemical reaction in which a deuterium atom is replaced by a hydrogen atom. It serves as a useful tool in assessing protein structure, as well as in studying thermodynamic properties of biological molecules. Specifically, it can be utilized to study the energetics of protein folding. This information is important because there is usually only one conformation of a protein that is biologically active. Denaturation or improper folding often results in loss of function, and can lead to disease. Studies on the thermodynamic stability of proteins can also be used to examine the strength of protein-protein and protein-ligand interactions. This is important for the development of drugs that aim to inhibit proteins through these types of interactions.

H/D exchange can be performed by exposing a molecule to a deuterated solvent. When a protein is exposed to a deuterated solvent, hydrogens that are more highly accessible will exchange quickly, while those that are buried in the hydrophobic core will exchange slowly, if at all. As a protein unfolds, more exchange sites become available. There are a variety of H/D exchange methods present in the literature that are used to explore the conformational structure of proteins. Most of these techniques are only
concerned with the exchange of backbone amide hydrogens since they tend to exchange more slowly than carboxylic and side chain hydrogens. Exchange characteristics of these amide hydrogens are also more strongly related to the tertiary structure of the protein. A two state model is often used to understand the relationship between the conformation and the extent of deuterium incorporation. In this model it is assumed that for any slowly exchanging amide hydrogen, a folded conformation where the hydrogen is buried, is in equilibrium with an unfolded conformation where the hydrogen is accessible. Since equilibrium exists between the open and closed states, it can be thought of as the molecule cycling back and forth between two conformations.

Closed (NH) ⇌ Open (NH) → Exchanged (ND)

Longer reaction times allow for more cycles, and therefore more total time spent in the unfolded state. This implies that a longer reaction time will result in a higher probability of exchange at each protected site.\(^7^8\)

Mass spectrometry has recently become a popular method for detecting H/D exchange. Although other types of spectroscopy, such as NMR, IR spectroscopy, and circular dichroism, can be used for detection, mass spectrometry offers several advantages. These include improved sensitivity, the ability to study higher molecular weight compounds, and significantly faster data collection.\(^9\) The use of mass spectrometry has also enabled the reaction to be studied in the gas phase. Although solution phase H/D exchange has become commonplace, the mechanisms of gas phase H/D exchange have not been studied extensively. It is beneficial to study the reaction in the gas phase, because it allows for the determination of chemical properties in the absence of solvent effects.
Since very little is known about the gas phase mechanism, it is useful to study the reaction using amino acids rather than proteins or other large biomolecules. These simpler species are easier to understand and their mechanisms can be modeled computationally. The study presented here examines the exchange mechanisms of the protonated amino acid lysine and its three shorter homologs, ornithine, diaminobutyric acid (DABA), and diaminopropionic acid (DAPA) whose neutral structures are depicted in Figure 1.2.

![Neutral Structures of Lys, Orn, DABA, and DAPA](image)

Figure 1.2 – Neutral Structures of Lys, Orn, DABA, and DAPA

1.3 Mass Spectrometry Based Proteomics

Proteomics is the large scale study of protein structure and function. The term proteome was first used by Marc Wilkins in 1994 and refers to all of the proteins expressed by an organism’s genome. Proteins are a vital component of all living organisms, and are involved in almost every biological process. They act as enzymes to speed up biochemical reactions, transport molecules between cells, initiate immune
responses, and play an essential role in cell differentiation. Proteomics often involves using sequencing methods to determine the amino acid make-up of a particular protein or set of proteins encoded by a genome. It can also be used to study post translational modifications and to detect biomarkers of disease. Detailed knowledge about the composition of genetically encoded biomolecules is important for better understanding genetic disorders, as well as for the development of drugs that target specific biological pathways.

Recent advancements in the analytical techniques for studying proteins have made proteomics a rapidly growing area of research. Mass spectrometry, along with advanced separation methods, has risen to the forefront of proteomic analysis. Mass spectrometry is such a powerful tool because it allows for identification and quantification of a large number of proteins from complex mixtures. The high sensitivity of mass spectrometry, combined with sophisticated separation techniques also allows for the detection of low abundance proteins and peptides.\(^{10}\)

There are currently two different mass spectrometric approaches for protein characterization. In top down proteomics intact protein ions are introduced into a mass spectrometer, and then fragmented into smaller peptides using tandem mass spectrometry. This method eliminates the need for protein digestion, and also allows for the complete sequence to be obtained. It does however produce convoluted spectra, and is generally only useful for analyzing a single isolated protein or simple mixture of proteins.\(^{11}\)

In bottom up proteomics a proteolytic digest containing a mixture of peptides is analyzed by mass spectrometry coupled with liquid chromatography. In this technique
gel electrophoresis is initially used to separate cellular proteins. Gel electrophoresis is performed by placing a protein mixture in a polyacrylamide gel matrix, and adding an anionic detergent, usually sodium dodecyl sulfate (SDS), to denature and bind to the proteins. This creates a uniform negative charge on all of the molecules, which are then exposed to an electric field, and separated based on differences in ion mobility. Since all of the molecules have the same charge, it is differences in mass that drive the separation. Given that it is possible for two different proteins to have similar masses, it often useful to separate the molecules in two dimensions. In two-dimensional gel electrophoresis proteins are separated by isoelectric point before they are separated by size. An electric potential is applied to the gel so that one end is more positively charged than the other. Proteins will then travel across the gel at different rates depending on their charge. Once this stage of separation is complete, digestion can be achieved using the protease enzyme trypsin. Trypsin cleaves proteins via hydrolysis at the carboxylic side of the amino acids lysine and arginine. Bottom up proteomics allows trace amounts of protein to be identified from complex samples. It does not however, provide complete sequence coverage, and in most cases only a small portion of the peptide fragments are identified.

The work presented here describes the construction of a nanospray ionization source for the interface of mass spectrometry with high performance liquid chromatography (HPLC). This instrument will be used to identify and characterize the expressed proteins of the Cluster K bacteriophage.
Chapter 2 – Experimental and Computational Methods

2.1 Gas Phase H/D Exchange Methods

All experiments were performed in a Finnigan LCQ DECA ion trap mass spectrometer that has been modified to allow the introduction of D₂O reagent gas. Amino acid solutions of betaine, lysine, ornithine, diaminobutyric acid (DABA) and diaminopropionic acid (DAPA) were made at concentrations of approximately 5x10⁻⁵ M. To induce protonation of the analyte molecules, solutions were made in 49.5:49.5 H₂O:MeOH solvent with 1% acetic acid.

D₂O reagent was leaked into the helium line of the mass spectrometer at flow rates ranging from 200 to 400 µl/hr using a 500 µl Hamilton syringe and an automated syringe pump. D₂O was given approximately 1 hour to flood the ion trap before any data was collected. This allowed water on the surface of the trap enough time to exchange, and ensured that D₂O abundance was high relative to H₂O. Sample solutions were introduced into an electrospray ionization source at 800 µl/hr also via a 500 µl Hamilton syringe and automated syringe pump. A 5,000 V potential was applied to the source to produce a charged aerosol spray. Ions were desolvated and carried through a heated capillary by nitrogen sheath gas at a low flow rate (20 arbitrary units). The temperature of the heated capillary was maintained at 125⁰C. After leaving the heated capillary, ions were focused into the low pressure region of the mass analyzer (approximately 1 mTorr). Finally, the ion optics, which consists of a quadrupole, interoctapole lens, and an octapole, sent ions through the entrance endcap electrode into the ion trap.
After sample introduction, conditions were optimized to provide the highest possible ion count for the protonated amino acid of interest. This was achieved using the automatic tuning feature of the LCQ Tune™ software, which functioned as the user interface to the mass spectrometer. This tuning program adjusts the voltages applied to the source and the offset voltages of the focusing lenses to attain maximum signal on a specified m/z ratio.

The amino acid under study was isolated in ms/ms zoom scan mode and allowed to react with D$_2$O for up to 20 seconds. For the lysine homologs, isolation widths between 15 and 20 amu were used, and for betaine isolation widths between 5 and 10 amu were used. Reaction time was gradually increased until all labile hydrogens exchanged, and then gradually decreased until the protonated parent was again the most abundant peak. This method of bringing the reaction time up and then back down
provided the best correction for drifts in pressure that tend to occur over long experiments. Spectra were averaged over 10 scans and saved at each reaction time.

When possible, all four lysine homologs were studied on the same day. This helped to keep conditions (particularly pressure) inside the trap as constant as possible. 500µl of 49.5:49.5 H₂O:MeOH with 1% acetic acid flushing solution was run through the electrospray source in between experiments to clean out the ion trap.

The H/D exchange reaction is highly dependent upon pressure, but there is no way to directly determine the pressure inside of the ion trap. The ion gauge is the only location in the vacuum region of the mass spectrometer where pressure can be measured. It was adjusted by changing the flow of helium via a microneedle valve, and was usually kept between 1.3 and 2.0 x 10⁻⁵ Torr. Betaine is a simple molecule with only one labile hydrogen of a known rate coefficient, and can be used to determine the pressure of D₂O inside the trap. Betaine exchanges its hydrogen at a rate of 1x10⁻¹² cm³mole⁻¹s⁻¹. This value, along with an experimentally determined relative rate coefficient, was used to find the concentration of D₂O inside the trap on each day that data was collected.

Qualbrowser was used to view raw data files. Ion intensities for masses corresponding to d0 through d6 at each activation time were exported into Excel and then normalized to account for differences in total ion signal. Normalized intensities were fit to a set of ordinary differential equations as a function of time using the Kinfit kinetic fitting program.¹⁴ This program provided relative rate coefficients for each observed exchange. The equations entered into Kinfit assumed pseudo first order reaction kinetics because of the belief that D₂O concentration was significantly higher than amino acid concentration. The simplest equations did not account for isotope peaks or the possibility
of back exchange. Back exchange refers to a reaction were a deuterium is replaced by a hydrogen at a site that has already undergone exchange. These simple equations are shown below. D0 through D6 represents relative ion abundance, and k1 through k6 are the rate coefficients for each exchange.

\[
D_0 = -k_1 \times D_0 \\
D_1 = (k_1 \times D_0) - (k_2 \times D_1) \\
D_2 = (k_2 \times D_1) - (k_3 \times D_2) \\
D_3 = (k_3 \times D_2) - (k_4 \times D_3) \\
D_4 = (k_4 \times D_3) - (k_5 \times D_4) \\
D_5 = (k_5 \times D_4) - (k_6 \times D_5) \\
D_6 = k_6 \times D_5
\]

Another set of equations was written to account for carbon -13 isotope peaks. Since an amino acid with a C13 atom will appear at a higher m/z ratio, it could be mistaken for an amino acid with a deuterium incorporated. An isotope pattern calculator was used to find the amount of C13 expected to exist for the molecule of interest. This is symbolized by t1 in the equations shown below.

\[
D_0 = -k_1 \times Y_0 \\
D_1 = (k_1 \times D_0) - k_2 \times (D_1 - (t_1 \times D_0)) \\
D_2 = k_2 \times (D_1 - (t_1 \times D_0)) - k_3 \times (D_2 - (t_1 \times D_1)) \\
D_3 = k_3 \times (D_2 - (t_1 \times D_1)) - k_4 \times (D_3 - (t_1 \times D_2)) \\
D_4 = k_4 \times (D_3 - (t_1 \times D_2)) - k_5 \times (D_4 - (t_1 \times D_3)) \\
D_5 = k_5 \times (D_4 - (t_1 \times D_3)) - k_6 \times (D_5 - (t_1 \times D_4)) \\
D_6 = k_6 \times (D_5 - (t_1 \times D_4))
\]
The final set of equations used accounted for back exchange reactions, and included 12 rate coefficients to describe both the forward and reverse reactions as follows.

\[
\begin{align*}
D_0 & \xrightleftharpoons[k_{12}]{k_1} D_1 \xrightleftharpoons[k_{11}]{k_2} D_2 \xrightleftharpoons[k_{10}]{k_3} D_3 \xrightleftharpoons[k_9]{k_4} D_4 \xrightleftharpoons[k_8]{k_5} D_5 \xrightleftharpoons[k_7]{k_6} D_6 \\
\end{align*}
\]

These equations are depicted below.

\[
\begin{align*}
D_0 &= (-k_1 \times D_0) + (k_{12} \times D_1) \\
D_1 &= (k_1 \times D_0) + (k_{11} \times D_2) - (k_2 \times D_1) - (k_{12} \times D_1) \\
D_2 &= (k_2 \times D_1) + (k_{10} \times D_3) - (k_3 \times D_2) - (k_{11} \times D_2) \\
D_3 &= (k_3 \times D_2) + (k_9 \times D_4) - (k_4 \times D_3) - (k_{10} \times D_3) \\
D_4 &= (k_4 \times D_3) + (k_8 \times D_5) - (k_5 \times D_4) - (k_9 \times D_4) \\
D_5 &= (k_5 \times D_4) + (k_7 \times D_6) - (k_6 \times D_5) - (k_8 \times D_5) \\
D_6 &= (k_6 \times D_5) - (k_7 \times D_6)
\end{align*}
\]

It was determined that the isotope peaks were relatively small and accounting for them in the differential equations had minimal impact on results, but that back exchange was significant and needed to be included. This was especially true at longer reaction times where the amino acids had more time to undergo back exchange. The same principles were used to construct equations for fitting betaine. The rate coefficients obtained from Kinfit are relative rate coefficients, and to determine absolute rate coefficients relative values were divided by the concentration of D₂O in the trap on the day the experiment was conducted.

2.2 Computational Methods

Density functional theory (DFT) calculations were performed using GaussView 2.1 and Gaussian 98W to assist in the interpretation of experimental results and to gain insight into the H/D exchange mechanisms. DFT is a quantum mechanical method for
studying chemical systems, and uses electron densities rather than wave functions to compute energies. It is based upon the Hohenberg-Kohn theorem, which states that a system is uniquely described by functionals of the electron density. This method is considered a semi-empirical method, because it uses empirical values as starting parameters. This helps to reduce calculation time, but also means that the calculations are no longer based solely on first principles.\textsuperscript{15}

In this study, starting structures were drawn in GaussView and then optimized using DFT at the B3LYP level of theory and 6-31+G* basis set. Potential energy surfaces of the H/D exchange mechanisms for ornithine, DABA, and DAPA were constructed, and barriers for the reactions were calculated. The potential energy surface for protonated lysine was calculated previously, but at a different basis set. At this time the potential energy surfaces are not complete, as some transition states still need to be found.
Chapter 3 – H/D Exchange Behavior of Protonated Lysine Homologs

3.1 – Biological Importance of Lysine Homologs

Lysine is one of the twenty amino acids that are encoded by the human genome for protein synthesis. Lysine has a high proton affinity and is often the favored site of protonation in proteins. In previous work the Poutsma group has found the gas phase proton affinity to be approximately 1004 kJ/mol. As a result of its high basicity, lysine has a positively charged side chain at physiological pH. These hydrophilic properties cause lysine-containing proteins to fold in such a way that lysine residues are located on the exterior portion of the structure. This makes them easily accessible for interaction with other molecules. For example, lysine is often responsible for stabilizing negatively charged substrates in the active sites of enzymes. A better understanding of the chemical properties of lysine could provide useful insight into protein structure and dynamics. Lysine and arginine residues are also the sites of cleavage in trypsin hydrolysis, so lysine has significant implications for proteomics as well. Ornithine is a homolog of lysine that contains one less methylene group in the side chain, and it plays an important role in the urea-cycle.

The shorter homologs, ornithine, DABA, and DAPA, are non-coding amino acids, and do not appear in naturally occurring proteins. Although these homologs seem to be very structurally similar, they have different chemical properties. Non-protein amino acids can misincorporate into biologically active proteins, and interfere with proper folding mechanisms which is thought to be a possible cause of disease. Even though non-protein amino acids are not as widely studied as protein amino acids, understanding their chemical behavior clearly has important biological applications.
3.2 – Gas Phase H/D Exchange Mechanisms

The H/D exchange mechanisms are thought to depend on a variety of factors including the relative basicities of the exchange site and deuterating agent, site of protonation, hydrogen bonding, and the overall gas phase conformation. The effect of the relative proton affinities of the deuterated reagent and the exchange site has been explored in previous studies. The extent of exchange is expected to depend on the endothermicity of a proton transfer from the analyte molecule to the reagent. Studies have shown that peptides containing highly basic residues such as arginine exhibit very slow exchange. This suggests that for isotopic exchange to occur it must be energetically feasible for the molecule to transfer a proton.

Beauchamp et. al. has explored the effects of deuterating reagent basicity on the exchange rates of protonated glycine oligomers in an Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer. As expected, they found that the rate and extent of exchange increases significantly with reagent basicity. ND$_3$ was the most basic deuterating agent used in this study and resulted in exchange at all labile hydrogens. It was also observed that peptides could undergo multiple exchanges in a single collision event with the reagent. An onium ion mechanism, in which a proton is transferred from the N-terminus to the ND$_3$ molecule, was proposed for amino hydrogen exchange. Even though this transfer is endothermic, it is made possible by the simultaneous stabilization of the ammonium ion by the resultant neutral oligomer. This helps to explain the occurrence of multiple exchanges in one step. A tautomer mechanism has also been suggested for amino hydrogen exchange with ND$_3$. It is a concerted reaction in which a
A proton is transferred from the N-terminus to an amide carbonyl oxygen while an amide proton is concurrently transferred to the reagent base.

D$_2$O has a significantly lower proton affinity than ND$_3$ and it is therefore not energetically possible for it to form an onium ion or tautomer intermediate. A relay mechanism has been proposed to explain the exchange of amino hydrogens with less basic reagents. This process involves the relay of a proton between the amino group and the slightly less basic carbonyl oxygen. In other words, the deuterium is not incorporated in the same place that the hydrogen is lost. Beauchamp et. al. also presents the salt bridge and flip-flop mechanisms for exchange at the carboxylic acid hydrogen. The formation of a salt bridge with ND$_3$ reagent occurs through deprotonation at the C-terminus and stabilization by the protonated amino group. The flip-flop mechanism occurs with less basic agents. All of these proposed mechanisms are depicted below in Figure 3.1.

![H/D exchange mechanisms for Glycine Oligomers](taken from Beauchamp et.al.)
The mechanisms just described were for molecules containing the very simple amino acid glycine. The side chain of glycine is just a single hydrogen atom, and will therefore not get involved in H/D exchange reactions. It is likely that side chains containing heteroatoms will alter the mechanisms by which H/D exchange occurs. The amino acids studied here contain an amino group in the side chain. We were interested in understanding what role the side chain plays in the exchange process. In particular, the effects of side chain length and intramolecular hydrogen bonding were investigated.

The gas phase structure and H/D exchange behavior of protonated lysine has been studied previously. Most recently Rozman et. al. found that lysine exchanges all six labile hydrogens in an FT-ICR instrument. Five equivalent exchanges and one slightly slower exchange were observed. This is slightly surprising considering the high gas phase basicity of lysine, and suggests that the reaction does not proceed via direct proton transfer. They also performed density functional theory calculations to probe the gas phase conformation of protonated lysine. These studies have shown that the side chain is the preferred site of protonation, and that the molecule is stabilized by two intramolecular hydrogen bonds. These interactions involve the protonated side chain bonding to the lone pair of electrons on the N-terminus and to the oxygen on the C-terminus. This is notably different from the zwitterionic structure of lysine in solution.

The possibility of gas phase zwitterions of protonated amino acids has been explored, and although the non-zwitterion conformation is more stable on its own, it is thought that a H$_2$O (or D$_2$O) molecule can have a stabilizing effect on ion-zwitterion structures. This is an important consideration because it could potentially compete with H/D exchange reactions. This is especially true for amino acids with a high gas phase
basicity such as arginine and lysine. Rozman et. al. constructed potential potential energy surfaces mapping out the perturbation of protonated lysine from the ion-molecule structure to the ion-zwitterion structure. Density functional theory calculations were also performed to construct potential energy surfaces for the proposed mechanisms of H/D exchange of lysine. This theoretical data has suggested that isotopic exchange is energetically favorable to the water induced conformational change.

Rozman et. al. also explored the possibility of the flip-flop mechanism at the carboxylic hydrogen and a relay mechanism for amine hydrogens with D$_2$O as the reagent base. The flip-flop mechanism is identical to the one presented above, but the relay mechanism for lysine differs slightly from the glycine oligomer relay mechanism. Lysine has a basic side chain amino group that serves as the relay site with the N-terminus, rather than the carbonyl oxygen.$^{21-22}$

Lysine has also been studied by Lebrilla et. al. who determined that the exchange efficiency of protonated lysine was only weakly dependent upon the difference in proton affinity between the deuterating agent and exchange site. This difference in proton affinity for amino acids with alkyl side chains on the other hand, was strongly related to the extent of exchange.$^{23}$ These results also support the existence of a relay mechanism between the amino groups in lysine. This mechanism helps to explain how lysine can undergo such efficient exchange despite having a high proton affinity.

3.3 – Results

3.3.1 – Experimental Results

The gas phase H/D exchange behavior of protonated lysine and its three shorter homologs ornithine, DABA, and DAPA were studied. Experiments were carried out as
described in section 2.1. All four molecules were observed to exchange all six 
exchangeable hydrogens. Protonated lysine showed 3 fast exchanges, 2 slightly slower 
one, and 1 considerably slower exchange. The other amino acids also followed this 
general trend, but at significantly different rates. Surprisingly, exchange rate coefficients 
did not vary proportionally with side chain length. DAPA, the homolog with the shortest 
side chain, exhibited the fastest exchange rates followed by lysine and then DABA. 
Ornithine exchanged at substantially slower rate than the other three amino acids. See 
Table 3.1 for the absolute rate coefficients of lysine, ornithine, DABA, and DAPA.

<table>
<thead>
<tr>
<th></th>
<th>k1</th>
<th>k2</th>
<th>k3</th>
<th>k4</th>
<th>k5</th>
<th>k6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys</td>
<td>14.3±0.539</td>
<td>12.9±1.53</td>
<td>10.5±1.44</td>
<td>7.67±1.40</td>
<td>5.29±1.08</td>
<td>2.69±0.781</td>
</tr>
<tr>
<td>Orn</td>
<td>0.531±0.272</td>
<td>0.467±0.239</td>
<td>0.393±0.210</td>
<td>0.292±0.157</td>
<td>0.214±0.112</td>
<td>0.139±0.074</td>
</tr>
<tr>
<td>DABA</td>
<td>9.89±2.96</td>
<td>9.31±2.77</td>
<td>8.46±2.55</td>
<td>6.63±2.35</td>
<td>4.62±1.88</td>
<td>2.49±0.895</td>
</tr>
<tr>
<td>DAPA</td>
<td>24.4±1.49</td>
<td>24.3±0.669</td>
<td>20.9±0.819</td>
<td>16.2±0.324</td>
<td>11.6±1.41</td>
<td>6.09±0.718</td>
</tr>
</tbody>
</table>

Table 3.1 - Rate coefficients for the H/D exchange of lysine and its homologs (x 10^{-11} cm^3 molecule^{-1} s^{-1})

Data had to be collected at two different ion gauge pressures to account for the 
varying exchange rates. It was not possible to observe all six of ornithines exchanges at 
pressures that were suitable for studying the other homologs. At these higher pressures, 
however, the decay of protonated DAPA could not be studied, as it exchanged too 
quickly under these conditions, and even at the lowest possible activation time D1 was 
the most abundant peak. The resulting absolute rate coefficient values can still be
compared because betaine was analyzed at both experimental pressures. Therefore the concentration of D$_2$O in the trap could be determined at each ion gauge pressure. As described earlier, to obtain absolute rate coefficients relative values were divided by the concentration of D$_2$O present in the ion trap. Sample spectra and Kinfit graphs are shown in Fig 3.2 and Fig 3.3 respectively.
Figure 3.2 – Sample Spectra of Lysine, Ornithine, DABA, and DAPA at reaction time 250ms
Figure 3.3 – Graphs of normalized intensity vs. time obtained by Kinfit for the lysine homoolgs
3.3.2 – Computational Results

Density functional theory calculations were performed to further our understanding of the H/D exchange process. Lowest energy structures of protonated lysine, ornithine, DABA, and DAPA were first found, and are displayed below in Figure 3.4.

![Geometry optimized structures of protonated lysine homologs](image)

All of these structures show that a hydrogen bond exists between the side chain amino group and the N-terminus. These interactions cause the amino acids to form intramolecular rings of varying sizes, and suggest that a relay mechanism is responsible for the H/D exchange of amino hydrogens. Potential energy surfaces were calculated for ornithine, DABA, and DAPA at the B3LYP/6-31+G* level of theory. Energy barriers for both reactions did not vary significantly amongst the amino acids (see Table 3.2). The potential energy surface for DAPA is shown below. The question marks indicate that the
barriers for the movement of a D$_2$O molecule from its position in the global minimum structure to its position in the reaction site have not yet been found.

<table>
<thead>
<tr>
<th></th>
<th>Orn</th>
<th>DABA</th>
<th>DAPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flip-Flop Barrier (kcal/mol)</td>
<td>13.8</td>
<td>12.6</td>
<td>13.0</td>
</tr>
<tr>
<td>Relay Barrier (kcal/mol)</td>
<td>10.5</td>
<td>6.7</td>
<td>8.6</td>
</tr>
</tbody>
</table>

Table 3.2 – Calculated energy barriers for the H/D exchange mechanisms

Figure 3.5 – Potential energy surface for the H/D exchange of DAPA
3.4 – Discussion

Consistent with other studies, lysine was shown to exchange all six labile hydrogens with rate constants on the same order of magnitude as those presented above. These earlier experiments found that lysine had 5 equivalent exchanges, and therefore 5 equivalent rate coefficients. The disparity between the results of that study and the one presented here come from differences in how data was analyzed. The Kinfit fitting routine used in this study does not allow any assumptions to be made regarding where deuterium incorporation occurred. For example, not all of the D1 species detected in the ion trap will have a deuterium atom in the same location. In other words, the rate coefficients calculated here are not site specific. When site specific fitting routines are used, as they were in the earlier study, each rate coefficient can be attributed to a specific location on the molecule. This allows for the comparison of rate coefficients at different sites on the protonated lysine molecule. In this study only differences in rates between different amino acids can be compared.

Ornithine exhibited much slower exchange than the other amino acids. There are a few possible explanations for this phenomenon that were explored. The rate coefficients of the homologs are not directly related to the side chain length, indicating that there is an entropic factor involved in the exchange process. This is thought to be dependent upon intramolecular ring size. We believe that a hydrogen bond between the side chain and the N-terminus, puts these two groups in an ideal position for undergoing exchange via the relay mechanism. If this hydrogen bond is too weak to form a stable intramolecular ring, exchange will be hindered. On the other hand, if the hydrogen bond
is too strong, it will be energetically unfavorable for a D₂O molecule to enter the relay site.

It was initially expected that the reaction barriers for the H/D exchange of protonated ornithine would be significantly higher than those for the other homologs. Computational modeling, however, has shown that the barriers do not vary enough among the lysine homologs to account for the large differences in rate coefficients between DAPA and ornithine. The most likely explanation for the varying exchange rates, therefore, is differences in the energy barrier for the initial insertion of the D₂O molecule into the exchange site. This would be consistent with ornithine having a strong intramolecular hydrogen bond. As mentioned previously, these barriers have not yet been found.

Overall, experimental and computational results indicate that the presence of an intramolecular hydrogen bond greatly influences the rates at which the lysine homologs undergo H/D exchange. The relay mechanism is proposed as the most likely explanation for exchange at the highly basic amino group hydrogens, and accounts for the high extent of exchange despite the large proton affinities of the homologs. Differing stabilities of the intramolecular rings among the lysine homologs explain why the exchange rates were not directly related to side chain length.
Chapter 4 – Proteomics

4.1 – Peptide Fragmentation

Understanding the way in which peptides fragment inside of a tandem mass spectrometer is crucial for performing proteomic analysis. Their fragmentation patterns depend on the method of fragmentation used as well as the amino acid makeup of the peptide. These patterns have been studied for a large number of peptide sequences, and there is now a growing database of ‘mass spectra fingerprints’ that can be used to unambiguously identify proteins. Gaining insight into the fragmentation patterns and mechanisms of small peptides has been an ongoing effort in the Poutsma group.

Collision-induced dissociation has been used in this study, and produces fragment ions using relatively weak collisions with an inert gas. These weak collisions tend to cause breakages at the amide bond, as well as loss of posttranslational modifications. For every collision two fragments are produced, but it is only the one that retains charge that can be detected by mass spectrometry. When the peptide breaks in such a way that the charge is maintained on the N-terminus, it is referred to as a $b_n^+$-type ion using the Roepstoff nomenclature scheme. If the charge is on the C-terminus it is referred to as a $y_n^+$-type ion (see Figure 4.1).
The most comprehensive model currently used to understand how protonated peptides dissociate under low energy collisions is called the mobile proton model. Peptides can be protonated at a variety of different sites, including the N-terminal amino group, amide oxygens and nitrogens, as well as the side chains of the more basic residues. The dissociation mechanisms are thought to be driven by the movement of this ionizing proton. When there is one site on the peptide that is preferentially protonated, more energy will be required to transfer that proton to other, less energetically favorable, sites. As increasing amounts of vibrational
energy is given to the molecule, more of these protonation sites become available. Molecular orbital calculations have been used to show that unless the ionizing proton is located on a highly basic side chain residue, it will tend to migrate to an amide oxygen upon excitation. This migration actually makes the amide bond stronger, and may seem counterintuitive as a means of directing cleavage at the amide bond. If enough excitation energy is applied it is possible to induce protonation at the amide nitrogen, but this rarely occurs under the low energy excitations utilized in CID.

The mobile proton model suggests that dissociation does not occur via direct cleavage of the amide bond, and that a more complex mechanism is responsible for fragmentation. A better understanding of these mechanisms is needed for predicting which fragments will appear in a mass spectrum for a given peptide. Experimental evidence has suggested that fragmentation proceeds by nucleophilic attack on the carbon center of the protonated amide oxygen to form an oxazolone derivative intermediate. Thermochemical properties decide which fragment keeps the ionizing proton. Wysocki et. al. have explored the effects of proton affinity on fragmentation pathways, and have observed that peptide fragments with higher gas phase proton affinities are more abundant in CID spectra. This suggests that the fragment with the higher basicity retains the extra proton.

Fragmentation rules are still being established, and experiments are showing that fragmentation is often preferential at specific amino acid residues. For example, proline-containing peptides tend to cleave at the amide bond N-terminal to the proline residue and produce primarily y-type ions. This trend is referred to as
the proline effect, and was initially thought to be caused by the high proton affinity of proline relative to other amino acids with alkyl side chains. This explanation was explored further by Vaisar and Urban, who studied peptides containing the non-protein amino acid pipecolic acid. Pipecolic acid is a homolog of proline that contains a six membered pyrolidine ring, and has a comparable gas phase basicity. CID spectra of these Pip containing peptides, however, did not show preferential formation of the expected y ions. This indicates that proton affinity is not responsible for the proline affect. Vaisar and Urban proposed that increased ring strain in forming a bicyclic b-ion leads to the observed proline affect, but this possibility is still being explored.27

Electron-capture dissociation and electron transfer dissociation are additional methods of dissociating protonated peptides and often result in fragmentation at the stronger bonds along the peptide backbone. They therefore produce information that is complementary to CID. Electron capture dissociation is performed by allowing low energy electrons to interact with trapped gas phase ions. Electron transfer dissociation is similar except that fragmentation occurs via interactions with radical anions rather than direct electron transfer. These methods of dissociation tend to produce c and z-type ions as shown in the figure above.

Information on peptide fragmentation patterns has led to the development of large databases for protein identification. These databases are still expanding, and a better understanding of dissociation pathways and patterns is integral to this growth. SEQUEST is one of the most common proteomic software programs used to analyze MS/MS spectra of unknown peptides.
4.2 Construction of Nanospray Ionization Source

High performance liquid chromatography (HPLC) can be used to separate mixtures of proteins or peptides, and is therefore often used in conjunction with mass spectrometry in proteomic analysis. In this study peptides were separated from a tryptic digest using liquid chromatography and a nanospray ionization source. Peptides are separated based on the number and strength of interactions with solid stationary particles, as well as with a liquid mobile solvent. A pump is used to move the mobile phase and sample components through a column packed with these solid phase particles. The column can be packed with either polar or nonpolar packing material. When polar molecules are used, nonpolar sample molecules will elute from the column quickly since they do not interact with the stationary phase. Highly polar molecules will be retained in the column for a longer period of time. When nonpolar particles are used, the opposite is true, and it is referred to as reverse phase chromatography. Reverse phase HPLC was used in this study.

The polarity of the solvent also affects how quickly analyte molecules are eluted through the column. One of the major advantages of using liquid chromatography is the ability to employ a solvent gradient. The polarity of the solvent can be gradually or instantaneously changed to improve separation. In this study, two solvent mixtures were used, and the ratio was changed so that the polarity was first decreased and then returned to the initial level. The solvents used were 90:10 H₂O:MeOH with approximately 0.5% formic acid and 0.01% trifluoroacetic acid, and 98:2 MeOH:H₂O with approximately 0.5% formic acid and 0.01% trifluoroacetic acid. These were referred to as solvent A and solvent B respectively. The gradient was started right after making an injection using the
autosampler of the HPLC. In this case 20 µl injections were made. A sample gradient that was used for some preliminary studies is given below.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>% Solvent B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>32</td>
<td>35</td>
</tr>
<tr>
<td>37</td>
<td>45</td>
</tr>
<tr>
<td>37.1</td>
<td>95</td>
</tr>
<tr>
<td>42.1</td>
<td>95</td>
</tr>
<tr>
<td>42.2</td>
<td>5</td>
</tr>
<tr>
<td>75</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4.1 – Sample Time Course Gradient

This shows that solvent B is kept at 5% for the first 2 minutes, gradually increased to 35% over 30 minutes, gradually increased to 45% over 5 minutes, then instantaneously changed to 95% for another 5 minutes, and finally returned to 5% for the remainder of the 75 minute experiment. Decreasing the polarity in this manner helps to elute very nonpolar molecules that are held up in the column. The flow rate of the mobile phase was also altered over the course of the experiment. For the first 2 minutes a high flow of approximately 2 ml/min was used to speed up the travel time of the sample. The flow rate was then decreased to 0.5 ml/min. The goal was to quickly get the sample to the beginning of the column, and then slow down the flow to achieve maximum sensitivity and avoid overloading the column.

A nanospray ionization source was constructed for the LCQ-DECA ion trap instrument described in the previous section and interfaced to a Shimadzu HPLC. The
high sensitivity of nanospray was required for detection of small quantities of peptides from tryptic digestions. This ionization method differs from conventional electrospray in a few ways. It uses flow rates on the order of nanoliters per minute, as opposed to the micro or milliliter per minute flows used in electrospray. The superior sensitivity of nanospray is a result of smaller droplet size, which improves the efficiency of desolvation. Since smaller droplets have a higher surface to volume ratio, more analyte ions are able to be desorbed from each droplet. There is no need for nitrogen sheath gas because of this improved efficiency. Superior sensitivity also means that less sample is required for analysis, and makes nanospray a convenient method for studying clinical samples where limited amounts of material are available.

A splitter was installed onto the HPLC to achieve sufficiently low flow rates. As solvent passes through the splitter, 99.9% of the flow is diverted to a waste tube. Only the remaining 0.01% of the solvent is sent through the sample loop to carry the analyte molecules toward the chromatography column. The sample output of the HPLC was connected to the column via a microtee setup. The column consists of a silica capillary that has been packed with approximately 6-8 cm of Zorbax C18 packing material. Ions are sprayed through a five micrometer tip at the end of the column, and then sent through the heated capillary into the mass spectrometer. The column was held in a metal fitting on a stage that can be adjusted in all three dimensions. These adjustments were used to align the tip of the column with the opening of the heated capillary. A 2000V potential was applied via an alligator clip from the voltage supply of the mass spectrometer to a gold wire. This wire was held in the last remaining port of the microtee. This instrument will eventually be used to for the separation and analysis of tryptic digests.
4.3 – Proteomic Analysis of Tryptic Digests

Some preliminary studies have been performed on the protein Bovine Serum Albumin (BSA) using the HPLC-MS instrument described in section 4.1. This protein was selected because it is readily available and can be purchased as a predigested mixture. A 20µl injection of 1 pm/ml BSA in 48.5:48.5 H₂O:MeOH with 2% acetonitrile and 1% trifluoroacetic acid was made using the HPLC autosampler and a solvent gradient was programmed as described previously. The Xcalibur software program was used to acquire data using the data dependent scanning mode. A method was set up to automatically select the most abundant peaks in the mass spectrometer and perform CID on those ions. Solvent peaks and other impurities that are consistently seen in the ion trap were entered into a global ‘ignore list’ to avoid being mistaken for sample peaks.

The ultimate goal of this project is to examine unknown samples containing tryptic digests of bacterial and viral proteins. Data collected in X-calibur will then be entered into the SEQUEST software program to identify the proteins present in the unknown sample. Bacterial proteomics has recently become very popular because of the importance of bacteria in medicine. They are also simpler than eukaryotic species, and therefore have smaller proteins that can be more easily identified. In the analysis of bacteriophage, it is expected that many of the interesting viral proteins will be present in small amounts. The nanospray source was constructed for the purpose of providing adequate sensitivity for the detection of fragments from low abundance proteins. The steps of mass spectrometry based proteomics are shown in Figure 4.2.
Figure 4.2 – Steps of mass spectrometry based proteomics
References


