

Copyright
by
Gregory Terrell Blakney
2003

**The Dissertation Committee for Gregory Terrell Blakney
certifies that this is the approved version of the following dissertation:**

**Investigations of Biological Interactions by Hydrogen Deuterium
Exchange Fourier Transform Ion Cyclotron Mass Spectrometry:
Novel Methods, Automated Analysis and Data Reduction**

Committee :

David A. Laude, Supervisor

Jennifer S. Brodbelt

Jason B. Shear

David A. Vandebout

Kevin N. Dalby

**Investigations of Biological Interactions by Hydrogen Deuterium
Exchange Fourier Transform Ion Cyclotron Mass Spectrometry:
Novel Methods, Automated Analysis and Data Reduction**

by

Gregory Terrell Blakney, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2003

Dedication

To my loving family,
Laura, Zachary, Brett, and Alexandra

Acknowledgements

In addition to my family, there are many that I need to thank. I would like to thank Dave Laude for his help during my stay in Austin and beyond. He encouraged me throughout my graduate career and never lost patience.

While many others were responsible for guidance and meaningful discussions, Jared Drader, Chris Hendrickson and Jessica Robinson come to mind as important scientific influences throughout my graduate studies. Jared Drader helped me to better understand the subtleties of ICR. Chris Hendrickson continues to help me hone that understanding. Jessica Robinson did much of the initial H/D exchange work that is extended throughout this work.

There are many others that have influenced my work both in Austin and at NHMFL in Tallahassee. From Austin, Tim Histen, Chad Ostrander, and Richard Arkin were instrumental to my work and to my sanity. While in Tallahassee, Mark Emmett, John Quinn, Ryan Rodgers, Dan McIntosh and Alan Marshall are key ingredients in an exceptional workplace. Also while at NHMFL I had met some exceptional people that were a joy to work and play with: Stone Shi, Steve Beu, Bruce Wilcox, TuKiet Lam, Terri Quenzer, and Mike Freitas.

Finally, I must thank my family. I would like to thank my wife Laura who has withstood this long ride. I would like to thank my son Brett and daughter Alexandra for the love they bring to my life. I would like to thank my moms (Nancy and Barbara) and dads (Dick and Roger) for their support throughout years. I would like to thank my brothers (Wade, Earle, Clay and Bryan) for reminding me that there are some things more important than science

**Investigations of Biological Interactions by Hydrogen Deuterium
Exchange Fourier Transform Ion Cyclotron Mass Spectrometry:
Novel Methods, Automated Analysis and Data Reduction**

Publication No. _____

Gregory Terrell Blakney, Ph.D.
The University of Texas at Austin, 2003

Supervisor: David A. Laude

Hydrogen/deuterium (H/D) exchange is used to investigate biological interactions by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS). For the first time, a series of oligonucleotides of varying length are interrogated by negative mode gas phase H/D exchange. Data presented describes the reactivity of these oligonucleotides and correlates reactivity to functional moieties of the model compound. The results of the study are subjected to center of mass analysis, a technique that uses the high mass resolving power of FT-ICR MS to facilitate an isotope-counting / abundance-weighted algorithm to determine deuterium incorporation. A maximum entropy method produces the relative reaction rates for each of the model compounds. Curve fitting of rates reveal >90% correlation between the areas of individual rate curves and available hydrogens. Observed data is consistent with literature reported reaction mechanisms.

This work describes the successful implementation of the thorough high resolution analysis of spectra by Horn (THRASH) algorithm for the analysis of electron capture dissociation (ECD) spectra. Speed improvements in THRASH arise from optimized libraries and use of modern processors. The nonergodic nature of ECD alleviates deuterium scrambling and affords improved localization of exchange data. An ECD spectrum of ubiquitin is subjected to automated THRASH analysis resulting in 85% sequence coverage.

Automated data analysis is extended to the batch processing of entire high performance liquid chromatography (HPLC) FT-ICR MS experiments. Center of mass calculations are determined by combining the data from multiple scans to maximize signal. Data is output as a single spreadsheet. Results of automated and manual data processing are compared. Complete analysis of a H/D exchange study is completed in two hours, rather than the two months required for manual analysis. Results based on this method are demonstrated for biological interactions of HIV capsid and Nop5-fibrillarin complex.

New FT-ICR hardware increases data station throughput and reduces the amount of back exchange in liquid phase studies. The improved data station shows a three-fold improvement in scan speed and collects scans every 1.25 s. No significant loss of performance results from increased scan rates. Faster scan rates result in better chromatographic resolution and decreased spectral complexity.

Table of Contents

Chapter 1: Introduction to H/D Exchange: Techniques and Chapter Overviews	1
1.1 Hydrogen/Deuterium Exchange.....	4
1.2 Electrospray Ionization.....	5
1.3 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.....	6
1.4 Comparison of FT-ICRMS and NMR Techniques For Hydrogen/Deuterium Exchange Studies of Biomolecules.....	7
1.5 Chapter Overviews.....	9
CHAPTER 2: Experimental Method for the Investigation of Negative Ion Gas Phase H/D exchange of Oligonucleotides.....	11
CHAPTER 3: Analytical Method for the Characterization of Oligonucleotides by Negative Ion Gas Phase H/D Exchange ESI FT-ICR MS.....	12
CHAPTER 4: Data Reduction of Multiply Charged Ion Species Using an Automated Implementation of the THRASH Algorithm for the PC.....	14
CHAPTER 5: Data Reduction in the Determination of Solution Based H/D Exchange for the Determination of the Sites of Protein: Protein Interaction.....	15
CHAPTER 6: Improved ICR Data Station for HPLC FT-ICR MS of Liquid Phase H/D Exchange of Proteins.....	16
CHAPTER 7: Conclusions and Future Directions.....	17

References.....	19
-----------------	----

**Chapter 2: Experimental Method for the Determination of Negative Ion
Gas Phase H/D Exchange of Oligonucleotide 25**

2.1 Introduction.....	25
2.2 Methods.....	27
2.2.1 Sample Preparation.....	27
2.2.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.....	28
2.3 Results and Discussion.....	30
2.4 Conclusions.....	41
References.....	44

**Chapter 3: Analytical Method for the Characterization of Oligonucleotides by
Negative Ion Gas Phase H/D Exchange ESI FT-ICR MS 47**

3.1 Introduction.....	47
3.2 Methods.....	48
3.2.1 Sample Preparation.....	48
3.2.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.....	48
3.2.3 Data Analysis.....	49
3.3 Results and Discussion.....	49
3.4 Conclusions.....	65
References.....	70

Chapter 4: Data Reduction of Multiply Charged Ion Species Using an Automated Implementation of the THRASH Algorithm for the PC	71
4.1 Introduction.....	71
4.2 Methods.....	72
4.2.1 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry.....	73
4.2.2 Thorough High Resolution Analysis of Spectra by Horn.....	74
4.3 Results and Discussion.....	75
4.4 Conclusions.....	85
References.....	88
Chapter 5: Data Reduction in the Determination of Solution Based H/D Exchange for the Determination of the Sites of Protein: Protein Interaction	92
5.1 Introduction.....	92
5.2 Methods.....	94
5.2.1 Sample Preparation and H/D Exchange.....	94
5.2.2 Protein Digestion and HPLC Conditions.....	95
5.2.3 Electrospray Ionization FT-ICR MS.....	96
5.2.4 Proteolysis.....	96
5.3 Results and Discussion.....	97
5.3.1 Automated Batch Analysis Algorithm.....	97
5.3.2 Application of Automated Analysis.....	106
5.4 Conclusions.....	110

5.5 Acknowledgements.....	110
References.....	111
Chapter 6: Improved ICR Data Station for HPLC FT-ICR MS of Liquid Phase H/D Exchange of Proteins	113
6.1 Introduction.....	113
6.2 Methods.....	116
6.2.1 FT-ICR MS.....	116
6.2.2 HPLC Conditions.....	117
6.3 Results and Discussion.....	119
6.4 Conclusions.....	123
References.....	129
Chapter 7: Conclusions and Future Directions	133
7.1 Conclusions.....	133
7.2 Future Directions.....	134
References.....	142
Bibliography	143
Vita	144

Chapter 1: Introduction to H/D Exchange: Techniques and Chapter Overviews

With the recently awarded 2002 Nobel Prize for chemistry earned by Fenn, Tanaka, and Wüthrich for their combined work on the analysis of large biomolecules; chemistry, biochemistry and biology seem to be inextricably linked in the twenty first century. It is worthy of note that the award winning work of both Fenn [1, 2] and Tanaka [3] came specifically in the field of mass spectrometry related to biological molecules; and the award winning work of Wüthrich involved hydrogen exchange studies [4]. This pioneering work ushered in an era in the analysis of large biomolecules in which there is still much work to be done. One the areas of increased interest is the investigation of protein conformation as it relates to biological interactions. Techniques related to the investigation biological conformations or biological interactions vary from huge distributed computing projects investigating protein folding [5] to advanced hydrogen/deuterium (H/D) exchange assisted 2D nuclear magnetic resonance (NMR) measurements [6]. Regardless of the individual experiment, it is hydrogen exchange techniques that have provided much of the significant advances in the field of biological conformation / interaction.

The purpose of this dissertation was to demonstrate new and novel methodologies for hydrogen/deuterium exchange electrospray ionization (ESI) Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS). Methods establish FT-ICR MS as a valuable tool for the study of interactions of biologically active systems. Methods were demonstrated that make use of the explicit information that is afforded a high resolution FT-ICR instrument. Experimental methods for both gas phase and liquid phase hydrogen/deuterium exchange studies were investigated. Since the hydrogen/deuterium methods

proposed herein generate immense volumes of data, accuracy, speed and ease of use were considered crucial when developing the techniques for the analysis of hydrogen/deuterium exchange data. The analytical techniques described in this work can be equally applied to gas phase or liquid phase exchange techniques.

Experimental methods are presented that investigate negative gas phase exchange properties of oligonucleotides as well as liquid phase exchange techniques amenable to both positive and negative mode mass spectrometry of many biomolecules. These methods were demonstrated initially with a model system of poly-deoxycytidine (poly dCn). The results of the model study suggested a strong correlation between charge state and deuterium incorporation with deuterium incorporation varying from 10% for the $[M-3H]^{3-}$ to 92% for the $[M-5H]^{5-}$ of poly dC10. The observation of charge state dependence on the amount of deuterium incorporation was consistent with literature accounts for gas phase exchange of oligonucleotides [7].

The analytical techniques presented here are unique in that they were developed specifically for ESI FT-ICR MS data. In fact, the high mass resolution and mass accuracy of FT-ICR was central to both the accuracy and automation of the algorithms described herein. Techniques are demonstrated that improve the quality hydrogen/deuterium exchange data by removing the “human factor” from many of the repetitive tasks traditionally required of the user performing hydrogen/deuterium exchange. An additional benefit arises from further reduction in analysis time results from the constant improvement in the processing power of personal computers. The raw exchange profiles were then analyzed by a maximum entropy method (MEM) with the resultant MEM curves fitted under Microcal Origin (Microcal Software, Inc., Northampton, MA). The areas under the reaction curves were compared to the known exchangeable hydrogens for each species, the

correlation of the experimental abundances to the known abundances was shown a fit of 90% or greater for oligonucleotides poly dC9 through polydC14. This method of direct observation of exchangeable hydrogens was limited to highly charged species and was consistent with literature reports of a charge mediated mechanism for negative ion hydrogen / deuterium exchange [7].

Additional work describes the successful implementation of the thorough high resolution analysis of spectra by Horn (THRASH) [8] algorithm for the automated analysis of multiply charged species. The proposed use of the THRASH algorithm is the analysis electron capture dissociation (ECD) spectra that would arise from the MS/MS of liquid phase H/D exchange species. Speed improvements in THRASH arise from optimized libraries and use of modern processors. The nonergodic nature of ECD alleviates deuterium scrambling and affords improved localization of exchange data. An ECD spectrum of ubiquitin is subjected to automated THRASH analysis resulting in 85% sequence coverage.

Automated data analysis is extended to the batch processing of entire high performance liquid chromatography (HPLC) FT-ICR MS experiments. Center of mass calculations are determined by combining the data from multiple scans to maximize signal. The user may choose to analyze individual or multiple targets for batch analysis. The selection of multiple targets per session minimizes total analysis time since selection and sorting of the many files involved in the analysis requires a significant portion (40-50 %) of the analysis time. Data from each target analysis output as a single spreadsheet. Results of automated and manual data processing are compared. Complete analysis of the raw data from a H/D exchange study is completed in two hours, rather than the two months required for manual analysis. Results based on this method are demonstrated for biological interactions of HIV capsid and Nop5-fibrillarin complex.

New FT-ICR hardware increases data station throughput and reduces the amount of back exchange in liquid phase studies. The improved data station shows a three-fold improvement in scan speed and collects scans every 1.25 s. No significant loss of performance results from increased scan rates. Further reduction in scan time was possible but resulted in decreased performance for the standard solution used for this study. The net result of faster scan rates result in better chromatographic resolution and decreased spectral complexity. Future work will provide similar performance gains for each component in the FT-ICR data station with the end goal of this work being the ability to collect data-dependent HPLC MS/MS H/D exchange FT-ICR MS data. The enhanced hardware is needed to perform such experiment on a time scale that is compatible with HPLC (<2 sec per scan) while back exchange is minimized.

Background

1.1 Hydrogen/Deuterium Exchange

The study of hydrogen/deuterium exchange effects is certainly not limited to the study of biomolecules. Literature cites a number of works in the 1930s in which hydrogen/deuterium exchange studies were performed. Some of these early experiments involved an exchange study of acetate determined by density [9] and mechanistic studies of the deuterated nitramide [10]. Dibeler and Taylor first combined mass spectrometry with a hydrogen/deuterium reaction in 1948. This work was limited to the study of n-butene [11].

Hydrogen/deuterium exchange of peptides was first reported in 1955, when Hvidt and Linderstrom-Lang released no less than four papers on

hydrogen/deuterium exchange studies on peptides [12-15]. Throughout the 1960s and into the early 1970s, hydrogen/deuterium exchange was used in conjunction with a gradient tube apparatus [16, 17] or early NMR instruments [18] to elucidate the exchange rates of selected amide hydrogens [19, 20].

With the development of the high resolution NMR in the early 1980s, hydrogen/deuterium studies took on new life as improvements in resolution permitted each amide bond in small peptides to be resolved [18, 21-23]. This allowed one to assign an exchange rate to individual amino acids within a peptide. Further work on NMR coupled hydrogen/deuterium exchange extended the technique to larger peptides and more advanced conformational studies [6, 22, 24].

1.2 Electrospray Ionization

Just as early hydrogen/deuterium exchange studies languished until NMR spectrometers achieved the resolution required to define exchange rates of individual amine sites, hydrogen/deuterium studies with mass spectrometry faced a similar hurdle: an ionization technique soft enough to safely deliver biomolecules to the gas phase ions. As referred to previously, two ionization techniques have been developed that address the soft ionization of biomolecules. Matrix assisted laser desorption ionization (MALDI) and electrospray ionization are quite of producing intact analyte ions from various biomolecules. While MALDI has been proven effective for large biomolecules [3, 25], electrospray ionization is a much better ion source for FT-ICR MS instruments [26-28].

For a brief history of electrospray, one must first look to work done in 1917 by Zeleny [29]. This work describes factors that affect current from a discharging point. Additional work by Dole in 1968 [30], involved the detection of macroions of

polystyrene sampled from a molecular beam with a Faraday cage. Finally, work by Fenn brought the modern electrospray ionization source to fruition in 1984 [1, 31, 32].

Electrospray ionization is well matched for FT-ICR MS. The production of multiply charged ions acts to extend the somewhat limited mass range of the FT-ICR instrument. The high resolution of the FT-ICR allows for the facile determination of the charge state even for highly charged species. Since the electrospray source operates in the solution phase, it provides the link between HPLC separations and the FT-ICR mass spectrometer. Electrospray ionization performs efficiently as an ion source with detection limits reported in the attomole level [26, 33]. Electrospray also has the capability of creating both positive and negative ions.

1.3 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Comisarow and Marshall developed FT-ICR MS in the mid 1970s. FT-ICR provides unrivaled mass resolution and mass accuracy [34, 35]. FT-ICR instrument can suffer from limited mass range but with introduction of multiply charged ions from modern electrospray ionization sources, this problem can be largely overcome [28].

FT-ICR has distinct advantages over many other types of mass spectrometers for performing hydrogen/deuterium studies. The primary reason to choose an FT-ICR instrument is the unsurpassed mass resolution and mass accuracy of the technique. The resolution of the FT-ICR affords the ability to unambiguously determine the charge state of the analyte even at relatively high charge states when performing HPLC FT-ICR exchange studies. Automated charge state determination

provides an important safeguard against incorrect assigned exchanged data [8, 36-38]. Charge state determination will become increasingly important in data-dependent FT-ICR experiments that are introduced in this work. One last advantage that the resolution of an FT-ICR MS permits is the ability to reject interferences that are overlapped with the isotope envelope of interest in both manual and automated analysis.

FT-ICR spectrometers also have the capability to maintain a trapped ion population for extended time periods, complementing the extended reaction periods required for gas phase H/D exchange experiments [39, 40]. FT-ICR MS also allows precise control over the ion population in the ICR cell; this allows selective ejection of ion populations that interfere with exchange experiments. Such control is key to minimizing interference when performing gas phase exchange experiments.

1.4 Comparison of FT-ICR MS and NMR Techniques For Hydrogen/Deuterium Exchange Studies of Biomolecules

NMR has a proven track record for hydrogen/deuterium exchange monitoring. When comparing an NMR technique to a mass spectrometry technique, it is important to note that the current NMR technique is considered to have the better exchange resolution than current mass spectrometric technique. This improved resolution is the result of the NMR measuring the deuterium content of residues individually while mass spectrometry techniques rely upon the specificity of the chosen digestion enzyme to provide the given analyte sequence. In the general case, a protease like pepsin will generate fragments of about 10 to 15 base pairs [41]. This would correspond to an exchange resolution of about 10 to 15 residues as a typical result for MS.

The advantages of a mass spectrometric technique for hydrogen/ deuterium exchange when compared with an NMR technique include decreased sample consumption, a lower detection limit, reduced dead time / less back exchange, and extended mass range [41]. The amount of sample required for a NMR study ranges from 5 to 50 μmol and the same mass spectrometric study would consume 2 to 5 nmol of sample. The millimolar concentrations required for NMR can lead to aggregation for analyte where the micromolar concentrations of mass spectrometry might not. The analysis time for 2D NMR experiment can last several hours during which time many of the deuterons have back exchanged and go undetected. The analysis time for HPLC-MS is typically 10 to 15 minutes, which allows the determination of all but the most extreme exchanges. However, it is the extended mass range that can be the most important reason chosen to use the mass spectrometric method.

With the noted exception of the GroEL-GroES work published this summer [6], NMR studies are limited too analytes with molecular weights of 50 kDa. Mass spectrometric techniques have been applied to α -crystallin analytes with molecular weights near 1000 kDa [42]. The GroEL-GroES (900 kDa) is to be considered a special case since the protein was grown in media where the isotope contributions of C, N and H were strictly controlled whether they be enriched or depleted. In contrast the α -crystallin analytes were extracted from a biological sample with no special growth preparations. Finally, the mass spectrometric technique benefits from the digestion of the analyte into manageable fragments. By fragmenting the sample prior to analysis, increasing the molecular weight of the analyte increases the number of fragments in the analyzer at any one time but the complexity of fragments remains relatively constant for a given enzyme system.

High Field FT-ICR has additional advantages over NMR and alternate MS techniques. For small analytes, gas phase exchange removes the possibility of significant back exchange since there are very few collisions at the typical base pressure associated with FT-ICR instruments. FT-ICR has the resolution to distinguish thousands of individual ion packets in each scan. The multiplex advantage inherent to ICR detection allows consistent scan times regardless of the number of analytes in the ICR cell. Electron capture dissociation (ECD) may provide the most important advantage to FT-ICR exchange experiments. Initial studies of ECD suggest that the dissociation is nonergotic; that is, the dissociation occurs on a timescale that precludes the energy from randomizing throughout the analyte. If ECD can be proven to dissociate the analyte without causing the deuterium scrambling [43], it would be possible to further localize the deuterium incorporation. The net effect might yield the similar exchange resolution of NMR studies without the additional sample handling.

1.5 Chapter Overviews

The purpose of the work in this dissertation is to provide the tools both analytical and experimental to examine the conformational aspects of biological systems. These tools are designed to make hydrogen/deuterium exchange more accurate, user friendly and flexible. In chapter 2 a general protocol is described for gas phase exchange experiments. A series of poly-deoxycytidine analytes are selected as the model system to demonstrate the protocol. In chapter 3 analytical techniques are described that automate data analysis. This is required because even the relatively small H/D exchange project described in chapter 2 generates 1260 data files. In chapter 4, automated deconvolution of analytes to zero charge is demonstrated for multiply charged ion packets. Automated analysis of electron capture dissociation data is presented. Chapter 5 describes an algorithm that reduces

the total data analysis time for HPLC FT-ICR liquid phase H/D exchange data from a process taking months to a process accomplished in a few days. Chapter 6 describes an improved FT-ICR data station with the specific purpose of improve HPLC FT-ICR performance. The final chapter (7) reviews the preceding chapters focusing on the relationship between current work and future experiments. Specifically, information about the further improvement to the ICR data station will be discussed that will introduce the ability to produce data-dependent fragmentation of the exchanged analytes. Preliminary data-dependent work is presented.

Chapter 2: Experimental Method for the Investigation of Negative Ion Gas Phase H/D Exchange of Oligonucleotides

In the studies presented, a systematic investigation of a series of oligonucleotides of varying lengths is performed. Poly-deoxycytidine (poly dC) was chosen as the best candidate for the study. The series of oligonucleotides from poly dC₂ through poly dC₁₅ were obtained and prepared for analysis. The gas phase structure of each poly dC candidate is examined through the use of gas phase H/D exchange negative ion micro-ESI FT-ICR MS. A protocol is developed that precludes chemical interference by removing all ions but those of the charge state of interest from the ICR cell prior to the H/D exchange reaction. Exchange curves are presented that demonstrate the H/D exchange reaction as a function of time and features and terminology are discussed that lay the groundwork for more advanced topics.

A simple deuterium calculation described as the last isotope method introduces the first quantitative measure of H/D exchange described in this work. Deuterium calculations based on the last isotope method can be performed quickly and as such are often described in the literature [7, 44-46]. Although more descriptive techniques for the monitoring of H/D exchange will be introduced in later chapters, last isotope measurements are provided that suggest marked charge state dependence upon H/D exchange of the poly dC analogues. Finally, the observation of multiple conformers within the same charge state of poly dC₁₄ suggest that even the poly dC system is not devoid of higher order influences and therefore does not perform as an ideal system.

Chapter 3. Analytical Method for the Characterization of Oligonucleotides by Negative Ion Gas Phase H/D Exchange ESI FT-ICR MS

This new work demonstrates methods for the analysis of H/D exchange studies created by high-field FT-ICR MS instruments. The primary goal of this work is to provide data-mining techniques for the user that produce more precise exchange information in a rapid and facile manner. The primary tool that is provided is the automated calculation of center of mass data. Center of mass calculations use a combination of isotope counting and abundance weighting to provide an exchanged value that is relatively immune to problems associated with the last isotope method. Potentially detrimental effects of user bias, isotope envelope spread and variable signal to noise are demonstrated. The impact of these effects on the accuracy of the last isotope method and the center of mass method is discussed. A direct comparison of the last isotope and center of mass methods is performed demonstrating that incorporation results calculated by the last isotope method are generally over-estimates the calculated deuterium incorporation when compared to those calculated by center of mass methods.

While the calculation of center of mass is relatively straightforward, manual calculation for each occurrence would prove tedious. In addition to the tedium, the probably of a data entry error as the user performed hundreds of such calculations would become substantial. Automated calculation of center of mass has become the cornerstone of H/D exchange studies within the Modular ICR Data Acquisition and analysis Station (MIDAS) [47]. The resolution of FT-ICR allows baseline separation of the isotope of the analyte; this allows for automated peak picking and from the peak list the center of mass is easily calculated. HD Helper is an auxiliary feature of MIDAS that allows one to rapidly remove any disparate peaks and recalculate the center of mass. Data generated by center of mass is subjected to

maximum entropy method (MEM) analysis [48]. The results of the MEM analysis show a greater than 90% correlation to structural moieties located in the analytes when those analytes contain sufficient charge. The requirement of high charge on the analyte for diagnostic levels of exchange to occur suggest that such gas phase H/D exchange studies are not generally amenable to the elucidation of biological interactions.

Chapter 4. Data Reduction of Multiply Charged Ion Species Using an Automated Implementation of the THRASH Algorithm for the PC

This work describes the successful implementation of Thorough High Resolution Analysis of Spectra by Horn (THRASH) algorithm [8] for the analysis of electron capture dissociation (ECD) spectra [43, 49, 50]. Automated analysis is a prerequisite for H/D exchange by data-dependent HPLC FT-ICR MS methods. Data-dependent MS/MS promises to extend the resolution of the deuterium exchange beyond current mass spectrometric limits. The key to the extended resolution relies upon the successful application of electron capture dissociation (ECD) to liquid phase H/D exchange FT-ICR MS methods. The nonergodic nature of ECD alleviates deuterium scrambling normally associated with gas phase dissociation events. In the absence of deuterium scrambling, ion dissociation event can be used to localize the deuterium exchange. Unfortunately, the complexity of ECD spectra leads to increased analysis time.

This work describes the successful implementation of Horn's THRASH algorithm on a PC platform using the C language. Large gains in processing speed arise from optimized libraries and bandwidth increases of modern processors. An ECD spectrum of ubiquitin is subjected to automated THRASH analysis resulting in 85% sequence coverage. The use of THRASH will improve the efficiency of the analysis of primary proteolytic fragments generated by liquid phase exchange HPLC FT-ICR MS experiments. These primary fragments will guide the application of the automated batch H/D exchange processing discussed in Chapter 5. A minimized version of the THRASH algorithm is planned to operate as an analysis module for data-dependent ICR discussed in chapter 7.

Chapter 5. Data Reduction in the Determination of Solution Based H/D Exchange for the Determination of the Sites of Protein: Protein Interaction

The work described involves the development of an automated analysis protocol that can effectively reduce the total analysis time while improving the accuracy for a liquid phase H/D exchange HPLC FT-ICR MS studies. This analysis protocol reduces the total analysis time for a complete H/D exchange study by liquid- phase HPLC FT-ICR MS from a period of months to an analysis period of about a week. The accuracy of the results is also increased since the algorithm does not require the user to track the analysis through a series of data transfer and collation stages.

A detailed description of the algorithm is provided and the advantages of the iterative nature of the algorithm are explained. Automated center of mass calculations are determined by combining the data from multiple HPLC scans to maximize signal. Data is output as a single spreadsheet. Results of automated and manual data processing are compared. Results based on this method are demonstrated for biological interactions of HIV-1 CA capsid protein and fibrillarin – Nop5 complex.

Chapter 6. Improved ICR Data Station for HPLC FT-ICR MS of Liquid Phase H/D Exchange of Proteins

Current FT-ICR data stations do not operate on the time scale that is best suited for the modern nanoscale HPLC apparatus. Until recently, there was little emphasis placed on improving the data station because other factors such as ion accumulation also required time scales on the order of 5 to 10 s to generate adequate signal. A recent advance has increased the efficiency of external accumulation that has made it possible to decrease the ion accumulation period to 1 s or less [51]. The development of this type of external source served to underscore the deficiencies of current data stations.

The Modular ICR Data Acquisition and analysis System (MIDAS) was developed at the National High Magnetic Field Laboratory (NHMFL) in order to provide a flexible solution to commercial data station shortcomings. The primary advantage of the MIDAS is the use of off the shelf components and industry standard bus systems to create a data station that has the ability to be immediately upgraded as new products reach the market. Programmer and scientific end users are co-workers or in some cases the same person. This interaction allows much better turn around when applied to software issues. New hardware implementation for improved data throughput was already underway when the improved external ion accumulation design was being developed. The replacement of an older technology VXI digitizer with a newer PXI digitizer led to a >40 fold improvement in data transfer speed. This allowed HPLC FT-ICR scans to be acquired at rates of greater than 1 scan per second whereas the older data station required 3.5s. HPLC FT-ICR MS results were demonstrated for a 3-peptide mixture and later a tryptic digest from a salt buffer solution.

Chapter 7: Conclusions and Future Directions

The use of H/D exchange techniques remains one of the most powerful tools to study the interactions of large biomolecules. Liquid H/D exchange followed by HPLC FT-ICR analysis has several advantages over similar NMR techniques. FT-ICR techniques require less sample and sample preparation, and are capable of analyzing a wider variety of samples. Work presented in this dissertation provides for the rapid and accurate analysis of large time-dependent exchange studies. An automated mass determination followed by automated batch processing eliminates months of low-level user analysis and delivers to the user reports appropriate for immediate high-level evaluation. In addition to improved data analysis techniques, improved FT-ICR data acquisition was described that uses improved hardware to generate superior HPLC FT-ICR MS results. The use of electron capture dissociation (ECD) was suggested to improve the H/D exchange resolution of the liquid exchange protocol; however, the utility of ECD must still be verified. Results of current H/D exchange studies have been to categorize the interactions of HIV-1 capsid proteins and Nop5 – fibrillarin complex.

Future work will involve the addition of hardware to increase the speed of the HPLC FT-ICR as well as the development of data-dependent data acquisition. Data-dependent isolation of a standard peptide mixture is demonstrated for MIDAS data station using currently implemented hardware. While the ICR scan time was relatively slow at ~ 13 seconds per scan, the data does show isolation using the front-end quadrupole on the 9.4 T FT-ICR system at NHMFL.

Studies to ascertain the effect of ECD on deuterium scrambling are currently underway. Given that ECD can be proven to be truly nonergodic with respect to H/D exchange data, then data-dependent selection and fragmentation should yield

increased localization of exchange sites. Additional studies are planned to characterize the response of ECD on the 9.4 T FT-ICR at NHMFL for various peptides in an attempt to determine the best conditions to utilize for automated fragmentation.

References

1. Fenn, J.B., et al., *Electrospray Ionization for Mass Spectrometry of Large Biomolecules*. Science, 1989. **246**: p. 64-71.
2. Fenn, J.B., et al., *Electrospray*. Mass Spectrom. Rev., 1990. **9**: p. 37-70.
3. Tanaka, K., et al., *Protein and Polymer Analyses up to m/z 100000 by Laser Ionization Time-of-Flight Mass Spectrometry*. Rapid Commun. Mass Spectrom., 1988. **2**: p. 151-153.
4. Roder, R., G. Wagner, and K. Wuthrich, *Amide Proton Exchange*. Biochemistry, 1985. **24**: p. 7396-7407.
5. Larson, S.M., et al., *Folding@Home and Genome@Home: Using distributed computing to tackle previously intractable problems in computational biology.*, in *Computational Genomics*. 2002, Horizon Scientific Press. p. in Press.
6. Flaux, J., et al., *NMR Analysis of a 900K GroEl-GroES complex*. Nature, 2002. **418**: p. 207-211.
7. Robinson, J.M., et al., *Hydrogen/Deuterium Exchange of Nucleotides in the Gas Phase*. Anal. Chem., 1998. **70**: p. 3566-3571.
8. Horn, D.M., R.A. Zubarev, and F.W. McLafferty, *Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules*. J. Am. Soc. Mass Spectrom., 2000. **11**: p. 320-332.
9. Liotta, S. and V.K. LaMer, *Hydrogen-deuterium exchange in acetate solution*. J. Am. Chem. Soc., 1938. **59**: p. 946.
10. LaMer, V.K. and s. Hochberg, *Hydrogen and deuterium exchange between nitramide and water*. J. Am. Chem. Soc., 1939. **61**: p. 2552-2553.
11. Dibeler, V.H. and T.I. Taylor, *Mass-spectrometric and infrared study of rates of deuterium exchange, isomerization, and hydrogenation of the n-butenes*. J. Chem. Phys., 1948. **16**: p. 1008-1009.

12. Hvidt, A., *Deuterium exchange between ribonuclease and water*. Biochim. et.Biophys. Acta, 1955. **18**: p. 306-308.
13. Hvidt, A. and K. Linderstrom-Lang, *Kinetics of deuterium exchange of insulin with D₂O. An Amendment*. Biochim. et.Biophys. Acta, 1955. **16**: p. 168-169.
14. Linderstrom-Lang, K., *Deuterium exchange between peptides and water*. Chem. Soc., 1955. **2**: p. 1-20.
15. Linderstrom-Lang, K., *The pH-dependence of the deuterium exchange of insulin*. Biochim. et.Biophys. Acta, 1955. **18**: p. 308.
16. Levi, H. and E. Zeuthen, *Microweighing in the gradient ube*. Compt.rend.lab. Carlsberg, Ser.chim, 1946. **25**: p. 273-287.
17. Hvidt, A., et al., *Exchange of deuterium and oxygen-18 between water and other substances. I. Methods*. Compt.rend.lab. Carlsberg, Ser.chim, 1955. **29**: p. 129-57.
18. Molday, R.S., S.W. Englander, and R.G. Kallen, *Primary Structure Effects on Peptide Hydrogen Exchange*. Biochemistry, 1972. **11**(2): p. 150-158.
19. Hvidt, A. and S.O. Nielsen, *Adv. Protein Chem.*, 1966. **21**: p. 287-286.
20. Englander, S.W., N.W. Downer, and H. Teitelbaum, *Hydrogen exchange*. Annu. Rev. Biochem., 1972. **41**: p. 903-924.
21. Pan, Y. and M.S. Briggs, *Hydrogen Exchange in Native and Alcohol Forms of Ubiquitin*. Biochemistry, 1992. **31**: p. 11405-11412.
22. Englander, S.W. and L. Mayne, *Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR*. Annu. Rev. Biophys. Biomol. Struct., 1992. **21**: p. 243-265.

23. Englander, S.W. and N.R. Kallenbach, *Hydrogen exchange and structural dynamics of proteins and nucleic acids*. Quart. Rev. Biophys., 1984. **16**: p. 521-655.
24. Li, R. and C. Woodward, *The hydrogen exchange core and protein folding*. Protein Sci., 1999. **8**: p. 1571-1591.
25. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Anal. Chem., 1988. **60**(20): p. 2299-2301.
26. Emmett, M.R. and R.M. Caprioli, *Micro-Electrospray Mass Spectrometry: Ultra-High-Sensitivity Analysis of Peptides and Proteins*. J. Am. Soc. Mass Spectrom., 1994. **5**: p. 605-613.
27. Valaskovic, G.A., N.L. Kelleher, and F.W. McLafferty, *Attomole Protein Characterization by Capillary Electrophoresis-Mass Spectrometry*. Science, 1996. **273**: p. 1199-1202.
28. Laude, D.A., E. Stevenson, and J.M. Robinson, *Electrospray ionization/Fourier transform ion cyclotron resonance mass spectrometry*, in *Electrospray Ionization Mass Spectrometry*, R.B. Cole, Editor. 1997, John Wiley & Sons, Inc.: New York. p. 291-319.
29. Zeleny, J., *Electric Discharge from Points*. Phys. Rev., 1917. **9**: p. 562-563.
30. Dole, M., et al., *Molecular beams of macroions*. J. Chem. Phys., 1968. **49**(5): p. 2240-2249.
31. Yamashita, M. and J.B. Fenn, *Electrospray Ion Source. Another Variation on the Free-Jet Theme*. J. Chem. Phys., 1984. **88**: p. 4451-4459.
32. Whitehouse, C.M., et al., *Electrospray*. Anal. Chem., 1985. **57**: p. 675-679.
33. Valaskovic, G.A., et al., *Microbore Capillary Electrophoresis ESI FT-ICR MS*. Anal. Chem., 1995. **67**: p. 3802.

34. Marshall, A.G. and P.B. Grosshans, *Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: The Teenage Years*. Anal. Chem., 1991. **63**: p. 215A-229A.
35. Comisarow, M.B. and A.G. Marshall, *Fourier Transform Ion Cyclotron Resonance Spectroscopy*. Chem. Phys. Lett., 1974. **25**: p. 282-283.
36. Covey, T.R., et al., *The determination of protein, oligonucleotide and peptide molecular weights by ion-spray mass spectrometry*. Rapid Commun. Mass Spectrom., 1988. **2**(11): p. 249-256.
37. Mann, M., C.K. Meng, and J.B. Fenn, *Interpreting Mass Spectra of Multiply Charged Ions*. Anal. Chem., 1989. **61**: p. 1702-1708.
38. Zhang, Z. and A.G. Marshall, *A Universal Algorithm for Fast and Automated Charge State Deconvolution of Electrospray Mass-to-Charge Ratio Spectra*. J. Am. Soc. Mass Spectrom., 1998. **9**: p. 225-233.
39. Wood, T.D., et al., *Gas-phase folding and unfolding of cytochrome c cations*. Proc. Nat. Acad. Sci. U.S.A., 1995. **92**(7): p. 2451-2454.
40. Freitas, M.A., et al., *Gas-Phase Bovine Ubiquitin Cation Charge State Conformations Resolved by Gas-Phase Hydrogen/Deuterium Exchange Rate and Extent*. Int. J. Mass Spectrom, 1999. **185/186/187**: p. 565-575.
41. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the noncovalent structure of proteins by amide hydrogen exchange and mass spectrometry*. J. Mass Spectrom., 1997. **32**(2): p. 135-146.
42. Liu, Y. and D.L. Smith, *Probing High Order Structure of Proteins by Fast-Atom Bombardment Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 1994. **5**: p. 19-28.

43. Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty, *Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process*. J. Am. Chem. Soc., 1998. **120**: p. 3265-3266.
44. Freitas, M.A., et al., *High Field FT-ICR Mass Spectrometry for Simultaneous Trapping and Gas-Phase H/D Exchange of Peptide Ions*. J. Am. Soc. Mass Spectrom., 1998. **9**: p. 1012-1019.
45. Freitas, M.A., et al., *Gas Phase RNA and DNA Ions. 1. H/D Exchange of the [M - H]⁻ Anions of Nucleoside 5'-Monophosphates (GMP, dGMP, AMP, dAMP, CMP, dCMP, UMP, dTMP), Ribose-5-monophosphate (R5P), and 2-deoxyribose-5-monophosphate (dR5P) with D₂O and D₂S*. J. Am. Chem. Soc., 1998. **120**: p. 10187-10193.
46. Freitas, M.A. and A.G. Marshall, *Gas Phase RNA and DNA Ions 2. Conformational Dependence of the Gas-Phase H/D Exchange of Nucleotide-5'-Monophosphates*. J. Am. Soc. Mass Spectrom., 2001. **12**: p. 780-785.
47. Senko, M.W., et al., *A High-Performance Modular Data System for FT-ICR Mass Spectrometry*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1839-1844.
48. Zhang, Z., et al., *Human Recombinant [C22A] FK506-Binding Protein Amide Hydrogen Exchange Rates from Mass Spectrometry Match and Extend Those from NMR*. Protein Sci., 1997. **6**: p. 2203-2217.
49. Zubarev, R.A., et al., *Electron Capture Dissociation of Gaseous Multiply-charged Proteins is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity*. J. Am. Chem. Soc., 1999. **121**: p. 2857-2862.
50. Zubarev, R.A., et al., *Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations*. Anal. Chem., 2000. **72**: p. 563-573.
51. Wilcox, B.E., C.L. Hendrickson, and A.G. Marshall, *Improved Ion Extraction from a Linear Octopole Ion Trap: SIMION Analysis and*

Experimental Demonstration. J. Am. Soc. Mass Spectrom., 2002. **13**: p. 1304-1312.

Chapter 2 Experimental Method for the Determination of Negative Ion Gas Phase H/D Exchange of Oligonucleotides.

2.1 Introduction

The study of nucleic acids and nucleotides has had almost universal appeal since the discovery of the biological function of DNA by Watson and Crick in 1953 [1, 2]. Although this discovery offered great insight into the role that DNA played in cellular replication and described a general mechanism for such actions, the details of the replication mechanism were not yet understood. Additionally, the process by which the information contained in the DNA was harnessed to produce functional proteins became an area of huge interest. In the fifty years since the discovery of DNA, much has been learned about DNA replication, transcription and translation. Most of the enzymes and cofactors important to gene expression and cellular replication are described in college textbooks [3] but what remains to be answered are the details. While today's details such as molecular recognition and interaction may only be subsets of the original details arising from the discoveries of Watson and Crick, science must find new and imaginative ways to probe for the answers of these new questions of today to better understand tomorrow.

While many techniques have been employed to study this phenomenon, nuclear magnetic resonance (NMR) and mass spectrometry (MS) are the most popular. NMR studies can provide detailed exchange information provided that all of the amines can be resolved [4]. Mass spectrometry has advantages in speed, sample quantity required, and mass range [5-9]. In addition to liquid phase work, the applications of H/D exchange studies have been extended to the study of gas phase conformations of biomolecules through the development of gas phase H/D exchange mass spectrometric techniques [10-22].

While the initial gas phase exchange study of Winger et al. was performed in a modified triple quadrupole [10], much of the rest of the gas phase exchange work has been performed using Fourier transform ion cyclotron mass spectrometry (FT-ICR MS) [11-19]. The volume of work associated with FT-ICR MS techniques demonstrates the unique capabilities of FT-ICR MS. FT-ICR MS has unparalleled resolution, extended ion hold times and extensive ion isolation capabilities. The resolution of FT-ICR allows one to readily discern all of the isotopes associated with the analyte and discard spurious peaks that override the spectra. The extended hold times of FT-ICR are particularly well suited for gas phase exchange experiments. Gas phase exchange FT-ICR studies lasting up to an hour have been described in the literature [13]. Since gas phase exchange is known to be charge state dependent and conditions used for gas phase exchange lead to charge stripping, it is necessary to isolate individual charge states to eliminate contamination. Isolation is also used to remove salt adducts from the analyte that might confuse deuterium incorporation results.

One of the more recent tools used to define the role of oligonucleotides in cellular expression is gas phase hydrogen/deuterium exchange in conjunction with negative mode electrospray ionization (ESI) FT-ICR MS. In negative mode ESI, the phosphate backbone of oligonucleotides provides a charge site that is independent of the nucleobase added. Uniform charge distribution is important in negative mode gas phase exchange since the exchange process is reported to occur through a charge-catalyzed process [16]. Initial studies of negative mode gas phase H/D exchange oligonucleotides focused on monomers and cyclic monomers. Factors such as reagent gas and localization of the charge were studied in detail [16, 17]. Work on single stranded oligonucleotides was performed but the selection of randomized oligonucleotide sequences proved analytically challenging [23].

In this new work, a model system of oligonucleotides is examined by negative mode H/D exchange FT-ICR MS. An experimental protocol is described that eliminates spectral interference from charge stripping, salt adducts, and spurious noise peaks. The model system chosen for this work is the ammonium salts of polydeoxycytidine (poly dC_n) where n represents 2 to 15 cytidine units. The model system was chosen to reduce salt adduction and eliminate additional chemical interferences as described later. Results for total deuterium incorporation are described and the charge dependence on H/D exchange is examined. Results indicating the presence of multiple conformations are presented and discussion of these multiple conformers in the model system is discussed.

2.2 Methods

2.2.1 Sample Preparation

Oligonucleotide samples were obtained as lyophilized ammonium salt derivatives from TriLink Biotechnologies (San Diego, CA). Ammonium salt derivative preparation was chosen to minimize mass spectral interferences that are common with sodium salt oligonucleotide derivatives. The model system for our experiment would be the system that provided the least interferences while providing the simplest system to study. Based on these criteria poly deoxycytidine (poly dC) was chosen as the best choice for our model system. Poly-deoxythymidine (poly dT) ion has been shown to cyclize in the gas phase [24]. The additional structure of the cycle will complicate the analysis of this data. Poly-deoxyadenosine is also restricted since any studies on duplexes would include poly dT. Poly dC only has a single class of exchangeable hydrogen associated with the nucleobase whereas poly-deoxyguanosine has two classes of exchangeable

hydrogens, which again serves to potentially complicate the study. Oligonucleotides were used without further purification. Stock solutions were prepared at 1 mg / ml in 1:1 (v: v) isopropanol (IPA): water. Final solutions were prepared at 10-100 μg /ml in 1:1 IPA: water with 40 μM ammonium acetate added.

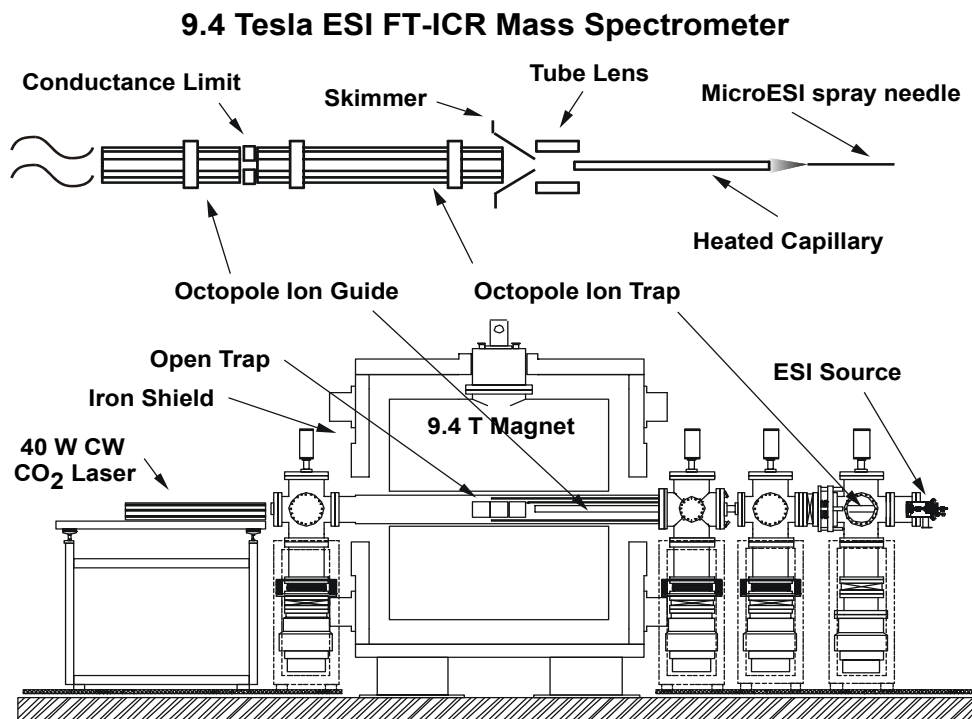
2.2.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

All experiments were performed using a custom built 9.4 T external electrospray source FT-ICR MS (National High Magnetic Field Laboratory, Tallahassee, FL) [25]. A schematic of this instrument as configured during these experiment is provided in Figure 2.1. This instrument employed a Chait type electrospray source with a 25 μm i.d. fused silica spray needle. Ion desolvation was accomplished using a resistively heated capillary.

Ions exiting the capillary were focused into a 0.8 mm i.d. octopole rod set with the use of an electrostatic tube lens. Ions were then accumulated in a 60 cm long octopole by biasing the entrance and exit conductance limits to the octopole to a negative potential [26]. After an appropriate accumulation period of 2 s to 60 s, the potential on the exit conductance limit was dropped to effect transfer of the ion through a second 2 m long octopole and into the analyzer cell of the FT-ICR instrument. [25]. The analyzer cell of this 9.4 T FT-ICR instrument is a 4 in diameter open cylindrical Penning ion trap.

Specific conditions that produced optimum sensitivity for each oligonucleotide were determined experimentally prior to collecting each set of hydrogen / deuterium exchange (H/D exchange) spectra. Samples were directly infused into the electrospray source at a flow rate of 300 to 500 nl / min controlled by a Harvard Apparatus syringe pump (South Natick, MA). Needle potentials of –

Figure 2.1. Schematic diagram of the 9.4 T FT-ICR MS Instrument at National High Magnetic Field Laboratory in Tallahassee, FL



2500 V to -2900 V were used to produce efficient analyte ionization. Prior to the H/D exchange period, each charge state of each analyte was isolated using a combination of chirp waveforms and a stored waveform inverse Fourier transform (SWIFT) waveform [27]. Isolation of a specific charge state insured that no spectral contamination arose from the H/D exchange of higher charge states of the analyte. Such contamination would arise from charge stripping of high charge states species during the H/D exchange event.

H/D exchange was performed in the analyzer cell of the FT-ICR instrument. A combination leak valve / double throw pulsed valve configuration allowed precise control of both the exchange reactant pressure and reaction period. For the experiments presented here, the reagent gas pressure was held at 6.0×10^{-7} torr during the prescribed reaction period. It should be noted that the pressure measurements were performed using an uncalibrated ion gauge mounted well away from the cell and as such the pressures noted were only useful for direct comparisons to H / D exchange experiments performed on the same instrument. The duration of the reaction period was controlled by a TTL pulse from the from the MIDAS data station [28]. Following the exchange period, the analyzer cell was allowed to pump down for sixty seconds prior to detection. The pressure in the analyzer at the time of detection was approximately 1×10^{-8} torr.

2.3 Results and Discussion

Figure 2.2 demonstrates a typical selection sequence. Poly dC15 analyte ions were produced with analytes representing $[M-4H]^{4+}$ through the $[M-9H]^{9-}$ charge states (Figure 2.2a). A zoom inset of the region around the $[M-7H]^{7-}$ charge state in

Figure 2.2. Typical Scenario for Isolation and H/D Exchange Demonstrated with [Poly dC15 -7H] ⁷⁻ Charge State

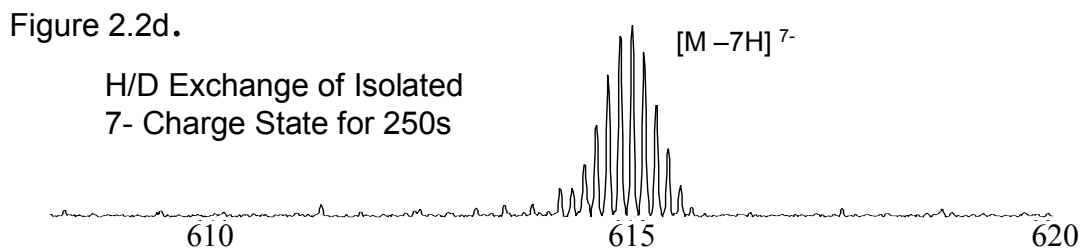
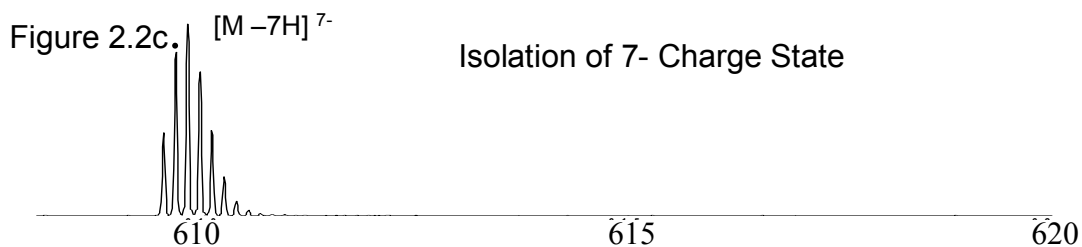
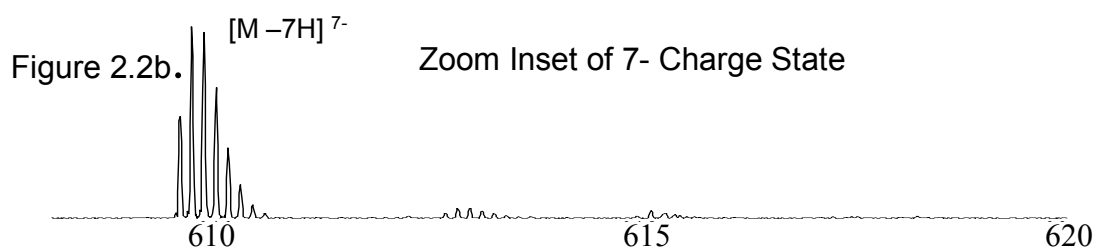
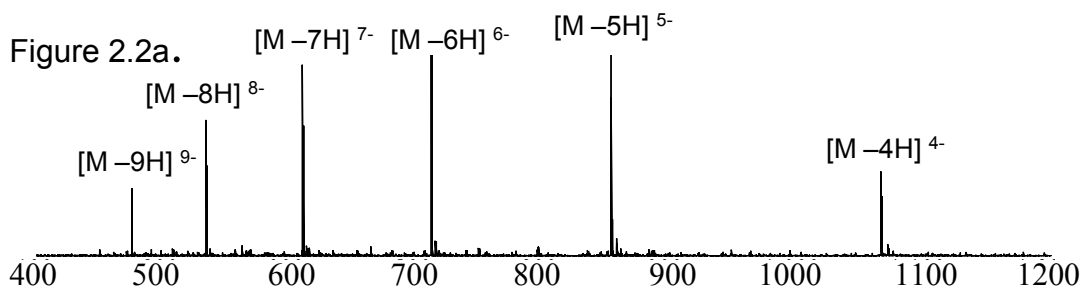
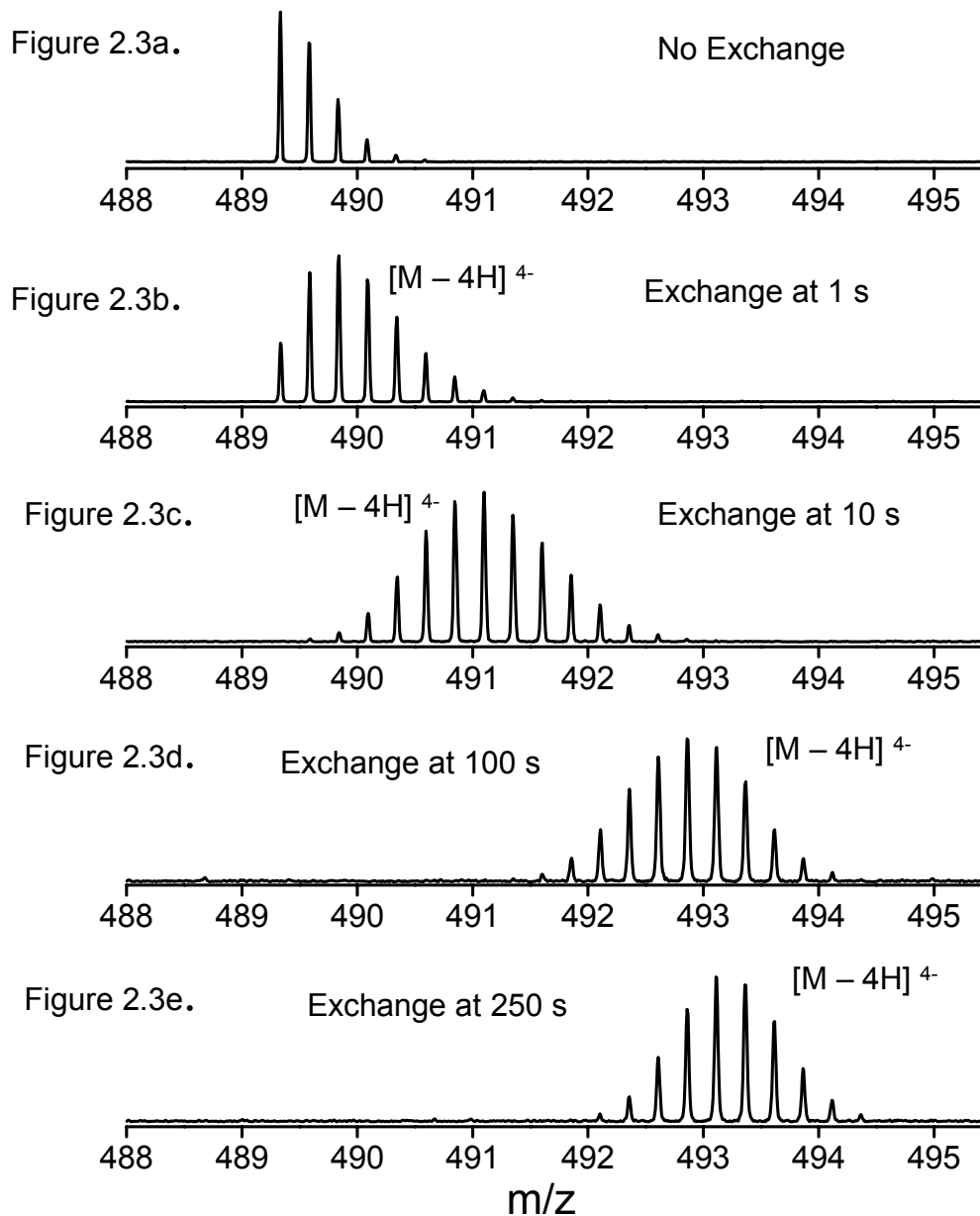


figure 2.2b showed the presence of sodium adducts. In order to achieve maximum isolation, non-target ions that were significantly removed in mass from the target ions ($\Delta m > 50$) were ejected from the cell with two independent swept frequency excite events. Any remaining non-target ions were ejected using a SWIFT waveform excite event. The combination of the two frequency chirps in conjunction with a narrow bandwidth SWIFT excite provided more efficient use of the arbitrary waveform generator memory than a single broad band SWIFT with the same isolation characteristics. The isolation afforded by this chirp-chirp-SWIFT is demonstrated in Figure 2.2c. Such isolation was critical to minimize confusion about the precursor ion as the H/D exchange experiment proceeded. Figure 2.2d demonstrates the later stage of H/D exchange experiment. In this case it was clear that the sodium adducts present in figure 2.2b would have served as interferences for the "exchanged" isotope envelope in figure 2.2d. Worthy of note that figure 2.2b and figure 2.2d had radically different abundances as noted by differences in the signal to noise of the two spectra. These spectra demonstrated one of the constraints of gas phase H/D exchange, ion loss through ion scattering and charge reduction occurred at extended reaction periods.

It is necessary to examine a "typical" exchange profile to establish some of the prominent features throughout the exchange period. In figure 2.3 one can track an H/D exchange experiment from zero exchange through the final exchange period of 250 seconds. The 250-second maximum was determined to be the point where the loss of ion signal outweighed the utility that additional time measurements provided. Additionally, experiments were designed such that each charge state of each analyte could be interrogated within a single day on a shared resource FT-ICR instrument. Figure 2.3a represents the $[M-6H]^{6-}$ charge state of poly dC13 with no H/D exchange (natural isotopic distribution). As evidenced in figure 2.3b

Figure 2.3. H/D Exchange Profile of [PolydC7 – 4H]⁴⁻ Charge State for Gas Phase Exchange with D₂O at ~ 6.0 x 10⁻⁷ torr



representing 1s H/D exchange and figure 2.3c representing 10 s H/D exchange, the H/D exchange reaction caused the isotope envelope to spread and there was a shift toward higher mass. As the study progressed to 100 s (figure 2.3d) and 250 s (figure 2.3e) the deuterium incorporation began to plateau although exchange was still proceeding.

In figure 2.4 another H/D exchange profile is presented; this profile represents the H/D exchange for the $[M-6H]^{6-}$ charge state of poly dC13. Figures 2.4a, 2.4b and 2.4c exhibit similar exchange characteristics to the analogous spectra of figures 2.3a, 2.3b and 2.3c. Figures 2.4d and 2.4e, however, demonstrate an interesting phenomenon that occurred when the $[M-6H]^{6-}$ charge state is allowed to react for a period of 250s. The isotope distribution appeared "bunched" with respect to the isotope distribution of figure 2.4c. Closer inspection revealed that the distribution in figure 2.4e bears some resemblance to the natural isotopic distribution demonstrated in figure 2.4a. When the initial isotope distribution appears shifted in mass but relatively intact, it is an indication that one has reached complete deuterium exchange. In this case the natural isotope distribution had not been completely returned, but the fact that figures 2.4d and 2.4e are nearly identical suggested that the reaction was approaching completion.

In order to compare H/D exchange characteristics of individual species objectively, a protocol must be established to catalog the extent of H/D exchange at each time period of the exchange profile. Perhaps the easiest technique would be to use the most abundant mass; however, this method works best when the isotope distribution remain unchanged. As evidenced in figures 2.3 and 2.4, it is seldom the case that the shape of isotopic distribution is not affected during the H/D exchange

Figure 2.4. H/D Exchange Profile of [Poly dC13 – 6H]⁶⁻ Charge State for Gas Phase Exchange with D₂O at ~ 6.0 x 10⁻⁷ torr

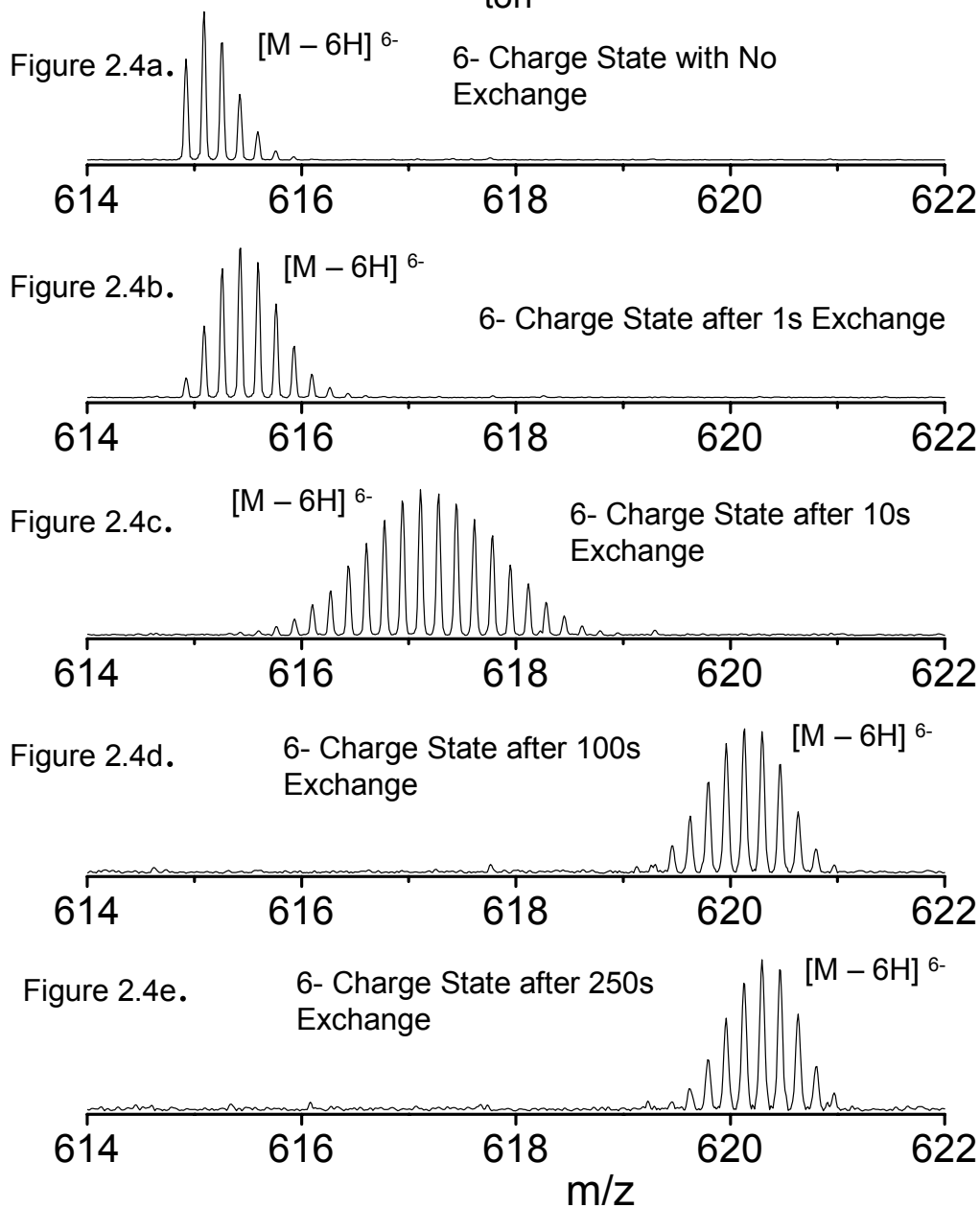
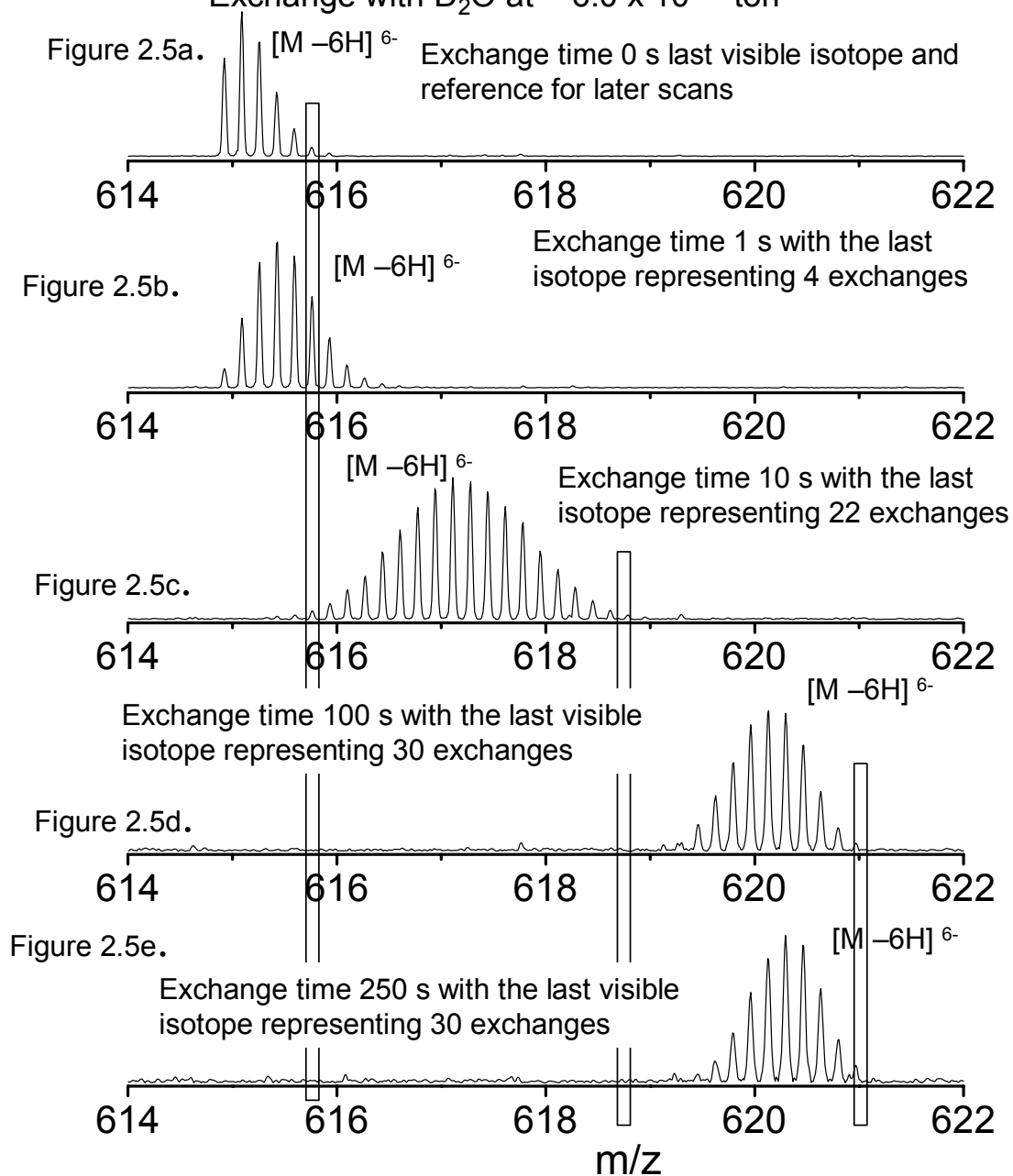


Figure 2.5. Determination of the Number of Exchanges of
[Poly dC13 – 6H]⁶⁻ Charge State for Gas Phase H/D
Exchange with D₂O at ~ 6.0 x 10⁻⁷ torr



experiment. An alternate protocol is often used to determine the amount of deuterium incorporated, which will be described as the last visible isotope method. Simply put the last resolved isotope from the current experiment is compared to the last resolved isotope in the unexchanged precursor. In figure 2.5 the spectra from figure 2.4 are used to graphically demonstrate this method. From the data, it was determined that poly dC13 [M - 6H]⁶⁻ charge state underwent 30 deuterium exchanges. There are a total of 34 hydrogens on poly dC13 [M - 6H]⁶⁻ charge state that were considered theoretically exchangeable. This yielded a total deuterium incorporation of 88%. Since the data suggested that the H/D exchange for this analyte was near completion, it is suggested that there were 6 of the theoretically exchangeable hydrogens that did not participate in H/D exchange of this system.

The previous data referred to exchangeable hydrogens. It is necessary to describe the term exchangeable hydrogens as it pertains to the poly dC system. The definition of an exchangeable hydrogen is any hydrogen attached to a heteroatom. Figure 2.6 demonstrates the exchangeable hydrogens of poly dC5 with neutral charge. Also worthy of note, negative charge is assumed to reside on the phosphate backbone of the oligonucleotide and as such one exchangeable hydrogen is assumed to have been removed from a phosphate moiety for each negative charge present [16].

One of the more novel results of the poly dC study involved the observation of multiple conformers during gas phase H/D exchange experiments in figure 2.7. Ideally, our model system would be without any higher order structure; however, the discovery of multiple conformers showed that our model system falls short of that mark. Multiple conformers created a challenge to objectively determine

Figure 2.6. Demonstration of Exchangeable Hydrogens in Oligonucleotides in an H/D Exchange Experiment

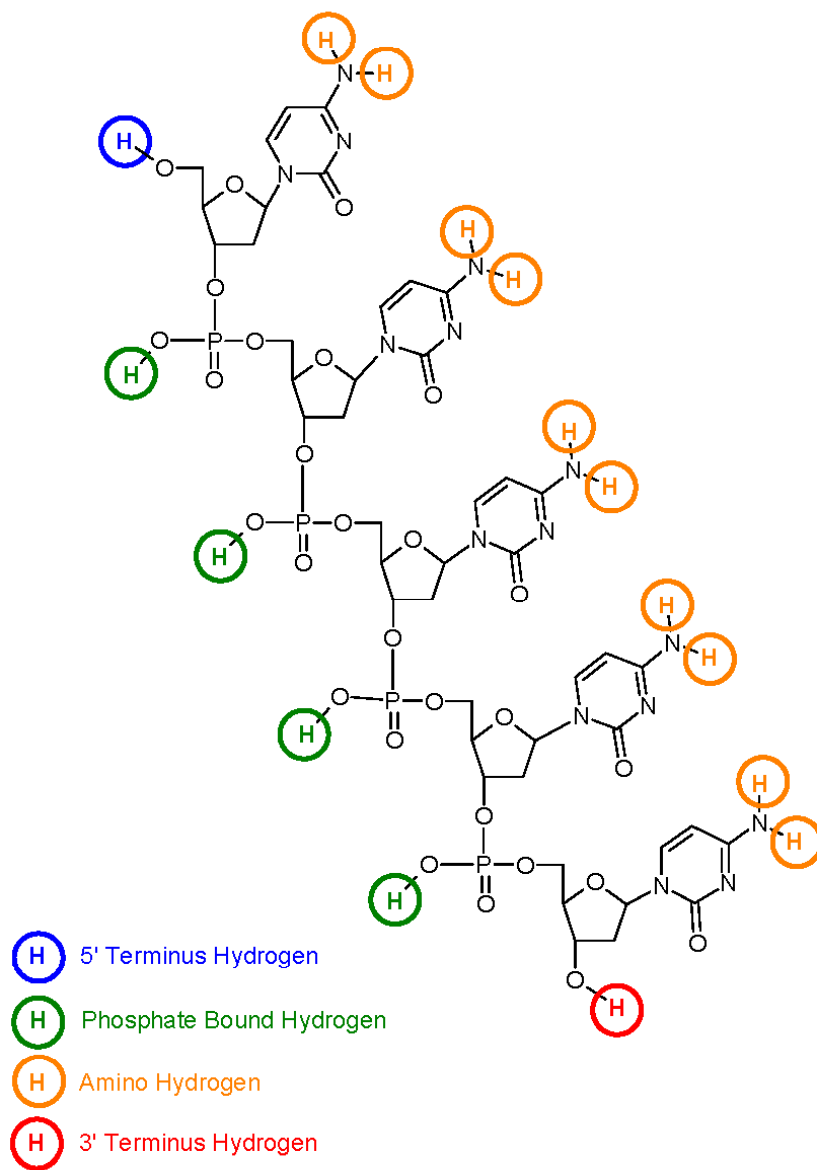


Figure 2.7. H/D Exchange Profile of [Poly dC14 -5H]⁵⁻ Charge State which Demonstrates Multiple Conformers

Figure 2.7a.

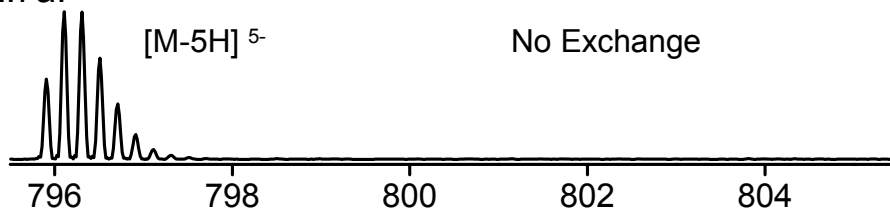


Figure 2.7b.

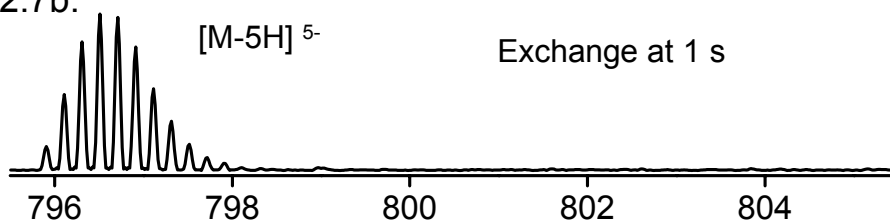


Figure 2.7c.

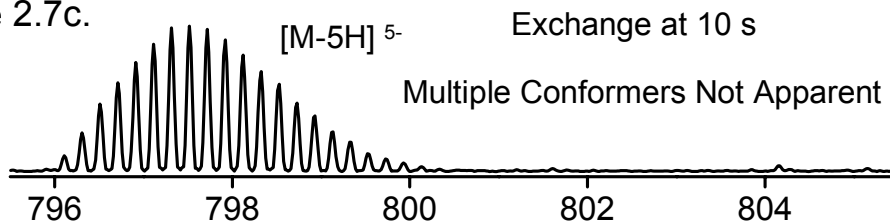


Figure 2.7d.

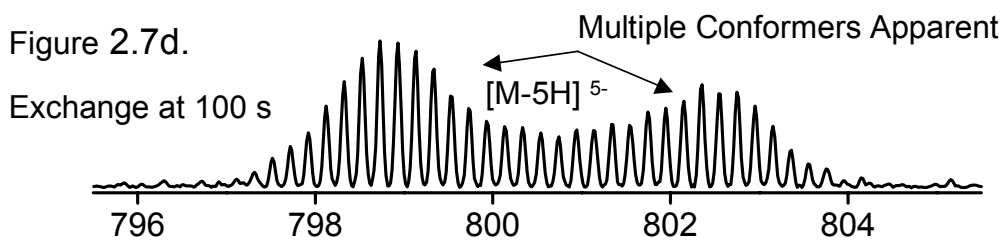


Figure 2.7e.

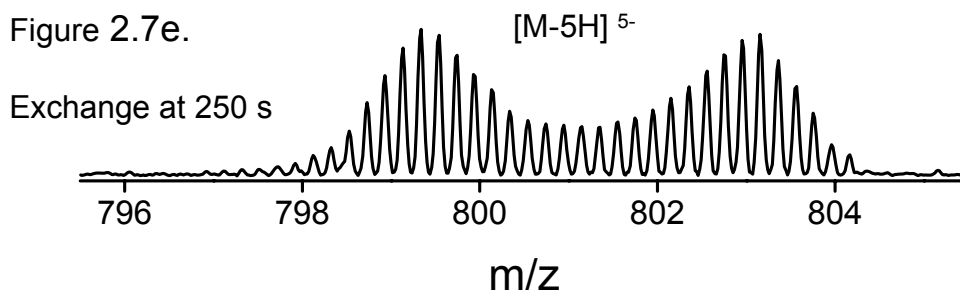
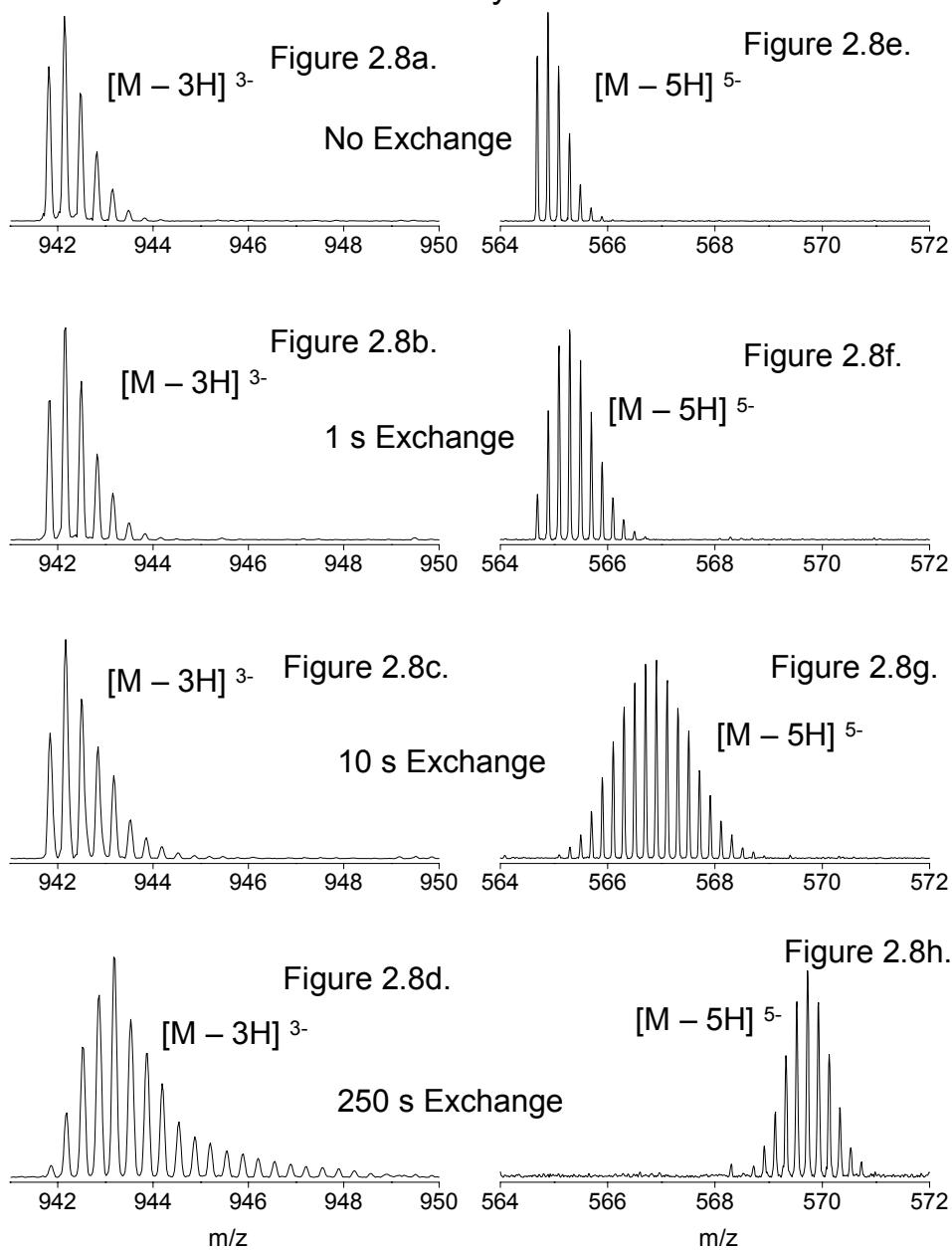


Figure 2.8. H/D Exchange Profiles of the 3- and 5- Charge States of Poly dC10



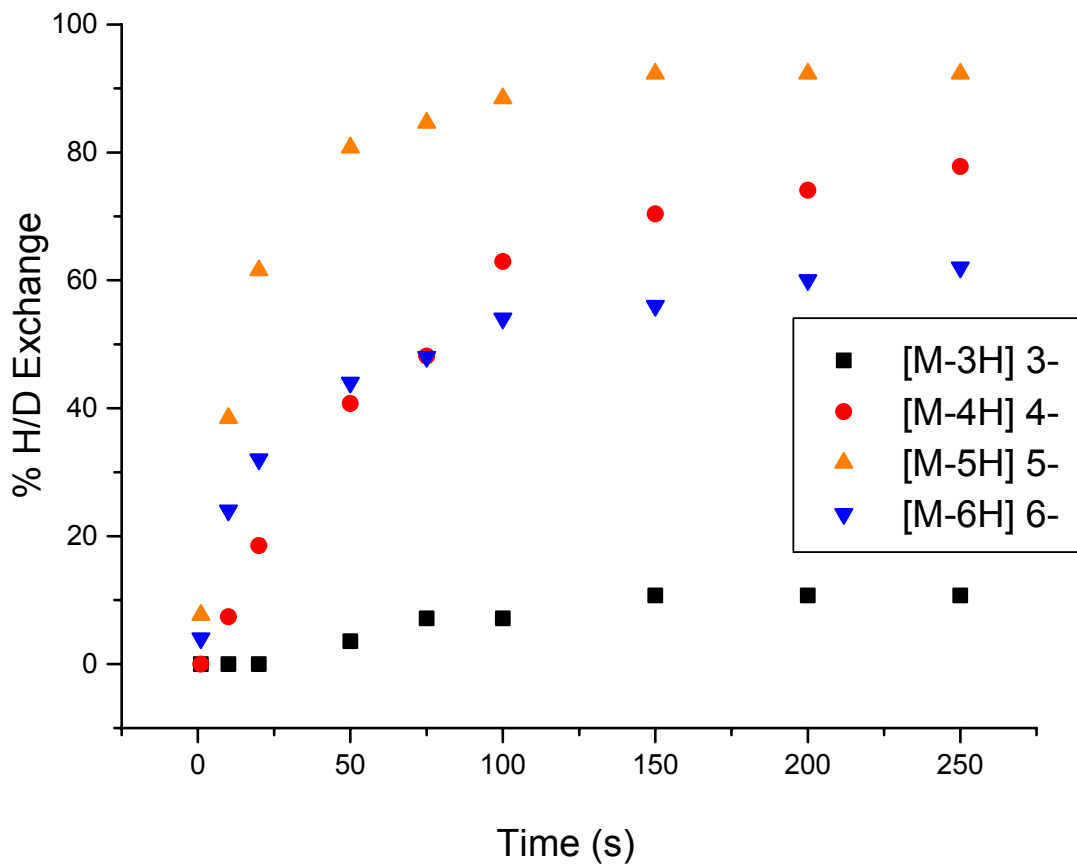
deuterium incorporation since any non-baseline resolved set of conformers made the determination of the last isotope of the lower mass conformer impossible. Additionally, comparing the reactivity of this charge state to other charge state leads to confusion as how to categorize the overall reactivity. The most troubling aspect of the discovery of multiple conformers in a system such as poly dC is whether the conformers differed primarily in physical accessibility to exchangeable hydrogens or in the reactivity in the gas phase of said hydrogens.

As previously mentioned, charge state is known to have an enormous effect on gas phase H/D exchange of negatively charged oligonucleotide. Poly dC is not immune to these charge effects. Figure 2.8 depicts the $[M - 3H]^{3-}$ and the $[M - 5H]^{5-}$ charge state of poly dC10. Figure 2.8a and 2.8e show the isolated unexchanged precursor ions. Figure 2.8b and 2.8f show the results of 1s of gas phase exchange with 6×10^{-7} torr of D_2O . Similarly figures 2.8c and 2.8g represent 10s of exchange for the respective charge state, and figures 2.8d and 2.8h show 250s of exchange. Using the last isotope method, the $[M - 3h]^{3-}$ of poly dC10 reached a maximum 50% deuterium incorporation and the $[M - 5H]^{5-}$ charge state reached 96%. It is unclear what affect the additional charge has on the physical conformation and what affect the charge has in facilitating the H/D exchange reaction.

2.4 Conclusions

Gas phase H/D exchange of oligonucleotides has been previously proven to be effective when applied to mononucleotides [16-18]. Based upon this initial work, projections were tested against larger systems. The results of the test were ambiguous, and it was apparent that more thorough fundamental work needed to be performed prior to extracting useful information from the gas phase H/D exchange

Figure 2.9. Percent Exchange as a Function of Charge State for Poly dC10



of such larger systems. Having constructed a database of poly dC analytes of lengths 2 to 15 cytidine units, gas phase experiments were performed using a stringent experimental protocol. This protocol precludes the interference of analyte ions through the use of complete isolation of each charge state within the trap prior to exchange reactants being admitted to the trap. Initial results suggested a tremendous charge state dependence upon H/D exchange. Figure 2.9 indicates such charge dependence for the various charge states of poly dC10. Figure 2.9 also demonstrates that the charge dependence was not absolute as the $[M - 6H]^{6-}$ charge state of poly dC10 undergoes less overall exchange than the $[M - 4H]^{4-}$ or $[M - 5H]^{5-}$ charge states. Percent H/D exchange was calculated as number of exchanges divided by the total theoretical number of exchanges. The initial results of this study show that it will be necessary to develop more sophisticated tools for the analysis of the gas phase H/D exchange of oligonucleotides before significant structural information will be attainable. The presence of multiple conformers in a series of poly deoxycytidine analytes was not predicted since it was not expected that short sequences of poly dC would coordinate into higher order structures.

REFERENCES

1. Watson, J.D. and F.H.C. Crick, *Nature*, 1953. **171**: p. 964-967.
2. Watson, J.D. and F.H.C. Crick, *Nature*, 1953. **171**: p. 737-738.
3. Voet, D. and J.G. Voet, *Biochemistry, Second ed.* 1995, New York: John Wiley & Sons, Inc.
4. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the noncovalent structure of proteins by amide hydrogen exchange and mass spectrometry*. *J. Mass Spectrom.*, 1997. **32**(2): p. 135-146.
5. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the Non-covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry*. *J. Mass Spectrom.*, 1997. **32**: p. 135-146.
6. Dharmasiri, K. and D.L. Smith, *Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins*. *Anal. Chem.*, 1996. **68**(14): p. 2340-2344.
7. Akashi, S. and K. Takio, *Probing higher order structure of proteins by hydrogen/deuterium exchange and mass spectrometry*. *J. Mass Spectrom. Soc. Japan*, 2000. **48**(2): p. 94-100.
8. Katta, V. and B.T. Chait, *Hydrogen/deuterium exchange electrospray ionization mass spectrometry: a method for probing protein conformational changes in solution*. *J. Am. Chem. Soc.*, 1993. **115**(14): p. 6317-21.
9. Anderegg, R.J., et al., *The Mass Spectrometry of Helical Unfolding in Peptides*. *J. Am. Soc. Mass Spectrom.*, 1994. **5**: p. 425-433.
10. Winger, B.E., et al., *H/D Exchange*. *J. Am. Chem. Soc.*, 1992. **114**: p. 5897-5898.
11. Suckau, D., et al., *Co-Existing Stable Conformations of Gaseous Protein Ions*. *Proc. Natl. Acad. Sci. USA*, 1993. **90**: p. 790-793.

12. McLafferty, F.W., et al., *Gaseous Conformational Structures of Cytochrome c*. J. Am. Chem. Soc., 1998. **120**: p. 4732-4740.
13. Freitas, M.A., et al., *High Field FT-ICR Mass Spectrometry for Simultaneous Trapping and Gas-Phase H/D Exchange of Peptide Ions*. J. Am. Soc. Mass Spectrom., 1998. **9**: p. 1012-1019.
14. Freitas, M.A., et al., *Gas-Phase Bovine Ubiquitin Cation Charge State Conformations Resolved by Gas-Phase Hydrogen/Deuterium Exchange Rate and Extent*. Int. J. Mass Spectrom, 1999. **185/186/187**: p. 565-575.
15. Cassady, C.J. and S.R. Carr, *Elucidation of isomeric structures for ubiquitin $[M + 12H]^{+12}$ ions produced by ESI-MS*. J. Mass. Spectrom., 1996. **31**: p. 247-254.
16. Robinson, J.M., et al., *Hydrogen/Deuterium Exchange of Nucleotides in the Gas Phase*. Anal. Chem., 1998. **70**: p. 3566-3571.
17. Freitas, M.A., et al., *Gas Phase RNA and DNA Ions. 1. H/D Exchange of the $[M - H]^-$ Anions of Nucleoside 5'-Monophosphates (GMP, dGMP, AMP, dAMP, CMP, dCMP, UMP, dTMP), Ribose-5-monophosphate (R5P), and 2-deoxyribose-5-monophosphate (dR5P) with D_2O and D_2S* . J. Am. Chem. Soc., 1998. **120**: p. 10187-10193.
18. Green-Church, K., et al., *Gas-Phase Hydrogen Deuterium Exchange of Positively Charged Mononucleotides by use of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 2001. **12**: p. 268-277.
19. Freitas, M.A. and A.G. Marshall, *Gas Phase RNA and DNA Ions 2. Conformational Dependence of the Gas-Phase H/D Exchange of Nucleotide-5'-Monophosphates*. J. Am. Soc. Mass Spectrom., 2001. **12**: p. 780-785.
20. Felix, T., M.L. Reyzer, and J.S. Brodbelt, *Hydrogen/deuterium exchange of nucleoside analogs in a QIT*. Int. J. Mass Spectrom. Ion Processes, 1999. **190/191**: p. 161-170.

21. Reyzer, M.L. and J.S. Brodbelt, *Gas-Phase H/D exchange Reactions of polyamine complexes : (M= h), (M + metal), (M+2H)*. J. Am. Soc. Mass Spectrom., 2000. **11**: p. 711-721.
22. Hofstadler, S.A., K.A. Sannes-Lowery, and R.H. Griffey, *Enhanced gas-phase hydrogen/deuterium exchange of oligonucleotide and protein ions stored in an external multipole ion reservoir*. J. Mass Spectrom., 2000. **35**: p. 62-70.
23. Robinson, J.M., *Examination of Gas Phase Conformations of Oligonucleotides Using Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*, in *Department of Chemistry and Biochemistry*. 1998, University of Texas at Austin: Austin, Texas. p. 172.
24. Bartlett, M.G., et al., *The Effect of Backbone Charge on the Collision-Induced Dissociation of Oligonucleotides*. J. Mass Spectrom., 1996. **31**(11): p. 1277-1283.
25. Senko, M.W., et al., *Electrospray Ionization FT-ICR Mass Spectrometry at 9.4 Tesla*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1824-1828.
26. Senko, M.W., et al., *External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 1997. **8**: p. 970-976.
27. Marshall, A.G., T.C.L. Wang, and T.L. Ricca, *Tailored Excitation for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Chem. Soc., 1985. **107**: p. 7893-7897.
28. Senko, M.W., et al., *A High-Performance Modular Data System for FT-ICR Mass Spectrometry*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1839-1844.

Chapter 3 – Analytical Method for Characterization of Oligonucleotides by Negative Ion Gas Phase H/D Exchange

3.1 Introduction

Initial studies of negative mode gas phase H/D exchange experiments of selected oligonucleotides demonstrated that the phenomenon of negative mode gas H/D exchange is a complex process requiring more in depth techniques than simple total deuterium incorporation [1, 2]. Many subtle differences in H/D exchange can be essentially invisible in such an initial investigation. Additionally, any analytical technique must be as objective and be as immune to noise as possible. Lastly, since the raw data sets for even a simple H/D exchange study can be very large, the ideal solution would allow easy implementation or better yet be fully automated.

Analytical tools are described which have been refined to mine more of the complex exchange information contained within the original oligonucleotide gas phase exchange study. The last visible isotope method was replaced by center of mass calculation for all subsequent calculations. Center of mass relies upon isotope-counting and abundance weighting of all isotopes and is more objective and immune to ion loss than the last isotope method. While complete automation of advanced techniques of maximum entropy analysis or the fitting of such results to a series of reaction rates is not currently available, there has been considerable effort to increase the accuracy, usability, and production of these techniques as well.

Oligonucleotide data was subjected to this rigorous protocol. Results were described and discussed in detail. Finally, the applicability of a gas phase negative

mode H/D exchange protocol was evaluated for effectiveness as a probe for biologically active systems.

3.2 Methods

3.2.1 Sample Preparation

Sample oligonucleotide solutions were prepared as 10 –100 $\mu\text{g} / \text{ml}$ in 1:1 IPA: water with 40 μM ammonium acetate added. Oligonucleotides were purchased from TriLink Biotechnologies (San Diego, CA) and were used without further purification. Final solution were prepared by serial dilution from stock solution of 1 mg / ml in 1:1 IPA and water.

3.2.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Experiments were performed on a 9.4 Tesla FT-ICR instrument at the National High Magnetic Field Laboratory in Tallahassee, FL ([3, 4]). The electrospray source was a Chait type source incorporating a resistively heated capillary for desolvation. Ions were focused into a linear octopole ion trap for external accumulation. After accumulation times of 2s to 60s, the ions are gated out of the octopole trap, through octopole transfer optics and into a 4 in diameter open cylindrical Penning ion trap.

Electrospray conditions were individually optimized for each to produce optimum ion signal. Spray conditions ranged from flow ranges of 300 to 500 nl / min , and needle voltage ranged from -2500V to -2900V . Individual charge states were isolated prior to a gas phase H/D exchange period ranging from 1s to 250s. Deuterated water was leaked into the ICR cell at an uncalibrated pressure of 6.0 x

10^{-7} torr for the prescribed reaction period. The MIDAS data station provided instrument control of the FT-ICR instrument [5].

3.2.3 Data analysis

H/D incorporation was calculated using the MIDAS data analysis package [5] developed at the National High Magnetic Field Laboratory (Tallahassee, FL.) Several techniques and algorithms within the MIDAS data analysis package were developed specifically to facilitate the analysis H/D exchange data. Automated center of mass calculation and the HD Helper functionality were developed by the author to speed the analysis of the formidable amount of data inherent to time dependent H/D studies. The maximum entropy method [6] functionality of MIDAS analysis was improved in both utility and stability for FT-ICR MS H/D exchange reactions. Origin Pro version 5.0 software package (Microcal Software, Northampton, MA) was employed to provide advanced curve fitting to the resultant exchange rate curves.

3.3 Results and Discussion

Work in the previous chapter demonstrated H/D exchange results calculated using the last visible isotope method. While the last isotope method has utility in the analysis of gross shifts in the exchange rate of analytes, the utility becomes limited when one begins to probe more subtle processes. The last isotope also suffers from some procedural shortcomings leading to systemic bias in the calculation of deuterium incorporation. The shortcomings are grouped into the following categories: user bias, ion loss at extended exchange period, and changes in the isotopic distribution envelope.

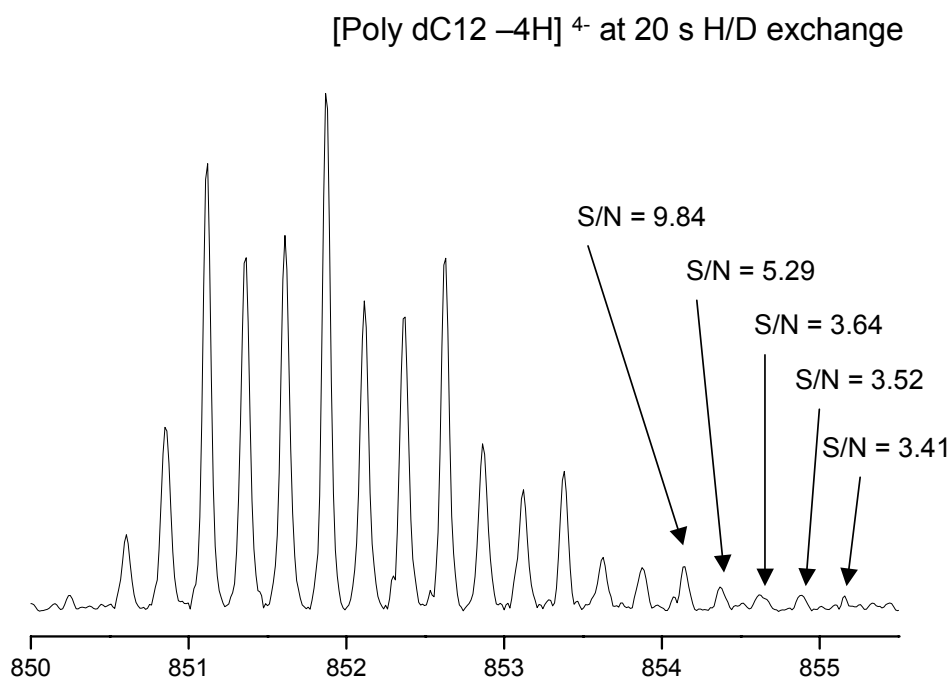
The last isotope method is dependent on the individual user to determine which peak to select as the final isotope visible. User bias is amplified since the last isotope visible is often at a low signal to noise ratio. Figure 3.1 illustrates the dilemma faced by a user determining which isotope to select as the last isotope.

Given the definition of signal to noise as described by Marshall [7], the signal to noise of each of the isotopes present in figure 3.1 were valid choices for the last isotope method. Nevertheless, one was faced with uncertainty about the accuracy of the smallest peak since the signal to noise of 3.46 could be considered borderline. With such low signal to noise peaks it becomes advisable to double-check the peak mass spacing to verify the identity of the peak. The acts of calculating signal to noise and verifying mass spacing negated the advantage of rapid H/D exchange determination garnered by using the last isotope method.

As an analyte undergoes exchange, the total ion signal becomes spread across additional mass channels created by the exchange process. The result is that for any specific isotope the abundance of the exchanged ion is less than the parent ion abundance for the same analyte ion population. Performing H/D exchange in the gas phase further complicates reliable application of the last isotope method for the determination of deuterium incorporation because total ion signal is decreased at extended exchange periods through collisional loss from the ICR cell. Additional effects of ion fragmentation and charge stripping often contribute to ion loss as well.

The combined effect of decreased spectrometric density and decreased total ion population leads to lower ion signal at increased gas phase exchange periods. Such loss of ion signal is the most pronounced on the lesser abundant analyte ions as many lesser abundant ions can shift from readily observable signal to noise ratios

Figure 3.1 Dilemma of the Last Isotope Method: Which One?



to questionable or non observable signal to noise ratios. The loss of low abundance ions is devastating to the accuracy of the last isotope method since it these low abundance ions that are the sole criteria that is used to determine the deuterium incorporation. The net effect of the loss of low abundance ions is a calculated deuterium incorporation that is mass shifted to lower incorporation when compared to other spectra calculated in similar fashion (at shorter exchange periods.) In cases of very low signal to noise at long reaction periods, the amount of deuterium can be calculated to be less than the amount of exchange at a shorter exchange period. An extreme example of such ion loss can be seen in figure 3.2. Figure 3.2a shows the ion population immediately after ion isolation in the cell where the signal to noise of 153 for the most abundant isotope. In figure 3.2b the signal to noise slips to 6.7 after 250 s of H/D exchange. It is likely that the last isotope visible in the spectra in figure 3.2b was affected by the signal to noise and that signal from additional isotope peaks was "lost" in the noise.

Figure 3.3 provides insight of the loss of ion signal at extended exchange periods; the spectrum represents [poly dC8 - 4H]⁴⁺ that has isolated in the ICR cell prior to a 100 s H/D exchange experiment. The observance of abundant fragment ions suggests that fragmentation is a major pathway for ion loss in this experiment. A combination of stringent ion isolation and the resolution of the FT-ICR MS were employed to insure that interferences from ion fragmentation and charge stripping do not bias the exchange data.

The final detractor for the last isotope method applied to H/D exchange is the assumption one accepts that changes in the relative isotope abundances are small in comparison to the total amount of exchange. This is by far the most serious shortcoming of the last isotope method for the determination of deuterium

Figure 3.2 Changes in Signal to Noise Causes Calculated Exchange to be Lower at Longer Exchange Periods for LIM

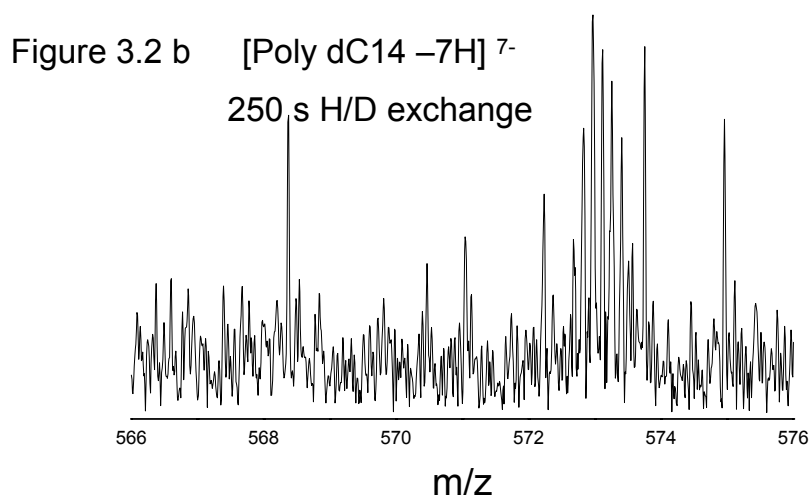
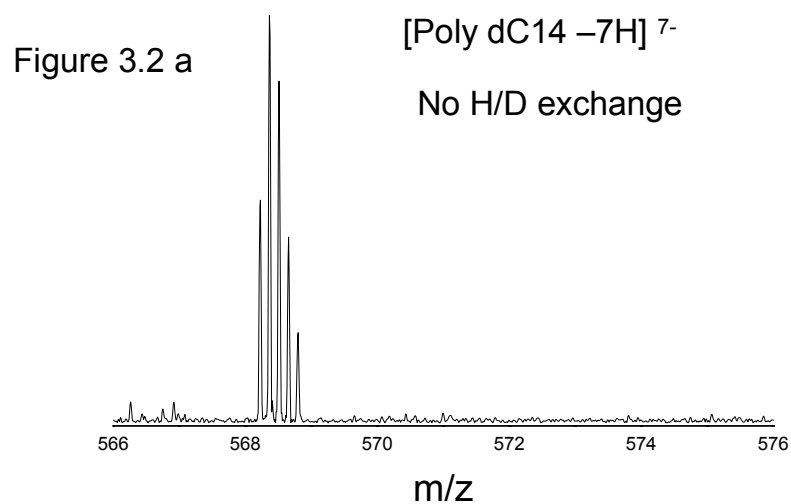


Figure 3.3 Spectra of [Poly dC8 -4H]⁴⁻ with Ion Dissociation Demonstrated (H/D Exchange for 100s at 6.0 x 10⁻⁷ torr)

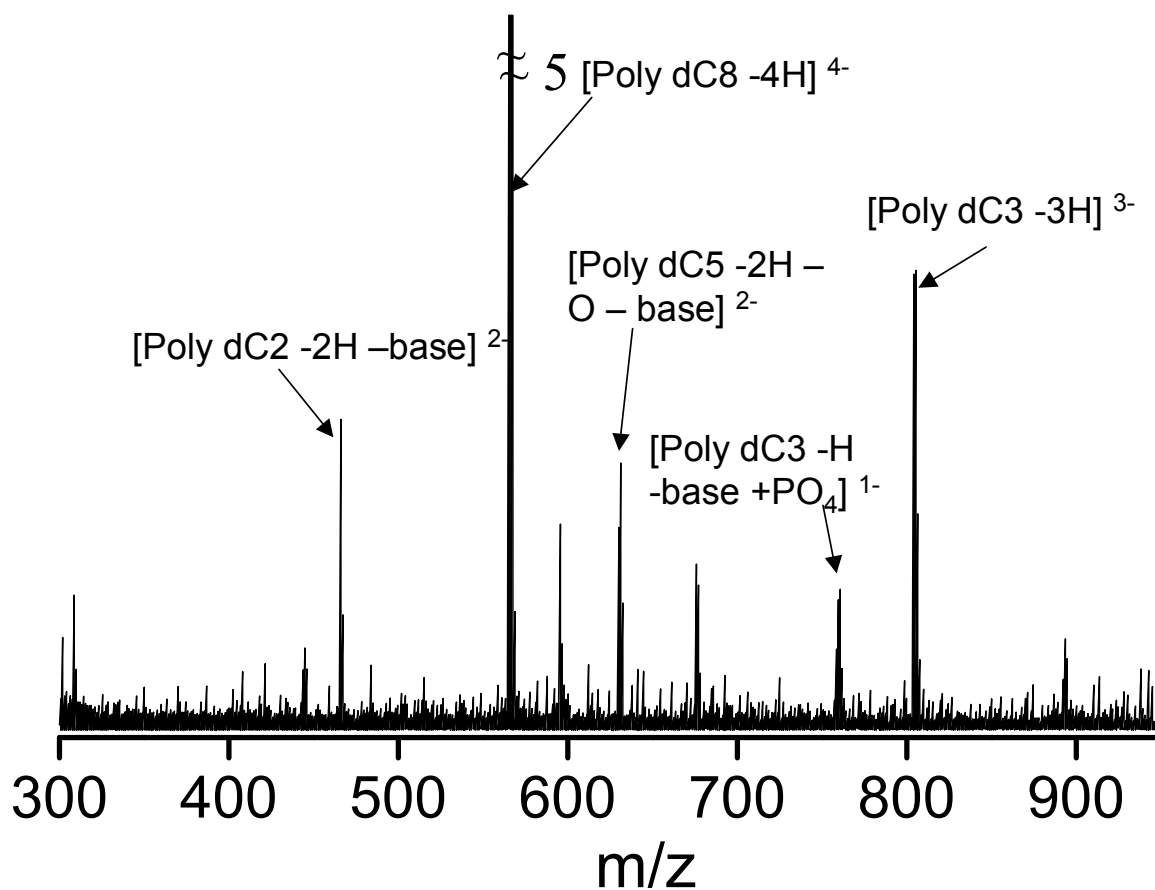
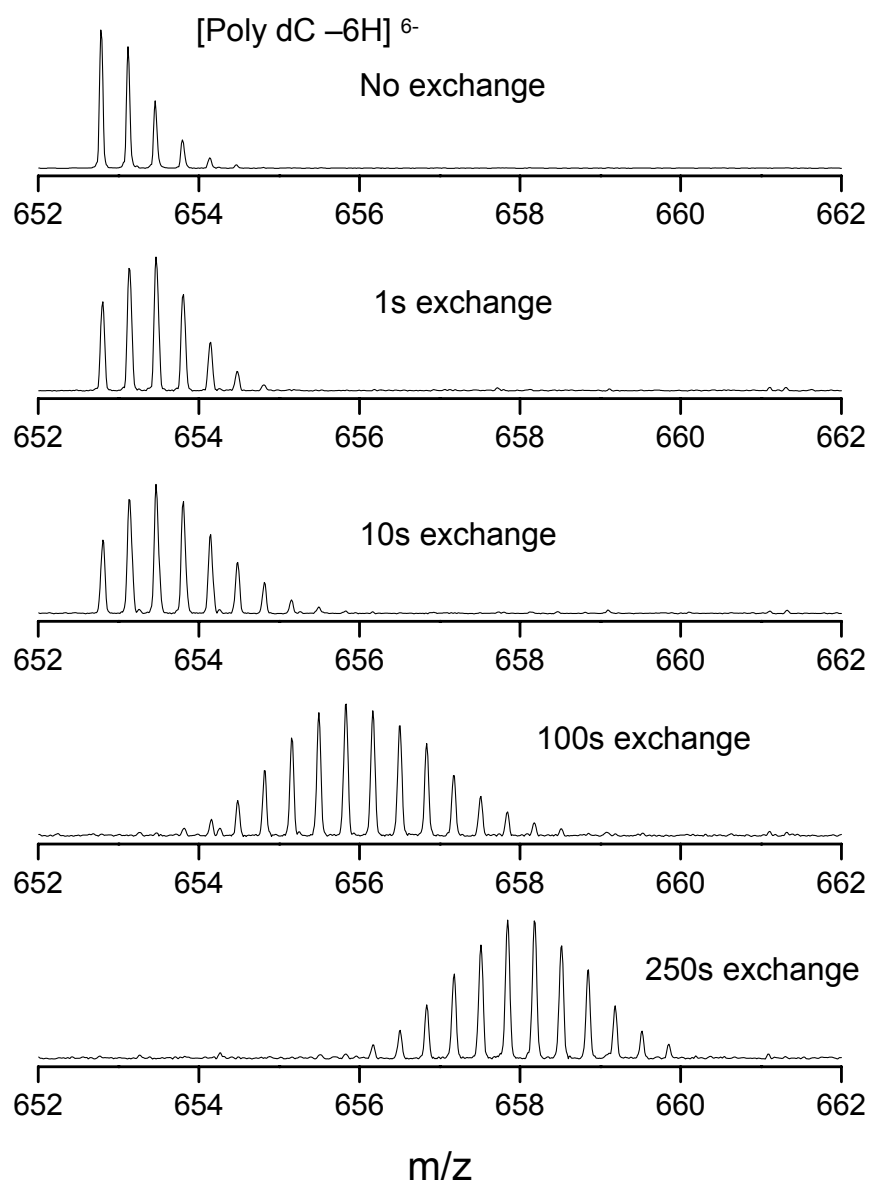


Figure 3.4 Typical H/D exchange demonstrating the spread of isotope distribution

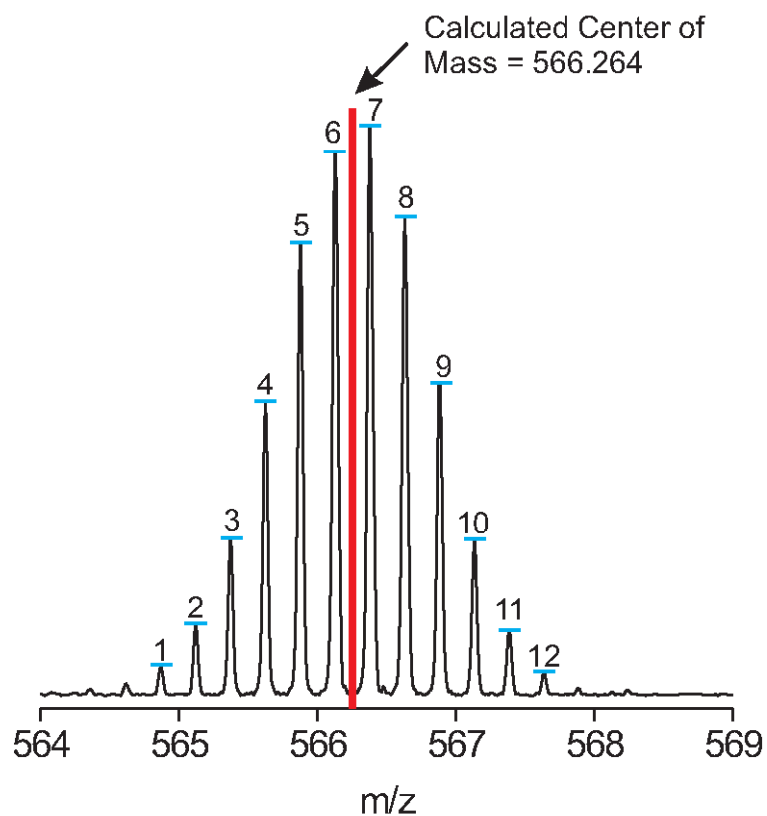


incorporation. Figure 3.4 shows a typical H/D exchange progression for a small oligonucleotide. The initial spectrum in figure 3.4a demonstrates the natural isotopic distribution of the oligonucleotide prior to H/D exchange. In figure 3.4b through 3.4d, the H/D exchange is increased from 1s to 10s and finally to 250s. As the exchange reaction proceeds, the isotope envelope spreads as the analyte shifts to higher mass.

Center of mass calculations have the advantage of using all of the analyte peaks and thus are less susceptible to many of the difficulties associated with the last isotope method. Center of mass involves a peak-picking algorithm followed by a mass-averaging algorithm. A schematic of the center of mass algorithm is depicted in figure 3.5. The mass of each peak is weighted by its abundance. Under normal FT-ICR conditions the abundance of a mass spectrometric peak can be described by peak height since this height is relative to the area under the curve of the peak. Because the signal to noise of each spectrum can vary especially at extended exchange period, the weighted abundance is subsequently scaled by the total ion abundance, yielding what is in effect an abundance weighted average mass that is called center of mass for the analyte in question.

The advantages of a center of mass calculation over the last isotope method focus on all the shortcomings listed for the use of the last isotope method in gas phase exchange reactions. The user bias is removed by the use of uniform peak picking parameters to describe the analyte peaks of each spectrum. While the user must select these parameters with care based on the characteristics of the spectra, the parameters once selected are applied to each spectrum in the series equally. While the loss of less abundant ions will affect the center of mass calculation, the effect is minimized by the fact that low abundance ions contribute the least weight

Figure 3.5 Spectrum of [Poly dC8 – 4H]⁴⁻ Graphically Representing the Determination of H/D Exchange at 200s Using Center of Mass Method



Calculation of Center of Mass for the Above Spectra is Represented by :

$$\text{Center of Mass} = \frac{\sum(\text{Mass}_n \times \text{Peak Height}_n)}{\sum(\text{Peak Height}_n)}$$

where n represents peaks 1 to 12.

to the center of mass calculation. Finally, the changes of relative isotope abundances are easily managed since changes in individual isotopes still contribute to the calculated value for deuterium incorporation.

There are some constraints one should consider in the application of center of mass to H/D exchange studies. Center of mass requires that the peak-picking algorithm easily identifies all of the isotopes. This limits its use for highly charged ions produced by electrospray ionization in many mass spectrometers. Additional limitations involve the number of calculations required to apply this technique to the hundreds or thousands of data files generated by H/D exchange studies. If one were to use a spreadsheet program and enter the data by hand the chance of data entry error would be a major limitation in the reliability of the calculated data. Finally, the effort required to perform these calculations would limit this technique to a proof of concept exercise.

The limitations of center of mass calculations are addressed in this work through the use of a mass spectrometer with excellent resolution and the automation of the center of mass calculation. The 9.4T FT-ICR at NHMFL easily resolved the isotopes for the present oligonucleotide H/D exchange study. With baseline resolution of the isotopes of the analyte, the peak-picking algorithm within MIDAS is able to accurately determine mass and abundance of the analyte. To combat both the data entry error and the data analysis overhead, automated center of mass calculation were implemented in MIDAS analysis.

The automation of center of mass calculation posed one additional challenge when applied to large data systems. Often one finds that the mass spectrum contains peak that are not related to the H/D exchange experiment. These interference peaks can arise from other analytes in the solution but are more often attributed to

contaminants or noise inherent to the instrument. In order to rapidly and effectively eliminate these interference peaks from the center of mass calculations a procedure called HD helper added to the MIDAS analysis package. A sample of the HD helper is provided in figure 3.6. HD helper provides the user with a visual reference to the peak list that was produced from the current spectrum. In addition the standard peak list parameters, HD helper provides the user with the relative mass spacing between adjacent peaks and a center of mass calculation for the current peak list (the third column in figure 3.6 labeled "Mass Diff to Next Peak"). Using the mass spacing column, the user can quickly scan the current peak list to verify that each member of the current peak list has the correct mass spacing given the charge of the initial analyte ion. If the user detects spurious peaks in the HD helper, these peaks can be selected and deleted from the peak list. The end product of the HD helper is a peak list with its center of mass calculated within the matter of seconds.

A direct comparison of the last isotope method and the center of mass methods is represented in figure 3.7. There are several interesting contrasts indicated for the two methods for the study of poly dC7. When comparing the results in figures 3.7a and figure 3.7b, one discovers that the quantized nature of the last isotope method suggest that no exchange has occurred for the 1 s exchange period. If one attempts to fit a rate constant to data with zero exchange for a given exchange period, the statistics suggest a very poor fit. Early attempts by the author to fit last isotope method data proved unsatisfactory, which led to the development of the center of mass technique. The effect of quantization can also be seen as a reaction reaches completion, in which case, the last isotope method reports identical exchange abundances for different exchange times. The center of mass method returns exchange data that indicated increased exchange at extended exchange time for each analyte in the oligonucleotide study.

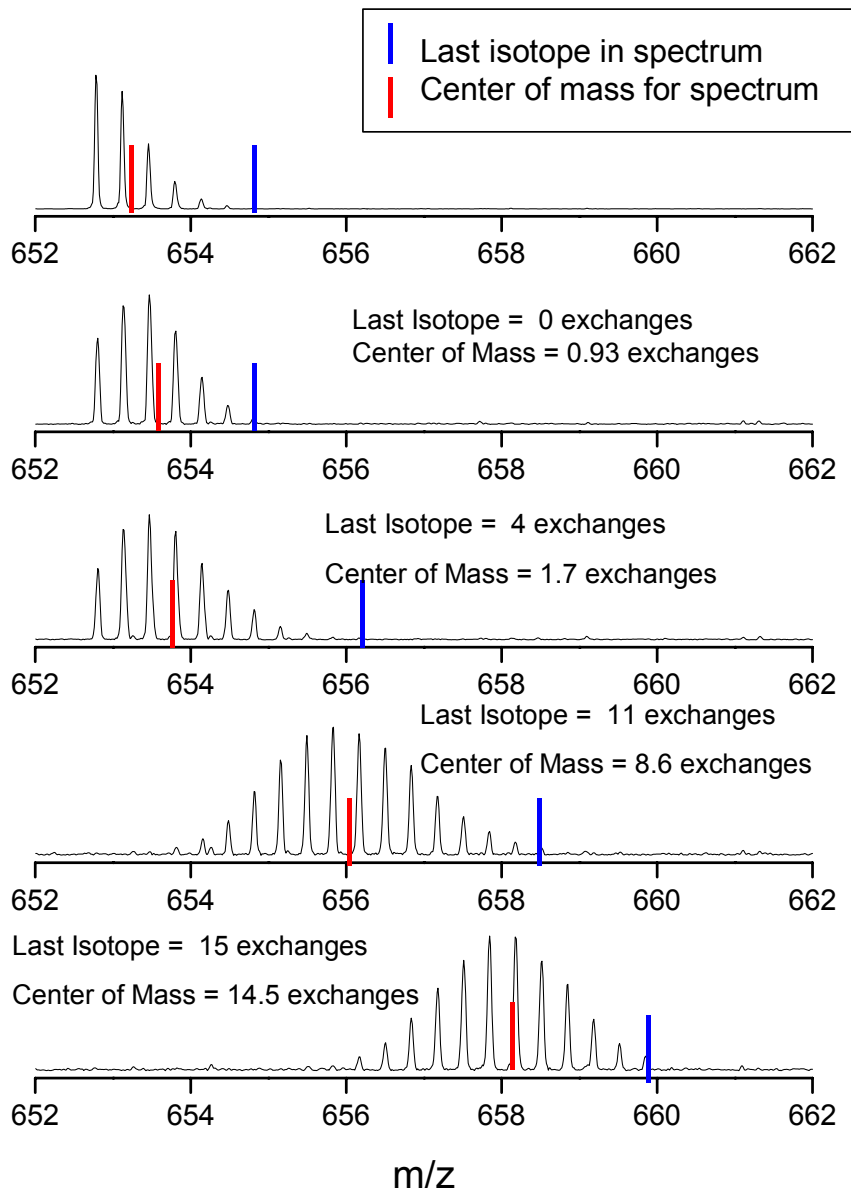
Figure 3.6 Screen Shot of HD Helper Tool from MIDAS Analysis

Number	m/z	Mass Diff to Next Peak	Frequency	Rel Abundance	Scaled Abundance	RP
1	613.9759	0.0000	235005.28	4.5675	2873735.7500	12322
2	614.1203	0.1444	234950.03	11.6291	7316782.0000	15399
3	614.2634	0.1431	234895.29	18.6689	11746047.0000	12316
4	614.4087	0.1453	234839.74	35.5471	22365404.0000	15391
5	614.5525	0.1439	234784.76	62.0343	39030500.0000	15388
6	614.6968	0.1443	234729.63	84.9644	53457576.0000	12308
7	614.8403	0.1434	234674.86	101.7997	64049924.0000	15381
8	614.9841	0.1438	234619.98	99.0304	62307544.0000	12302
9	615.1279	0.1438	234565.15	77.6353	48846284.0000	12299
10	615.2717	0.1438	234510.30	49.7871	31324878.0000	15370
11	615.4142	0.1425	234456.02	26.1952	16482005.0000	12293
12	615.4808	0.0666	234430.65	2.5385	1597168.8750	15364
13	615.5570	0.0763	234401.60	12.0756	7597694.5000	12290
14	615.7057	0.1487	234344.99	4.2113	2649620.5000	12287

Center of Mass

Close Save... Delete selected peaks 614.872

Figure 3.7 Direct Comparison of Last Isotope and Center of Mass Methods for $[\text{PolydC7} - 3\text{H}]^{3-}$



The other distinction evident when comparing the last isotope method and the center of mass method is that the last isotope method returns higher calculated exchange values in general than the center of mass method. This phenomenon is visible in figures 3.7c, 3.7d and 3.7e. This is a consequence of the isotope envelope "spreading" to higher mass as mentioned previously.

A maximum entropy method (MEM) technique has been shown to produce H/D exchange rate constants for mass spectrometric that correspond to those of NMR [6]. This MEM technique assumes that is possible for each exchangeable hydrogen in the analyte to have an unique rate constant and that the H/D exchange reaction follows pseudo first-order kinetics. The result for the exchange study of [poly dC10 -5H] ⁵⁻ is presented in figure 3.8. The algorithm returns a successful fit given the available exchangeable hydrogen with a standard deviation of 0.1 exchanges. The distribution of exchange rate constants for the above study is presented in figure 3.9. This represents the raw output of the MIDAS-based MEM algorithm. It should be noted that absolute value for the rate constants was not calculated because precise calibration for D₂O of the instrument's ion gauges was not practical prior to each instrument session. Given the relatively short period of 3 months over which the entire series was performed on the instrument qualitative comparisons between individual oligonucleotides are reasonable.

Figure 3.10 represents a MEM fit of the poly dC11 H/D exchange data for the [M - 6H] ⁶⁻ charge state. The MEM fit is present in black, Microcal Origin was used to produce the red curve that represented the best fit to the prominent peak in the data set. Integrating the total area under the curve and subtracting the area represented by the red curve yielded the subsequent boundary areas. On the right edge of figure 3.10, the MEM represented highly reactive hydrogens where H/D exchange proceeded at a rate beyond the FT-ICR instrumental ability to accurately

Figure 3.8 Maximum Entropy Method for the Determination of H/D Exchange Reaction Rates – Calculated Fit to [Poly dC10 – 5H] ⁵⁻ H/D Exchange Profile

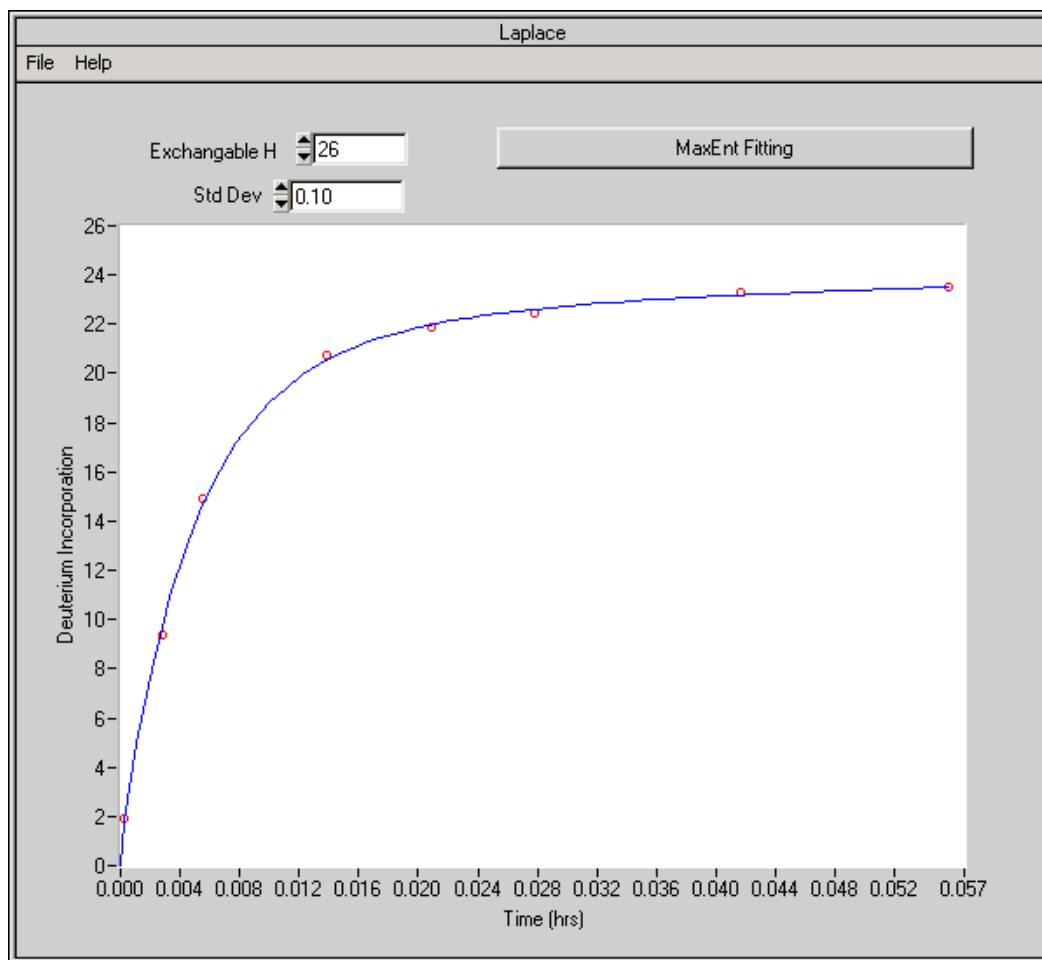
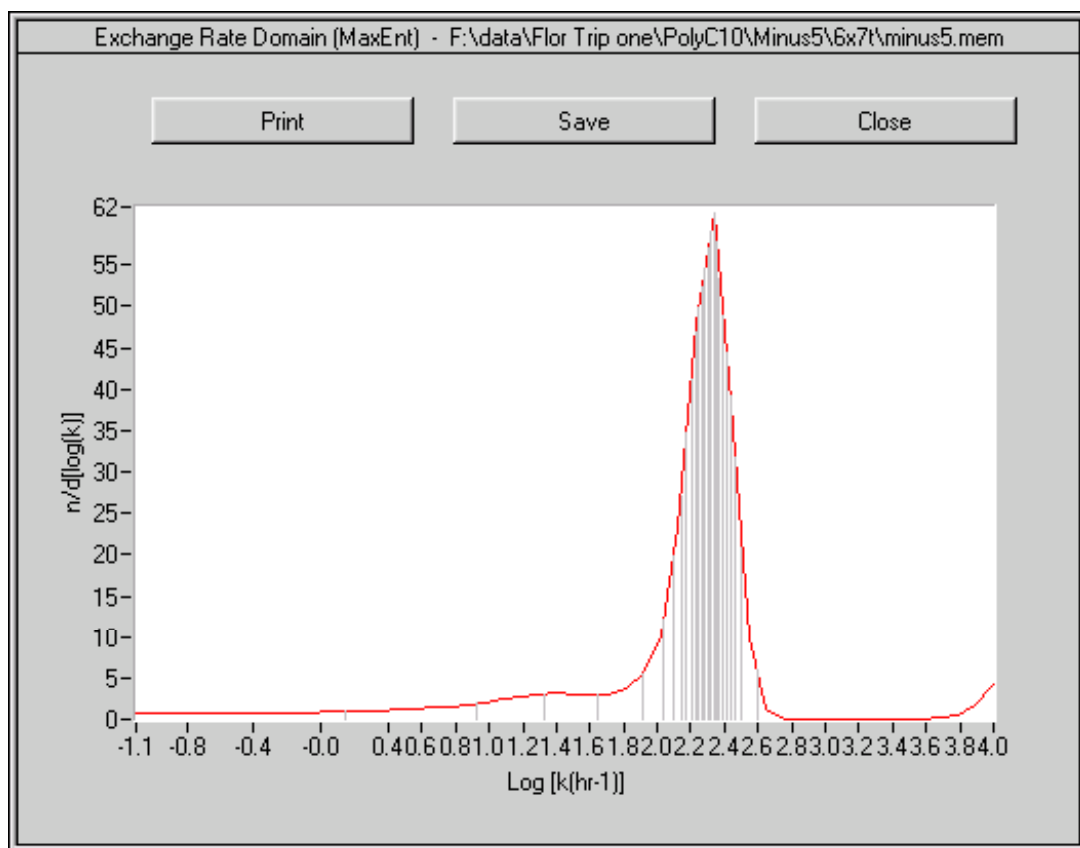


Figure 3.9 Maximum Entropy Method from MIDAS for [Poly dC10 – 5H] ⁵⁻



measure the reaction. In the left edge of figure 3.10, the MEM represents hydrogens that were reacting too slowly to be accurately measured within the 250 s reaction period of the experimental study. In order to convert these areas back into number of hydrogens, the area for each curve was normalized by the total area and then multiplied by the total number of exchangeable hydrogens of the poly dC11 species.

Given that figure 3.10 has three distinct regions of reactivity and the analyte poly dC11 has three functional types of exchangeable hydrogens (assuming that the end terminal hydrogens were treated equally), it was reasonable to assume that the reactivity was related to the functional structure of the hydrogens. When one compares the relative abundances of the area of the MEM and the relative amounts of exchangeable hydrogens, the relationship is obvious. In fact, the fit of the data to the theoretical abundance was approximately 98 % for poly dC11 $[M - 6H]^{6-}$. Figure 3.11 shows additional similarities between the experimental data and the available hydrogen abundances. In this case the fit was approximately 92 % for poly dC9 $[M - 5H]^{5-}$. The red curve represents the summing of the green curves to visually denoted the fit to the overall rate curve given overlapping peak areas.

While figures 3.10 and 3.11 demonstrate excellent agreement between the experimental and theoretical data, figure 3.12 demonstrated a much different scenario. With a fit of only 30 %, it was apparent that there was little correlation between the experimental data and the structure of the analyte. When the entire study was reviewed, it becomes obvious that only the higher charge states of poly dC undergo H/D exchange that allowed one to identify structural types of hydrogens related to the analyte species.

3.4 Conclusions

Figure 3.10 Results from MEM Study of [Poly dC11 – 6H]⁶⁻
Charge State

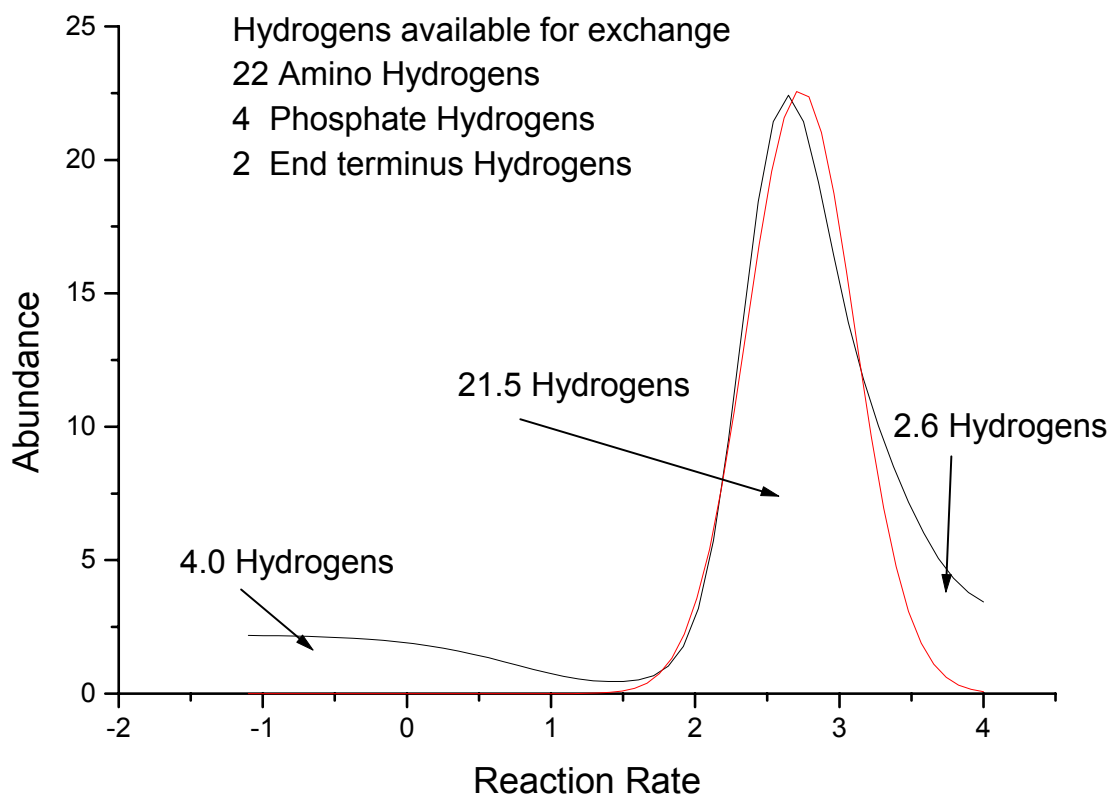


Figure 3.11 Results from MEM Study of [Poly dC9 – 5H]⁵⁻
Charge State

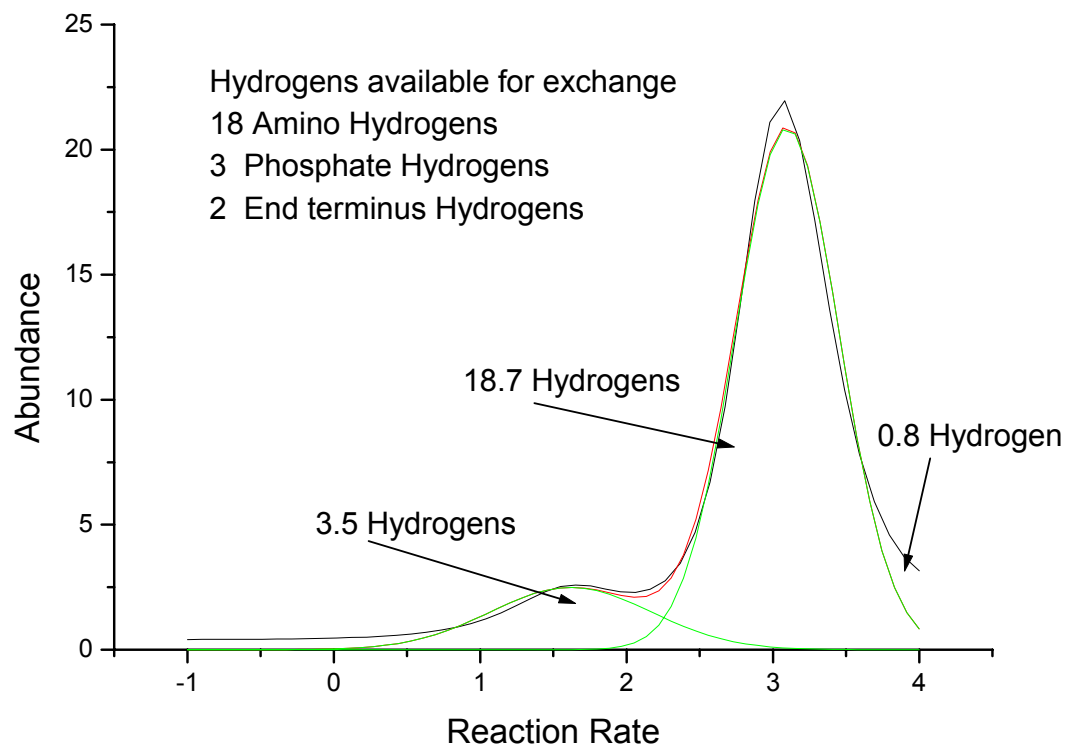
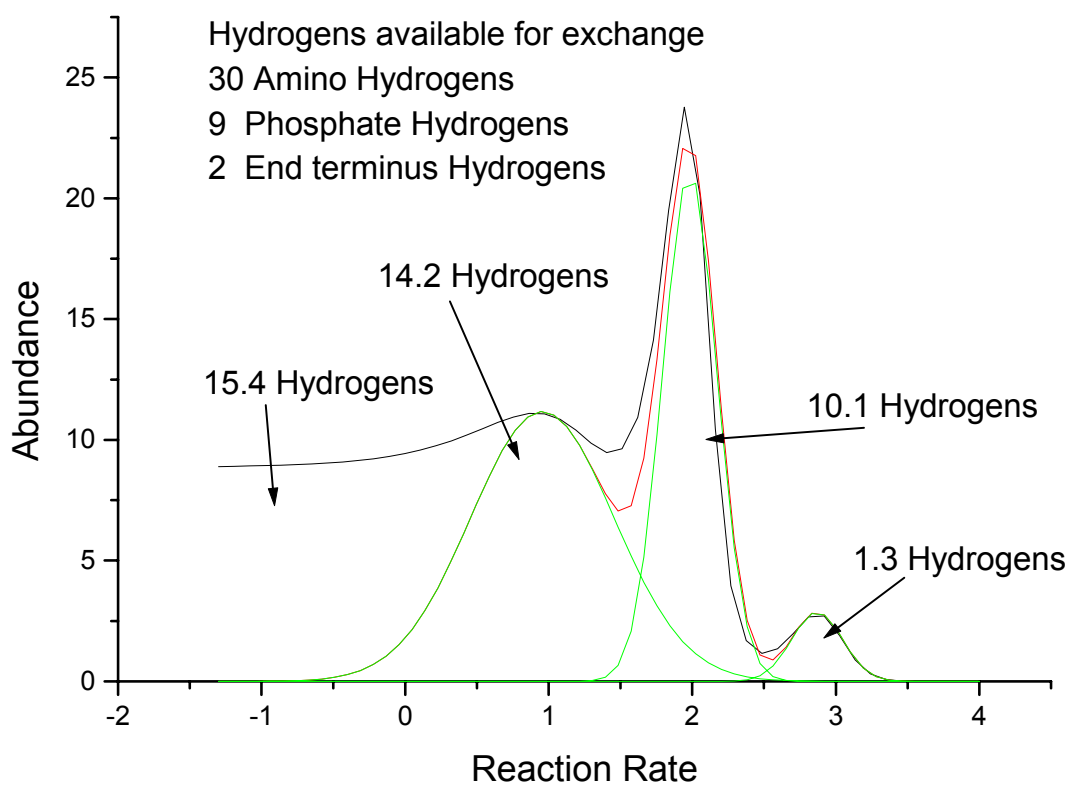


Figure 3.12 Results from MEM Study of [Poly dC15 – 5H] ⁵⁻ Charge State



Gas phase exchange studies were performed on a class of poly dC_n where the n represented residue of 2 to 15 cytidine units. The results showed that it was possible to determine the number of structural hydrogens present on these oligonucleotides at high charge states. The prerequisite for such high charge states limits the applicability of these methods to biological system. While it is still debatable as to whether any gas phase structure is comparable to the solution phase counterpart, it is even less likely that the highly charged ions are similar to biologically active species. While gas phase H/D exchange may not solve the mysteries of biological interaction of oligonucleotides, the techniques, such as center of mass and MEM, that were developed for the gas phase studies can be applied to improve the accuracy of future H/D exchange studies in the gas and liquid phase.

References

1. Griffey, R.H., et al., *Gas Phase Hydrogen-Deuterium Exchange in Phosphothioate d(GTCAG) and d(TCGAT)*. Rapid Commun. Mass Spectrom., 1999. **13**: p. 113-117.
2. Robinson, J.M., *Examination of Gas Phase Conformations of Oligonucleotides Using Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*, in *Department of Chemistry and Biochemistry*. 1998, University of Texas at Austin: Austin, Texas. p. 172.
3. Senko, M.W., et al., *Electrospray Ionization FT-ICR Mass Spectrometry at 9.4 Tesla*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1824-1828.
4. Senko, M.W., et al., *External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 1997. **8**: p. 970-976.
5. Senko, M.W., et al., *A High-Performance Modular Data System for FT-ICR Mass Spectrometry*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1839-1844.
6. Zhang, Z., et al., *Human Recombinant [C22A] FK506-Binding Protein Amide Hydrogen Exchange Rates from Mass Spectrometry Match and Extend Those from NMR*. Protein Sci., 1997. **6**: p. 2203-2217.
7. Marshall, A.G. and C.L. Hendrickson, *Fourier transform ion cyclotron resonance detection: principles and experimental configurations*. Int. J. Mass Spectrom., 2002. **215**: p. 59-75.

Chapter 4. Data Reduction of Multiply Charged Ion Species Using an Automated Implementation of the THRASH Algorithm for the PC

4.1 Introduction

With the discovery of electrospray ionization (ESI) by Fenn [1-4], the use of mass spectrometry was changed forever. This “soft ionization” technique along with the discovery of matrix-assisted laser desorption ionization [5-7] created the field of biological mass spectrometry. ESI was able to ionize without fragmentation large biomolecules [3, 8-12]. Multiply charged ions generated by ESI extended the effective mass range of many mass spectrometers, perhaps most notably FT-ICR due to the ability to resolve very high charge states [13, 14]. It has also been noted in the literature that with increasing charge comes increasing complexity, and there have been many techniques to simplify or indeed automate the process [15-20]. As each new wave of spectral deconvolution is developed often with the benefit of past successes to guide the development, the complexity of such algorithms increases.

The current state-of-the-art in spectral deconvolution is arguably the thorough high resolution analysis of spectra by Horn (THRASH) [20]. The THRASH algorithm represents the culmination of individual tools developed for the identify either the charge state [18] or identify the most probable monoisotopic mass [21] as well as new features unique to THRASH. While THRASH is considered the best algorithm for the analysis of multiply charge ions, the use of the algorithm has been heavily constrained by two criteria. THRASH is programmed using PV-Wave (Visual Numerics, Houston, TX) that only operates under SunOS and the algorithm is only compatible with Odyssey FT-ICR data files (Finnigan, Madison, WI).

The work in this chapter describes the successful translation of the THRASH algorithm to the personal computer (PC). The PC-based THRASH is fully integrated into the modular ICR data acquisition and analysis system (MIDAS) developed at the National High Magnetic Field Laboratory (NHMFL, Tallahassee, FL). The advantages of the PC based algorithm include improved performance through the use of optimized libraries, improved performance based on increased process power and the source is based upon the industry standard C programming language. Additionally, the MIDAS program is openly distributed without cost to the members of the scientific community as well as supporting for multiple vendors ICR data file formats. Data is presented demonstrating the functionality of the THRASH algorithm. THRASH is also demonstrated through the analysis of electron capture dissociation (ECD) data.

ECD is a relatively new fragmentation technique with a unique ability to generate fragments essentially at every backbone amide linkage of a peptide [22-31]. ECD has been described as having a nonergodic fragmentation pathway. It has been shown that energy imparted into a positively charged ion upon capturing an electron, does not undergo the energy randomization that is typical of many common fragmentation techniques such as infrared multiphoton dissociation (IRMPD) or collisionally induced dissociation (CID) [28]. ECD has also been shown to selectively cleave the peptide backbone while leaving post-translational modifications such as glycosylation [24, 30] and phosphorylation [32] intact. Further studies are underway in which the effect of ECD on hydrogen scrambling is being investigated.

4.2 Methods

4.2.1 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

All experiments were performed using a custom built 9.4 T external electrospray source FT-ICR MS (National High Magnetic Field Laboratory, Tallahassee, FL) [33, 34]. This instrument employed a Chait type electrospray source with a 25 μ m i.d. fused silica spray needle. Ion desolvation was accomplished using a resistively heated capillary.

Ions exiting the capillary were focused into a 0.8 mm i.d. octopole rod set with the use of an electrostatic tube lens. Ions were then accumulated in a 60 cm long octopole by biasing the entrance and exit conductance limits to the octopole to a positive potential [34]. After an appropriate accumulation period of 2s to 60 s, the potential on the exit conductance limit was dropped to effect transfer of the ion through a second 2-m long octopole and into the analyzer cell of the FT-ICR instrument. The analyzer cell of this 9.4 T FT-ICR instrument is a 4 in diameter open cylindrical Penning ion trap.

Samples were directly infused into the electrospray source at a flow rate of 300 to 500 nl / min controlled by a Harvard Apparatus syringe pump (South Natick, MA). A needle potential of 2100 V was used to produce efficient analyte ionization. Prior to the ECD, the charge state of interest was isolated using a combination of chirp waveforms and a stored waveform inverse Fourier transform (SWIFT) waveform. Isolation of each charge state insured the identity of the precursor for each fragment ion.

ECD was accomplished with a 1 cm diameter dispenser cathode mounted approximately 13 cm away from the ICR cell. Typical ECD conditions involved the cathode biased at 0.1V with approximately 100 nA of emitted current [35].

Irradiation time for the ECD spectra presented was 51 ms with 40 ICR scans co-added to increase signal to noise ratio.

4.2.2 Thorough High Resolution Analysis of Spectra by Horn

The THRASH algorithm is an automated iterative process capable of isolating isotopic clusters, determining the charge state, and performing monoisotopic mass – zero charge calculation for every isotopic cluster within an ICR spectrum [20]. A novel peak subtraction allows the determination of overlapping isotopic clusters. The use of least-square analysis to compare experimental data to predetermined average isotope distributions yields the best-fit monoisotopic mass for clusters of sufficient mass that the actual monoisotopic peak is not observable in the spectra. An additional benefit of the THRASH algorithm is that fragment ions can be determined directly within the MIDAS analysis package if the consensus sequence and fragmentation type are known. Fragment ion can be determined for both ion dissociation fragments and proteolytic fragments.

Specific improvements described in this work to the THRASH algorithm focus on speed and portability. The primary hurdle to implementing THRASH on a PC was to port the code from PV-Wave to the C language. Once this initial step was performed, improvements to the THRASH algorithm could be facilitated through the use of commercial available import libraries that have been rigorously tested for speed and reliability. The Intel signal processing library version 4.5 (SPL) was one such library used. The Intel SPL provided fast Fourier transform (FFT) functions that are customized for each type of Intel processor. Since FFTs are used extensively in the computer gaming industry for three-dimensional visual effects, it is in Intel's best interest to continually tune such functions and provide them free of charge to programmers in order to maintain Intel's market share in the PC arena.

This market pressure allows the MIDAS to piggyback performance gains from Intel for which the FT-ICR community would never support alone.

Charge state determination in the THRASH algorithm is determined through the use of Fourier transform techniques [18, 20]. Figure 4.1 demonstrates the performance of the FFT function from the Intel SPL for various processors and packet sizes. Additional performance gains arise from the optimized memory allocation and array mathematical functions also present in the Intel SPL. Libraries other than the Intel SPL also increase THRASH performance including the replacement of user defined fit function for the commercially available polynomial fit and cubic spline interpolation functions from the analysis package of the CVI language (National Instruments, Austin, TX).

4.3 Results and Discussion

The user input and THRASH verification panel are presented as figure 4.2a and 4.2b respectively. The input commands are fairly straight forward and are similar to peak picking. Minimum and maximum charge state commands allow the user to optimize for the given data set. Peak level indicates minimum peak height to initiate a THRASH analysis; the units of peak level are in signal to noise. The description of the noise level in THRASH is unique in that noise is automatically determined for each 1 m/z window as THRASH moves through the spectrum. The noise level can be determined in the presence or absence of ion signal. THRASH determines by calculating the spectral density of the total m/z region. It is assumed that the region where the spectral density is the greatest is the noise level [18, 20]. The spectral density can be defined as a plot of the abundance versus the number of

Figure 4.1 Comparison of Fourier Transform Speed versus Data Set Size for Various Computer Processors

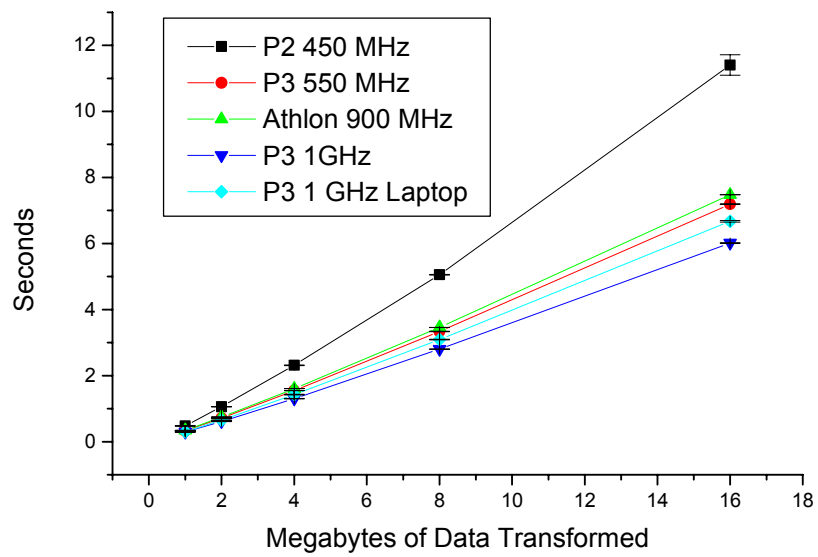
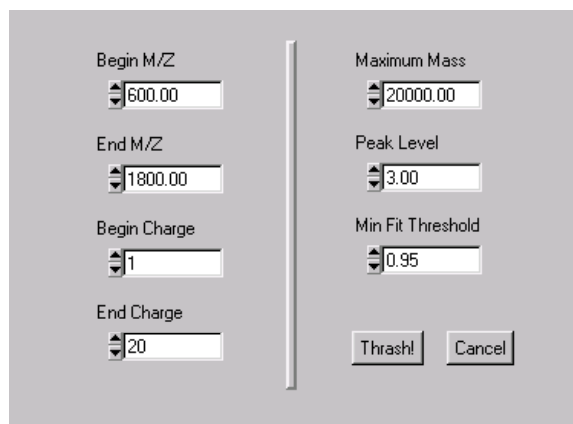


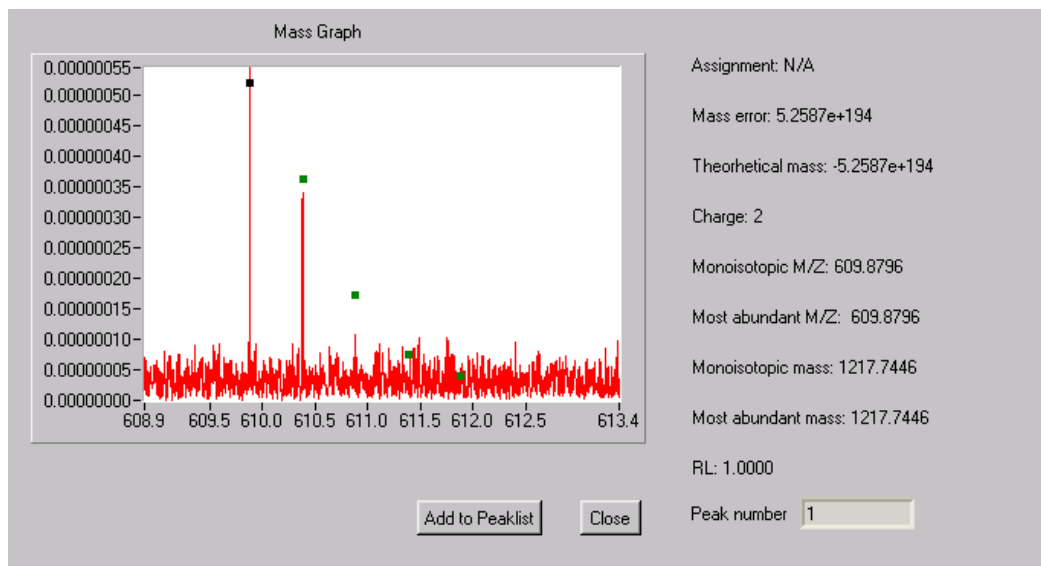
Figure 4.2 MIDAS Version of THRASH Algorithm in Action

Figure 4.2a. THRASH Input Panel



The THRASH Input Panel is a dialog box with a light gray background. It contains two columns of input fields, each with a vertical scrollbar. The left column includes: 'Begin M/Z' with a value of 600.00, 'End M/Z' with a value of 1800.00, 'Begin Charge' with a value of 1, and 'End Charge' with a value of 20. The right column includes: 'Maximum Mass' with a value of 20000.00, 'Peak Level' with a value of 3.00, and 'Min Fit Threshold' with a value of 0.95. At the bottom right, there are two buttons: 'Thrash!' and 'Cancel'.

Figure 4.2b. THRASH Progress and Verification Panel



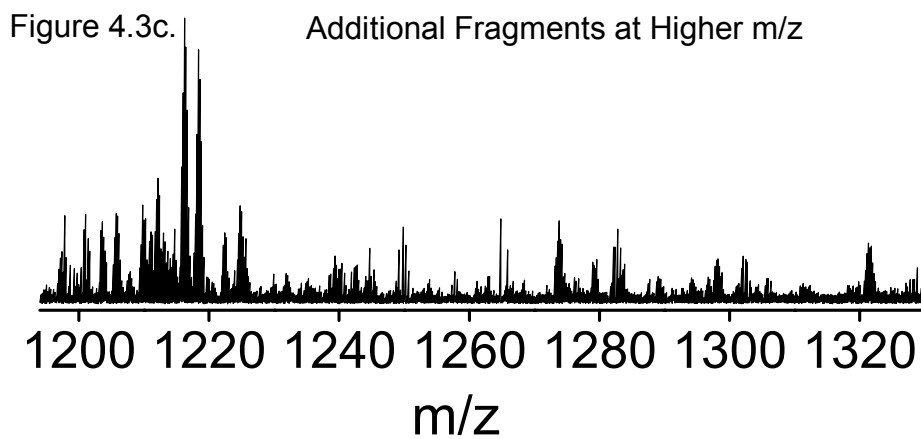
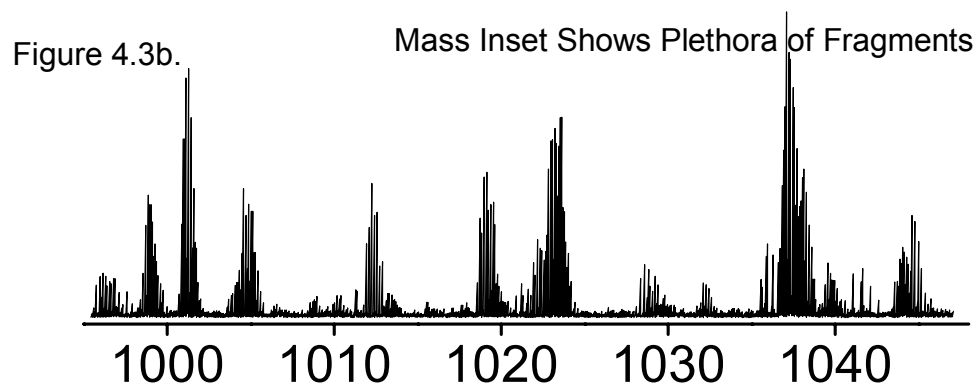
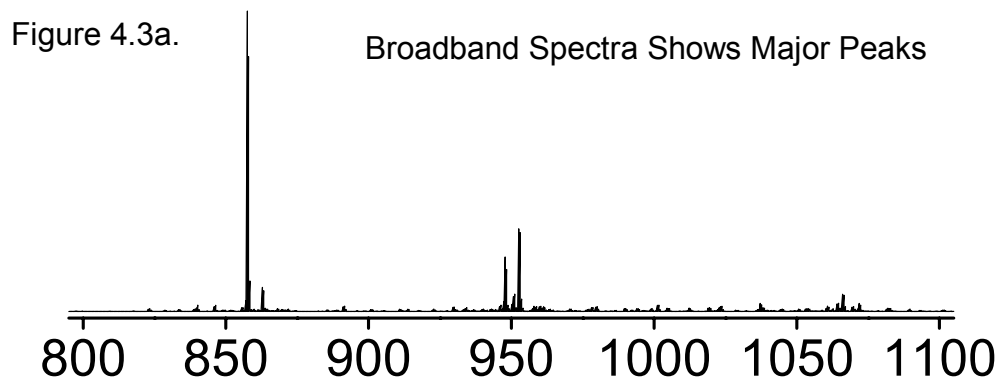
data points in the region that lie below this abundance. This method of noise determination is robust, requires no user input and requires little computational time.

The verification panel of THRASH in figure 4.2b presents the user with a tangible result for each THRASH assignment. The spectrum is overlaid with the isotope abundances that are predicted using a statistically average model system in this case of an average peptide. It is expected that most peptides, especially small peptides, will not contain the exact isotopic makeup; however, the isotopic structure of the statistically average peptide, averagine, will mimic the isotopic makeup of most peptides. The small squares above the isotope peak in figure indicate the averagine prediction for both relative abundance and mass for the given isotopic cluster. The color of the squares indicates the nature of the predicted peaks. A black indicates a monoisotopic peak while green squares indicate other peaks.

An ECD spectrum of ubiquitin $[M + 10H]^{10+}$ is presented in figure 4.3. Figure 4.3a shows a broadband region of the spectrum with the parent ion still apparent. Figure 4.3b and 4.3c demonstrate different mass insets of broadband spectrum; the complexity of the ECD experiment is apparent in these insets. While manual interpretation is possible for this type of experiment, it is time consuming. If one were to rely solely on manual interpretation, it would not be realistic to expect ECD to be performed routinely upon HPLC FT-ICR MS analysis.

Before one begins analysis of ECD spectra it is important to understand what types of fragment ions one should expect. ECD is reported to favor the cleavage of disulfide bonds and the C-terminal side of tryptophan while there is little to no

Figure 4.3 Sample Electron Capture Dissociation Experiment with Ubiquitin as Analyte



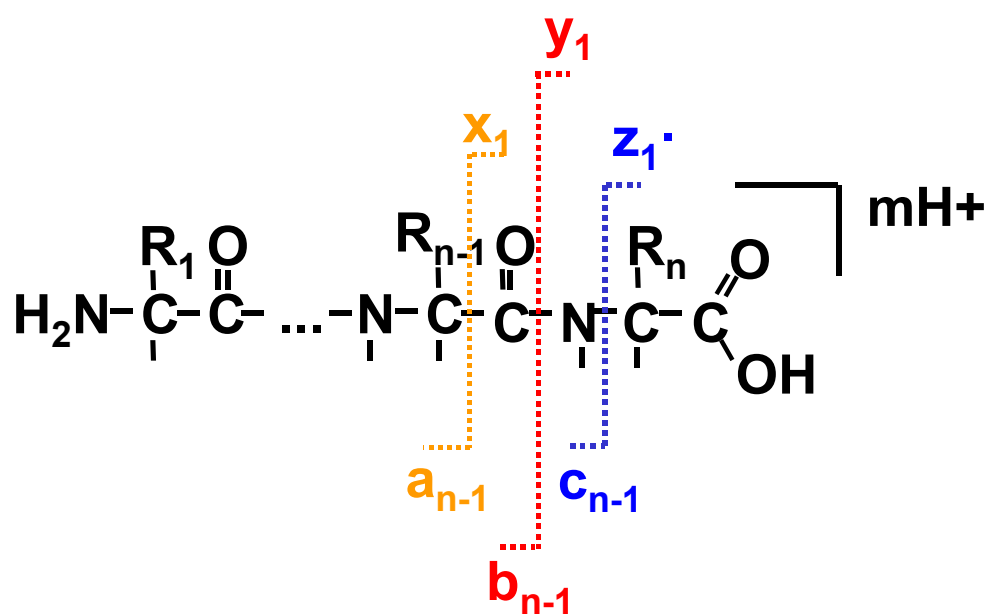
fragmentation seen on the N-terminal side of proline [25, 28]. ECD also creates different fragments than other ion dissociation events such as IRMPD or CID. ECD produces primarily c and z ions whereas IRMPD and CID produce mainly b and y ions. Figure 4.4 provides a general description of fragmentation along the backbone of a peptide.

Subjecting the ubiquitin ECD spectrum to THRASH yields the discovery of 222 isotopic clusters, of which 115 clusters were assigned as ECD fragments of ubiquitin. The total analysis time for the THRASH analysis of this spectrum was 162 seconds from a m/z range from 600 to 1600 if one compares this to the literature for a THRASH analysis isolated 282 isotopic peaks in m/z 550 –1800 in 19 minutes. Results from the THRASH analysis are shown in the user panel from the MIDAS analysis package in figure 4.5. Results from this panel can be saved in comma separated values (.csv) format and imported into commercial spreadsheet directly. Additional control allow the user to review individual results using the verification panel (figure 4.2b) and remove clusters which are not of interest.

Figure 4.6 recreates the original spectrum from figure 4.2b with the results from the automated THRASH analysis labeled above the clusters of interest. Several of these clusters overlap; a scenario that precludes from most types of automated analysis routines. The total sequence coverage of ubiquitin that was identified by THRASH is presented in figure 4.7. 66 out of 75 individual cleavages were identified using a 5 minute data acquisition and a 3 minute of analysis time.

While THRASH represent the most advanced automated deconvolution algorithm, the question remains “how good is THRASH?” THRASH is obviously not without faults. The most common concern with THRASH data fit involves the

Figure 4.4 Fragmentation Patterns of Peptides Emphasizing Electron Capture Dissociation (ECD)



ECD produces mainly c and z fragments but a and y fragments are not uncommon.

Figure 4.5 Automated Thrash Results from ECD Ubiquitin Spectra shown in Figure 4.3.

	Charge	Mono Mass	Abun Mass	Mono M/Z	Abun M/Z	Assignment	Mass Diff	RL
200	4	3103.0412	3112.0030	1276.4170	1278.1730	c - 43	-0.137330	0.9992
209	2	2562.3215	2563.3288	1282.1680	1282.6717	c - 23	-0.076298	1.0000
210	5	6481.1808	6485.2098	1297.2434	1298.0492	c - 58	0.723074	1.0000
211	3	3900.9739	3902.9884	1301.3319	1302.0034	None	N/A	0.9848
212	5	6597.3170	6601.3461	1320.4707	1321.2765	None	N/A	1.0000
213	5	6643.2806	6647.3097	1329.6634	1330.4692	c - 59	-0.240435	1.0000
214	2	2691.3438	2692.3511	1346.6792	1347.1828	c - 24	-0.09663	1.0000
215	6	8240.1961	8245.2324	1374.3733	1375.2127	None	N/A	0.9995
216	6	8286.2518	8291.2881	1382.0492	1382.8886	None	N/A	0.9998
217	5	6952.4621	6956.4912	1391.4997	1392.3055	None	N/A	1.0000
218	3	4209.1663	4211.1808	1404.0627	1404.7342	c - 38	-0.130716	0.9921
219	6	8457.2361	8462.2725	1410.5466	1411.3860	a - 75	-0.361313	0.9997
220	6	8471.3172	8476.3536	1412.8935	1413.7329	None	N/A	0.9994
221	6	8501.3062	8506.3425	1417.8916	1418.7310	c - 75	-0.304925	0.9999
222	6	8517.2264	8522.2628	1420.5450	1421.3844	None	N/A	0.9998

Figure 4.6 Ubiquitin Spectra with THRASH Results Designated
(Spectra From Figure 4.3b.)

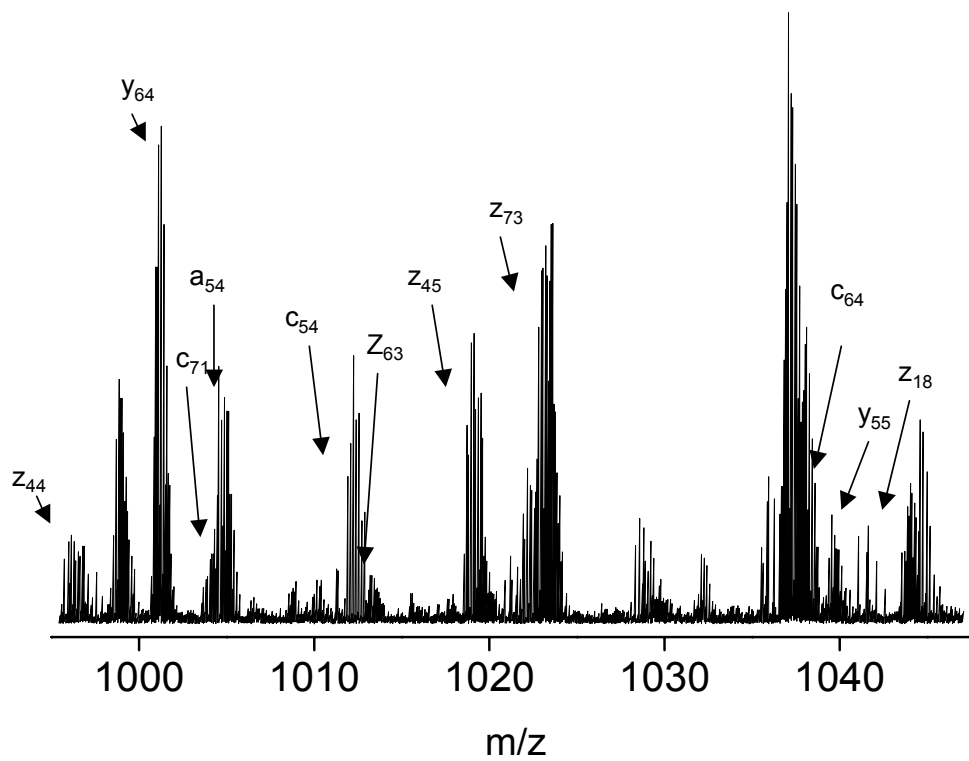


Figure 4.7 Sequence coverage of Ubiquitin by Automatic THRASH Assignment from ECD Spectra.

└ A cleavage
 └ C cleavage
 └ Y cleavage
 └ Z cleavage



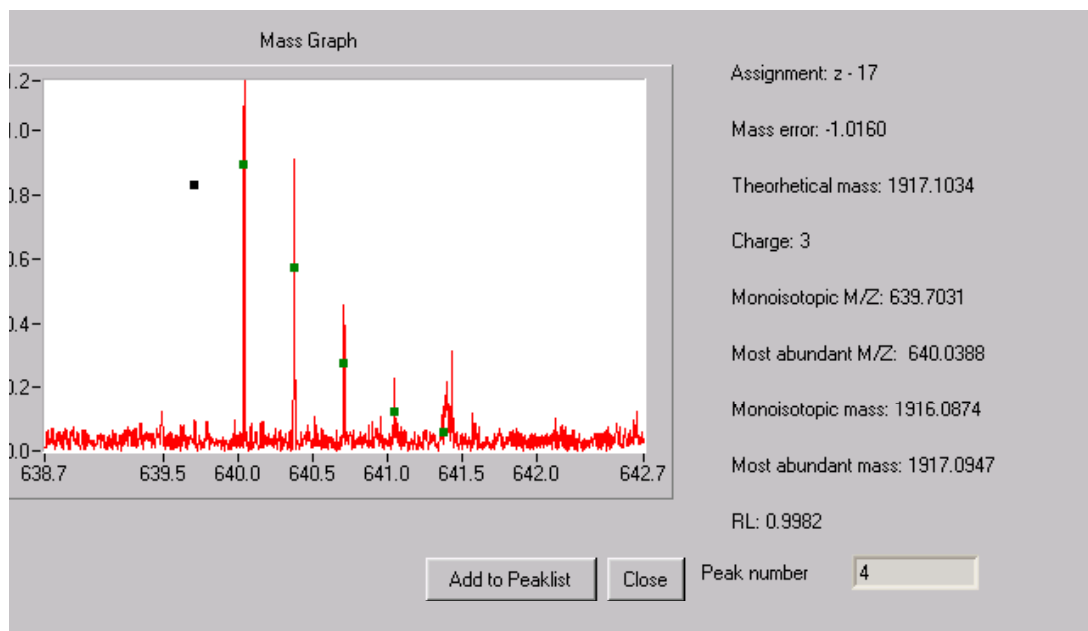
correct assignment of the monoisotopic peak. This is to be expected since this also the most common problem in the field of FT-ICR for the assignment of accurate mass. The problem arises when the abundances present in the ICR cell vary slightly from the predicted isotopic abundances. This leads to the mis-assignment of the most abundant peak and a mass shift 1 Da occurs in the calculated monoisotopic mass. THRASH excels at the assignment of highly charged larger species where the user is often fooled when using the “Chi by eye” approach. It is the lower charge states where the statistics are lower where THRASH can be tricked. User verification will often recognize the mis-assigned peak readily. Figure 4.8 demonstrates a mis-assigned cluster under THRASH. It is readily apparent that the black square representing the theoretical peak monoisotopic is located 0.33 m/z units low for this $[M + 3H]^{3+}$ species. The mass error for the assignment in the upper right corner of the panel as a z^{17} fragment is also off by approximately 1 Da.

4.4 Conclusions

THRASH is a important tool in the automated analysis of multiplied charged ions. The work presented to port THRASH to the PC has the advantages of speed and portability. The addition of THRASH to the MIDAS analysis allows users of several commercial systems other than the Finnigan Odyssey data station (that is no longer produced). Automated data analysis will soon be an important process in the data acquisition as well. Data dependent acquisition will increasingly require more sophisticated tools to be effective. A C language version of THRASH is readily implemented within the data analysis module of MIDAS.

Additional benefits afforded by ECD in its application to H/D exchange are contingent upon a better understanding of the effect that ECD fragmentation has

Figure 4.8 Demonstration of THRASH Isotope Matching where the Measured Peptide Distribution is Incorrectly Assigned by a Mass Unit



with respect to H/D scrambling. If ECD can be shown to promote fragmentation in the absence of scrambling, it will prove a powerful tool for future H/D exchange studies.

4.5 References

1. Yamashita, M. and J.B. Fenn, *Electrospray Ion Source. Another Variation on the Free-Jet Theme*. J. Chem. Phys., 1984. **88**: p. 4451-4459.
2. Whitehouse, C.M., et al., *Electrospray*. Anal. Chem., 1985. **57**: p. 675-679.
3. Fenn, J.B., et al., *Electrospray Ionization for Mass Spectrometry of Large Biomolecules*. Science, 1989. **246**: p. 64-71.
4. Fenn, J.B., et al., *Electrospray*. Mass Spectrom. Rev., 1990. **9**: p. 37-70.
5. Tanaka, K., et al., *Protein and Polymer Analyses up to m/z 100000 by Laser Ionization Time-of-Flight Mass Spectrometry*. Rapid Commun. Mass Spectrom., 1988. **2**: p. 151-153.
6. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Anal. Chem., 1988. **60**(20): p. 2299-2301.
7. Karas, M., U. Bahr, and F. Hillenkamp, *UV Laser Matrix Desorption/Ionization Mass Spectrometry of Proteins in the 100000 Dalton Range*. Int. J. Mass Spectrom. Ion Proc., 1989. **92**: p. 231-242.
8. Henry, K.D., et al., *Fourier-Transform Mass Spectrometry of Large Molecules by Electrospray Ionization*. Proc. Natl. Acad. Sci. U.S.A., 1989. **86**: p. 9075-9078.
9. Henry, K.D., J.P. Quinn, and F.W. McLafferty, *High-Resolution Electrospray Mass Spectra of Large Molecules*. J. Amer. Chem. Soc., 1991. **113**: p. 5447-5449.
10. Nohmi, T. and J.B. Fenn, *Electrospray Mass Spectrometry of poly(ethylene glycols) with molecular weights up to five million*. J. Am. Chem. Soc., 1992. **114**(9): p. 3241.

11. Beu, S.C., et al., *Improved FTICRMS of Large Biomolecules*. J. Amer. Soc. Mass Spectrom., 1993. **4**: p. 190-192.
12. Beu, S.C., et al., *FT Electrospray Instrumentation for Tandem High-Resolution MS of Large Molecules*. J. Amer. Soc. Mass Spectrom., 1993. **4**: p. 557-565.
13. Kelleher, N.L., et al., *Unit Resolution Mass Spectra of 112 kDa Molecules with 3 Da Accuracy*. J. Am. Soc. Mass Spectrom., 1997. **8**: p. 380-383.
14. Shi, S.D.-H., C.L. Hendrickson, and A.G. Marshall, *Counting Individual Sulfur Atoms in a Protein by Ultrahigh-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Experimental Resolution of Isotopic Fine Structure in Proteins*. Proc. Natl. Acad. Sci. U.S.A., 1998. **95**: p. 11532-11537.
15. Mann, M., C.K. Meng, and J.B. Fenn, *Interpreting Mass Spectra of Multiply Charged Ions*. Anal. Chem., 1989. 61: p. 1702-1708.
16. Ferrige, A.G., M.J. Seddon, and S. Jarvis, *Maximum entropy deconvolution in electrospray mass spectrometry*. Rapid Commun. Mass Spectrom., 1991. **5**: p. 374-379.
17. Senko, M.W., S.C. Beu, and F.W. McLafferty, *Mass and charge assignment for electrospray ions by cation adduction*. J. Amer. Soc. Mass Spectrom., 1993. **4**: p. 828-830.
18. Senko, M.W., S.C. Beu, and F.W. McLafferty, *Automated Assignment of Charge States from Resolved Isotopic Peaks for Multiply Charged Ions*. J. Am. Soc. Mass Spectrom., 1995. **6**: p. 52-56.
19. Zhang, Z. and A.G. Marshall, *A Universal Algorithm for Fast and Automated Charge State Deconvolution of Electrospray Mass-to-Charge Ratio Spectra*. J. Am. Soc. Mass Spectrom., 1998. **9**: p. 225-233.

20. Horn, D.M., R.A. Zubarev, and F.W. McLafferty, *Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules*. J. Am. Soc. Mass Spectrom., 2000. **11**: p. 320-332.
21. Senko, M.W., S.C. Beu, and F.W. McLafferty, *Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions*. J. Am. Soc. Mass Spectrom., 1995. **6**: p. 229-233.
22. Zubarev, R.A., N.L. Kelleher, and F.W. McLafferty, *Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process*. J. Am. Chem. Soc., 1998. **120**: p. 3265-3266.
23. Axelsson, J., et al., *Electron Capture Dissociation of Substance P using a Commercially Available FT-ICR Mass Spectrometer*. Rapid Commun. Mass Spectrom., 1999. **13**: p. 474-477.
24. Mirgorodskaya, E., P. Roepstorff, and R. Zubarev, *(ECD for sequencing of O-glycosylated peptides)*. Anal. Chem., 1999. **71**: p. 4431-4436.
25. Zubarev, R.A., et al., *Electron Capture Dissociation of Gaseous Multiply-charged Proteins is Favored at Disulfide Bonds and Other Sites of High Hydrogen Atom Affinity*. J. Am. Chem. Soc., 1999. **121**: p. 2857-2862.
26. Horn, D.M., Y. Ge, and F.W. McLafferty, *Activated Ion Electron Capture Dissociation for Mass Spectral Sequencing of Larger (42 kDa) Proteins*. Anal. Chem., 2000. **72**(4778-4784).
27. Zubarev, R.A., et al., *Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations*. Anal. Chem., 2000. **72**: p. 563-573.
28. McLafferty, F.W., et al., *Electron Capture Dissociation of Gaseous Multiply Charged Ions by Fourier-Transform Ion Cyclotron Resonance*. J. Am. Soc. Mass Spectrom., 2001. **12**: p. 245-249.

29. Shi, S.D.-H., et al., *Phosphopeptide/Photphoprotein Mapping by Electron Capture Dissociation Mass Spectrometry*. Anal. Chem., 2001. **73**: p. 19-22.
30. Hakansson, K., et al., *Electron Capture Dissociation and Infrared Multiphoton Dissociation MS/MS of an N-Glycosylated Tryptic Peptide to Yield Complementary Sequence Information*,. Anal. Chem., 2001. **73**(18): p. 4530-4536.
31. Hakansson, K., et al., *High Sensitivity Electron Capture Dissociation Tandem FTICR Mass Spectrometry of Microelectrosprayed Peptides*. Anal. Chem., 2001. **73**(15): p. 3605-3610.
32. Chalmers, M.J., et al., *Liquid Chromatography - Fourier Transform Ion Cyclotron Resonance Mass Spectrometric Characterization of Protein Kinase C Phosphorylation*. Submitted to J. Proteome Research, 2003.
33. Senko, M.W., et al., *Electrospray Ionization FT-ICR Mass Spectrometry at 9.4 Tesla*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1824-1828.
34. Senko, M.W., et al., *External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 1997. **8**: p. 970-976.
35. Quinn, J.P., et al. *Rapid Electron Capture Dissociation with Simultaneous Laser Access for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. in *50th American Society for Mass Spectrometry & Allied Topics*. 2002. Orlando, FL, USA.

Chapter 5. Data Reduction in the Determination of Solution Based H/D Exchange for the Determination Protein: Protein Interaction Sites

5.1 Introduction

Gas phase H/D exchange has always been challenged to prove a relationship between gas phase and biological conditions and conformations [1]. Consequently, there is an advantage to performing H/D exchange in "biological" conditions, followed by a peptide digest, and then coupling the experiment to an FT-ICR mass spectrometer. An additional benefit of liquid phase H/D exchange is increased signal at long exchange periods relative to gas phase exchange studies and unambiguous deuterium labeling. Since the energy required for in cell ion dissociation is absent, the energy that induces deuterium scrambling in the fragment ion is also absent [2-5]. A final detriment to gas phase H/D exchange is that such studies subject the ICR cell to extended periods of relatively high pressures of the exchange gas. An afternoon of gas phase exchange experiments can lead to degraded instrument performance for multiple days and contributes to increased instrument "down-time" and maintenance.

Studies linking liquid-phase exchange to mass spectrometry have their own problems, primarily related to the matrix effects on the mass spectrum and the occurrence of H/D back exchange. However, literature cites many successful applications involving liquid phase exchange and mass spectrometry [6-10]. The work described in this chapter employs two distinct strategies to effectively perform liquid phase exchange via FT-ICR MS. To combat matrix effects, HPLC was used to concentrate and desalt the sample prior to injection to the mass spectrometer. The complex nature of deuterium back exchange poses a non-trivial problem. Factors that affect back exchange can include solvent effects, temperature, and primarily

time. To minimize these effects, a protocol developed at the NHMFL describes a method to combat back exchange such that at the end of the liquid phase H/D exchange period that reaction is quenched by lowering the pH of the sample as well as performing the remainder of the experiment in an ice bath with the pH of the solution maintained at ~2. The sample undergoes peptic digestion for a short (~2 minutes) time period 0°C injection loop. After the digestion the sample is injected onto a reverse phase column, desalted and eluted into a micro-ESI source for introduction to the FT-ICR instrument.

While the collection of such data is by no means trivial, the resulting data required months to analyze. The results generated during a single HPLC FT-ICR experiment that lasts ten minutes can occupy over a gigabyte of disk space. H/D exchange studies require triplicate measurement for each of multiple H/D exchange periods. The primary study described in this chapter involved 2250 independent data files. The initial investigation took more than 2 months for two graduate students to complete. This investigation was performed using the center of mass technique and HD Helper tool described in a previous chapter. It was readily apparent that a more efficient method was needed before such studies could be considered practical.

Additional work described in this chapter involves the development of an automated analysis protocol that can effectively and routinely reduce the total analysis time for a liquid phase H/D exchange HPLC FT-ICR MS study. The analysis protocol reduced the total analysis time for a complete H/D exchange study by liquid- phase HPLC FT-ICR MS from a period of approximately eight weeks to an analysis period of a week. The accuracy of the results is also increased since the algorithm does not require the user to track the analysis through a series of data transfer and collation stages. Assumptions made for this analysis are described in detail and are consistent with traditional FT-ICR data analysis.

5.2 Methods

5.2.1 Sample preparation and H/D exchange

Recombinant HIV-1 capsid CA protein was obtained from Peter Prevelidge et al [11]. The protein was expressed in *Escherichia coli* and precipitated in ammonium sulfate. After purification by ion exchange chromatography, the resultant solution was aliquoted and the salt content of one aliquot was raised to 1.0 M NaCl to initiate assembly of CA. Assembly continued at ambient temperature for 30 minutes. The monomer and assembled CA were further aliquoted to perform triplicate analysis by H/D exchange.

Additional samples of fibrillarin and fibrillarin-Nop5 complex were expressed in *E.Coli*. The samples were purified by metal affinity chromatography and dialysis. Approximately 6 µg of sample were aliquot and were analyzed in triplicate H/D exchange experiments.

Several stock solutions were prepared to facilitate the H/D exchange procedure. Stock CA solutions were prepared with 500 µM purified monomer or hexamer in 50 mM Na₂HPO₄ at pH 8.0. A separate stock solution was prepared for each of the triplicate runs. H/D exchange solution was created containing 25mM ammonium acetate in deuterium oxide for a pH of 6.8. A digest solution of 80 µM porcine pepsin (Roche) was prepared in 0.5% formic acid (FA). A quench solution for the H/D exchange reaction was prepared as a 0.5% FA solution at pH 2.1. The use of gradient flow elution on the HPLC requires two solutions: solution A and solution B. Solution A contains 10 % acetonitrile, 89.5% water and 0.5 % FA. Solution B contains 90 % acetonitrile, 9.5 % water and 0.5 % FA.

H/D exchange was initiated as a 2 μL aliquot of stock CA solution was diluted into 18 μL of the H/D exchange solution. Hexamer CA samples contained 2.0 M NaCl to stabilize the assembled polymers. After exchange period ranging from 0 to 4080 minutes, the exchange was stopped by the addition of 20 μL of the quench solution. Each of the triplicate samples were vortexed and frozen in preparation for LC-MS analysis. Zero and blank controls were performed, and the 4080-minute exchange point was described as the 100% control. Further details of the sample preparation and exchange conditions are described in the literature [12].

5.2.2 Protein Digestion and HPLC Conditions

HPLC experiments were performed on a Shimadzu HPLC equipped with a 5 μL manual injector and Vydac microbore C_4 column (1.0 x 50 mm). The injector and column were submerged in a 0° C ice bath throughout the experiment. 20 μL of the digest solution was added to the CA protein solutions. The solutions were immediately mixed and then 5 μL was injected onto the pre chilled manual injector. The digestion was allowed to continue for 2 minutes while in the loop. The digest was then injected onto the chilled column and eluted by gradient HPLC. After injection, the peptides are desalted for 2 minutes on-column using 95% solution A and 5 % solution B. A 4-minute gradient flow elution was used to separate the proteolysis products prior to introduction to the FT-ICR MS instrument. Flow rate was reduced from 50 $\mu\text{L}/\text{min}$ to ~ 400 nL/min using a post column split-flow interface. The gradient was optimized to minimize the back exchange that occurs on-column and as such provides only marginal chromatographic separation. The lack of chromatographic resolution is easily overcome by the high resolution of the FT-ICR MS. The HPLC conditions and procedure is described in more detail elsewhere [12].

5.2.3 Electrospray ionization FT-ICR MS

All experiments were performed on the 9.4 T FT- ICR instrument at the NHMFL in Tallahassee, FL [13, 14]. A 50- μm fused silica micro-ESI needle was directly coupled to a Microtee flow rate splitter (Upchurch) with the electrical connection to the needle applied through the metal waste line of the splitter. The source employed a heated metal capillary for desolvation. The ions were guided through an accumulation octopole and into a linear octopole trap. The analyte ions were accumulated for 2 s before transmission to the 4-inch open cylindrical ICR cell. Swept frequency excitation was followed by the acquisition of a 512 k data point 0.8 s time domain transient. FT-ICR control and data acquisition was performed using the Modular ICR Data Acquisition System (MIDAS) [15, 16]. A total of 50 ICR spectra were collected for each LC run, but only a few of these spectra contained peptide fragments. External mass calibration was performed on the blank control and the mass error was determined to be less than 20-ppm.

5.2.4 Proteolysis

The determination of sites of biological interaction is accomplished by comparing the H/D exchange rates of the individual components to the H/D exchange of the assembled unit. This requires that proteolysis occur in conditions that limit back-exchange of the products. Another condition for successful proteolysis is that the resolution of the exchange data is proportional to the number of fragments observed. These conditions make porcine pepsin ideal for this experiment. Pepsin preferentially cleaves on the C-terminal side of F, L, E, W, Y and I with again cleavages occurring at various rates. Porcine pepsin is especially

suited for this study because it exhibits optimal activity at pH 2.1 and 0° C [17-19] where back exchange is minimized.

5.3 Results and Discussion

5.3.1 Automated Batch Analysis Algorithm

Primary data analysis is performed using MIDAS analysis with subsequent plots generated in Excel (Microsoft). Manual H/D exchange is calculated for each file in the HPLC run using center of mass calculation as described earlier. After each individual center of mass value is calculated the overall center of mass is calculated in an Excel spreadsheet as the average value of the individual files. This calculation is repeated for the additional 2 replicates. Replicate data is combined into an Excel spreadsheet for the average center of mass calculation and the subsequent standard deviation for a single H/D exchange time period. This process is then repeated for each subsequent exchange period.

For the case of the HIV-1 capsid CA protein described throughout much of this work, manual H/D exchange determination was an involved and tedious process. The user is required to perform as many as 2250 center of mass calculations, manually enter these values into a spreadsheet, and then manually transfer the resulting 2310 averages into additional spreadsheets. Having completed the calculation for one proteolysis product, the user can now start on proteolysis fragment #2.

Acknowledging that the manual method of H/D exchange calculation was prohibitively user dependent, the author developed an automated analysis procedure. This automated analysis is accomplished for a single fragment in 30

minutes. Each additional fragment is determined in less than 20 seconds. The disparity between the first and subsequent scans arises from the peak-picking algorithm that must be implemented as part of the initial analysis. The peak picking involves unzipping the data, processing the time domain data to the mass domain and the actual peak picking of the mass data. The peak list is stored in a small text file with the same name as the original file (different extension.) Subsequent analyses need only to interrogate the peak list files resulting in the dramatic speed increases.

Figure 5.1 demonstrates the MIDAS user interface for the automated analysis of H/D exchange data. The top section of the panel in figure 5.1 demonstrates parameters specific to H/D exchange targets, the information in this section are required for the analysis to proceed. The user must have knowledge of the peptide fragments that are expected to undergo exchange. The knowledge can be obtained by THRASH analysis of the control scans, by conventional ICR analysis of the control or the user can use predicted fragment masses given the use of pepsin and the analyte of interest. The user inputs the expected mass range for the species of interest into the “minimum m/z” and “maximum m/z”. This mass range must be chosen such that both the initial isotope of the zero control and the final isotope of the 100% control are within the selected m/z range.

The controls for “Selected Scans only”, “minimum scan” and “maximum scan” allow the user to limit the search to a limited fraction of the chromatographic run. Limiting the scan number is important if there are a large number of leading or trailing scans in an LC experiment that describe only the FT-ICR response to the LC solvents and not peptides of interest. The “Mass Error” in section 5.1a describes the maximum deviation that is allowable between isotopes of the same analyte. Further

Figure 5.1 MIDAS Control for Automatic Batch Processing of HPLC FT-ICR H/D Exchange Experiments

Selected Scans only <input checked="" type="checkbox"/>	minimum scan 7	maximum scan 10
Enable Table Lookup <input type="checkbox"/>	minimum charge 1	maximum charge 9
Mass Error (%) 2.00	minimum m/z 100.00	maximum m/z 200.00

HD Directory
c:\Documents and Settings\blakney\Desktop\30 minute exchange_008

<input checked="" type="checkbox"/> Peak Picking	<input checked="" type="checkbox"/> Baseline Correct
Truncations 0	Zerofills 1
Apodization Hanning	
Scale Factor 1.00E+0	
Start search from m/z: 500	<input checked="" type="checkbox"/> Silent
End search at m/z: 2000	
Minimum Peak Width 3	
Threshold (%) 2.000000	<input type="checkbox"/> Relative
Noise Threshold (%) 1.000000	<input checked="" type="checkbox"/> Absolute

<input checked="" type="checkbox"/> Recalibrate	A) 1.4426233173E+8
	B) -7.9965209428E+8

Auto Save Result
Save Filename
c:\Documents and Settings\blakney\Desktop\8 minute.csv

Peak Picking Progress

0 10 20 30 40 50 60 70 80 90 100

discussion of “Mass Error” will be provided as the selection algorithm is described in detail. Controls indicated by “minimum charge” and “maximum charge” may seem odd since the charge of the analyte must be known to the user prior to selecting the proper m/z for the analyte. The minimum and maximum charge provide an important role in the selection algorithm. The last control is “Enable Table Lookup” that disables all of the parameters of this section of this panel and opens the additional user panel shown in Figure 5.2 for multiple fragment target analysis in a single analytical session. The use of the additional panel helps manage display space during analysis.

The second section of figure 5.1 provides for the selection and displays the directory path to the H/D exchange data. This path may describe the path to the raw data files or to the raw data files plus the peak list files. It is important the directory and all subdirectory not be write-protected if peak picking is to be performed since the peak list files will be written to the same directory. In order to automate the file selection process, certain rules must apply to the directory structure where the data is stored. The primary folder must contain separate folders for each H/D exchange period. The folders are sorted by the last four characters of the folder name. The decision to use the last four characters allows one to easily retroactively rename a folder with a minimum of confusion. Within each exchange period folder, there must be three folders. Each of these folders must contain one of the three replicate runs. The replicate folders are also sorted by the last four characters of the folder’s name. Within each of the replicate folders there must be numerous LC FT-ICR files and may also contain the peak list files. Peak list files are required for the subsequent analysis and can be created either prior to the batch H/D exchange process or a peak-picking algorithm can be invoked as the first step of the batch H/D process that will automatically generate the necessary peak list files.

Figure 5.2 Auxiliary Control for the Automatic Batch Processing of Multiple Targets for HPLC FT-ICR MS H/D Exchange Experiments

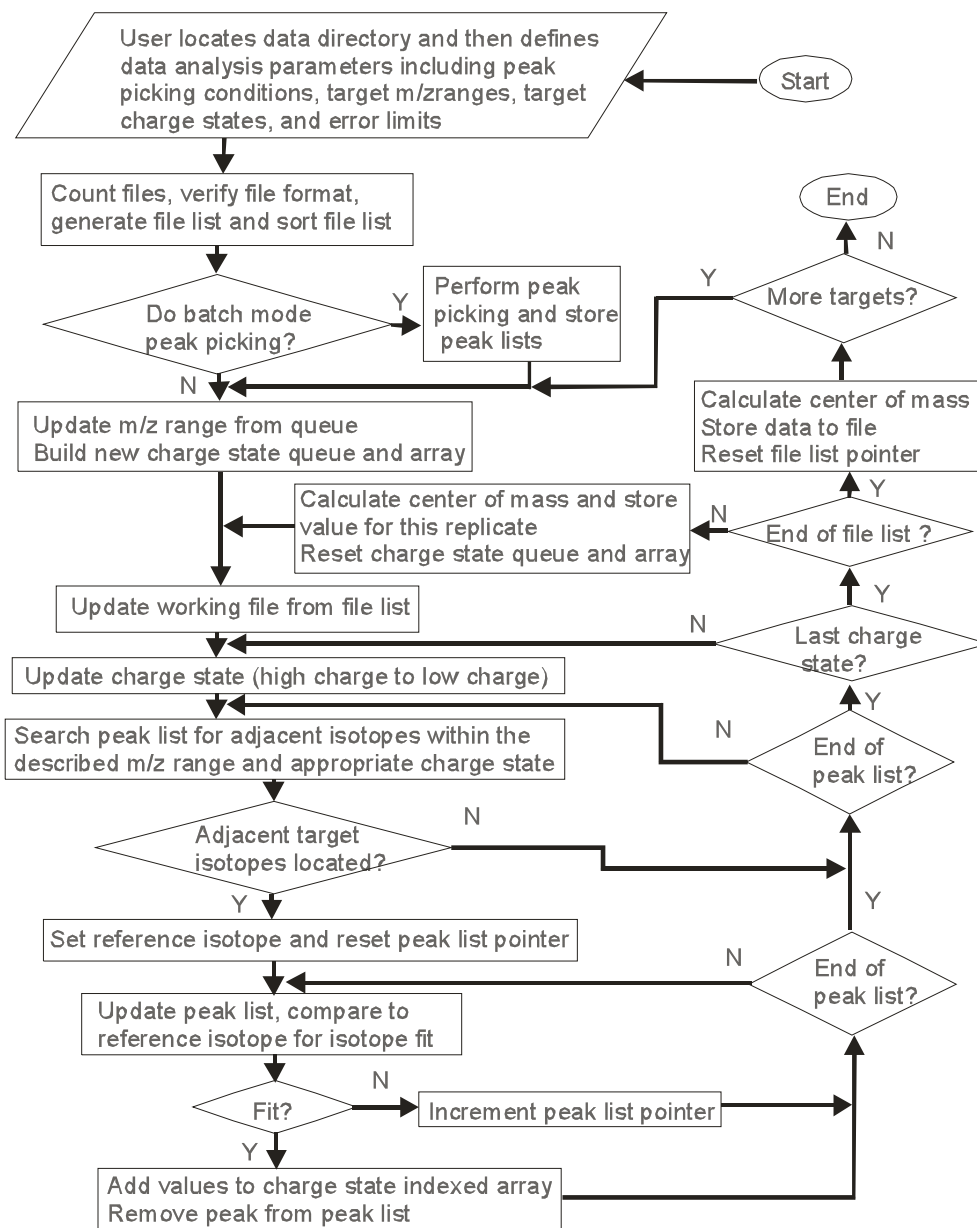
Check to Select Item(s) to Delete	Max m/z	Min m/z	Max Z	Min Z	Tolerance (%)
<input type="checkbox"/>	518.0000	512.0000	9	1	2.00
<input type="checkbox"/>	551.0000	542.0000	9	1	2.00
<input type="checkbox"/>	605.0000	602.0000	9	1	2.00
<input type="checkbox"/>	625.0000	621.0000	9	1	2.00
<input type="checkbox"/>	712.0000	708.0000	9	1	2.00
<input type="checkbox"/>	725.0000	722.0000	9	1	2.00

The batch-processing algorithm consists of a relatively large overhead mechanism that serves to manage, track and organize the files and a smaller mathematical unit that locates and isolates the H/D exchange data and performs the center of mass calculations. The general flow of the algorithm is presented in figure 5.3. For this flow chart, some standard shapes (ANSI) are used to indicate processes within the chart. Ovals indicate beginning and ending points. Parallelograms indicate user input. Rectangles indicate an operation or computation. Finally, diamonds represent decision-making steps or branches.

As can be seen from the diagram, the batch analysis process is highly iterative. One of the advantages of this iterative nature is that the center of mass calculation is only calculated once for each target for each entire HPLC run. This method allows true abundance weighting for every qualified peak within the study. Methods that calculated individual center of mass calculations for each spectrum are combined to produce an average center of mass for the HPLC run. Since each spectrum's center of mass is considered equal when calculating such an average, the contribution of spectra with lower total abundance will be disproportionately large.

A second advantage afforded by the iterative nature of the algorithm is that it allows identified peaks to be selectively removed from the local copy of the peak list. Multiple targets may be overlapped within the same target m/z region. Removing the selected peaks and then reprocessing the remaining peak list can produce additional results more quickly as the peak list is reduced. In the scenario, the identified peaks for the current charge state are removed. The next iteration uses the reduced peak list for the next charge state. The algorithm was designed to work from higher charge states to lower charge states. Since the higher charge state

Figure 5.3 Flowchart Representing the Batch H/D Exchange Algorithm



isotope peaks have already been removed from the peak list, non-sequential isotopes from higher charge state can no longer interfere with lower charge state determination. To better understand this potential interference, one should consider a case where the user expects a target that is singly charged. The isotope spacing for this analyte is approx 1 m/z unit. If there were a doubly charged species that fell within the same target range and approximate LC scan number, the first and third isotope of the doubly charge ion would have the same mass spacing as a singly charged species. Similar scenarios could describe for any charge state n and interference by the integer multiples of n charge state. By setting a modest target range above the charge of interest, higher charged interfering species are removed prior to peak picking of the primary target.

Typical results from the batch H/D exchange algorithm are presented in figure 5.4. These results were saved to a comma separated value (.csv) file that can be easily transported into Microsoft Excel or Microcal Origin for further analysis. These results represent the analysis of one target m/z range for HIV-1 CA capsid protein. The analysis of these 1200 files was completed in less than five minutes on a Pentium 4 computer. Analysis time does not include the generation of the peak list files. The results for charge states 9+ through 1+ were calculated for each of 8 exchange periods. The primary target was the 2+ charge state which appears in the lower part of the panel. The 2+ charge state is also the only charge state where qualifying peaks were selected for each trial of each exchange period. For any trial where for which there were no qualifying peaks a center of mass of 0 was recorded for the trial and the standard deviation (labeled SD) registered a value of -1. The average HDX was calculated for any exchange period with at least one qualified trial. The column labeled "Time" is not currently active other than a place holder for the user to enter the exchange time by hand in Excel; however, a planned enhancement will allow the user to designate this value during data acquisition.

In order to verify that the batch algorithm performed correctly, data from a H/D exchange study was processed and compared to the manually calculated results. This comparative study is presented in figure 5.5. The manual calculations tabulate a center of mass for each scan and then combine these individual centers of mass into an average center of mass. For the batch algorithm, only one center of mass calculation is performed per trial with each the abundance of each peak from each LC scan represented equally. The difference in calculation techniques would naturally lead to slightly different values even on identical data. The graph in figure 5.5b showed excellent agreement between the manual and automated algorithms.

5.3.2 Application of Automated Analysis

Results from the automated exchange analysis of an active biological interaction can be directly compared to those in an inactive state. Positive and negative differences in exchange patterns indicate sites of interaction. The interaction of Nop5 and fibrillarin is investigated in figure 5.6. Figure 5.6a and 5.6b indicate similar exchange patterns for the fibrillarin monomer and the assembled complex. With the exchange rates so similar, it was apparent that the fragments "VKARSIDSTAEPEEVF" and "KIVKHGSLMPYHRDHIFIHA" did not participate in the interaction of the fibrillarin in the complex. On the other hand, there was a clearly an interaction for the fragments "YLGAASGTTVSHLADI" and "VEKVDLIYQDIAQK" evidenced in figures 5.6c and 5.6d.

Figure 5.7 demonstrates the ultimate goal for the use of liquid phase exchange FT-ICR MS. A complete comparison of individual fragment reactivity

Figure 5.5 Comparison of Automated and Manual Calculation of H/D Exchange using Center of Mass

Figure 5.5a

Exchange time	Automated		Manual	
	Center of Mass	Standard Deviation	Center of Mass	Standard Deviation
0	2401.038	0.13	2401.092	0.16
0.5 min	2402.368	0.03	2402.544	0.03
1 min	2402.441	0.05	2402.510	0.06
2 min	2402.630	0.07	2402.687	0.04
4.5 min	2402.835	0.04	2402.863	0.05
6 min	2402.957	0.02	2403.012	0.01
8 min	2402.918	0.01	2402.983	0.01
10 min	2403.056	0.02	2403.075	0.02

Figure 5.5b
Average
Center of Mass

Peptide Fragment From Fibrillarlin

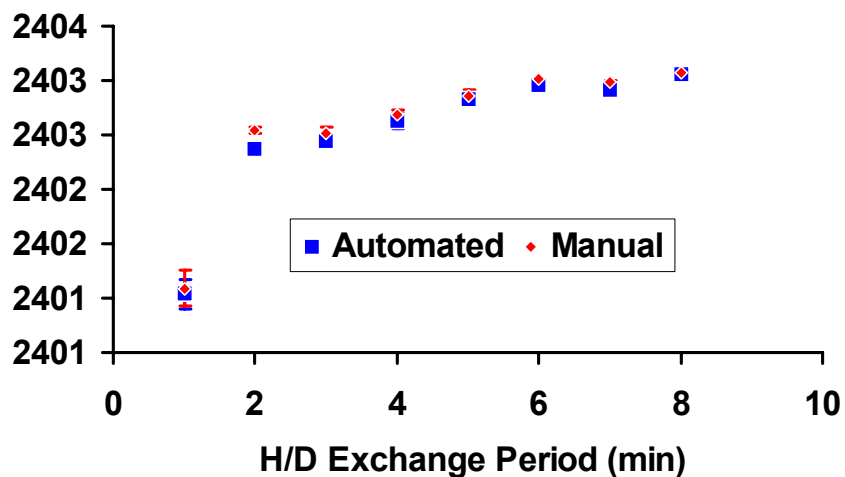


Figure 5.6 H/D Exchange of Complexed and Non-complexed Fibrillarin

Number of
Deuteriums Incorporated

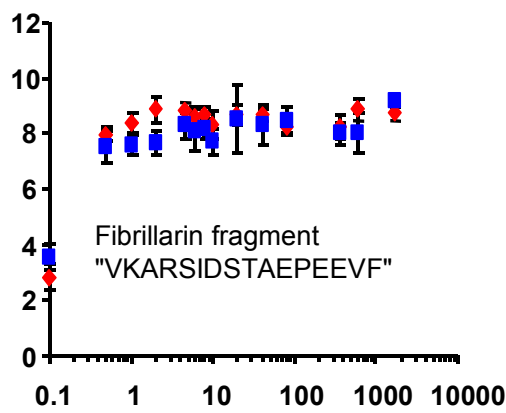


Figure 5.6a

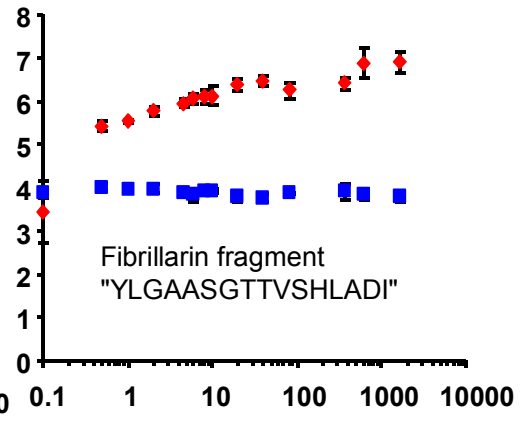


Figure 5.6c

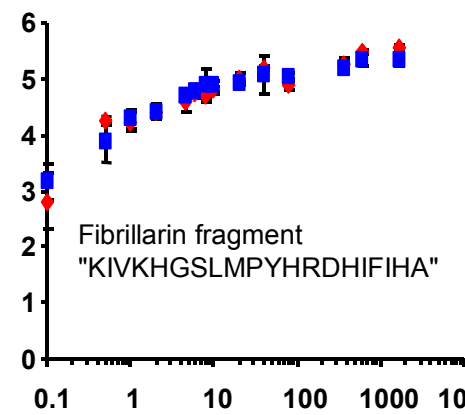


Figure 5.6b

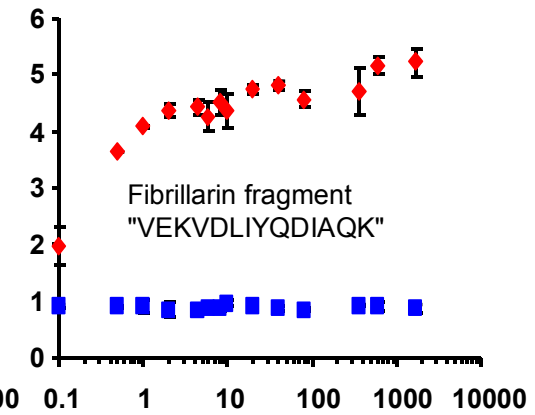
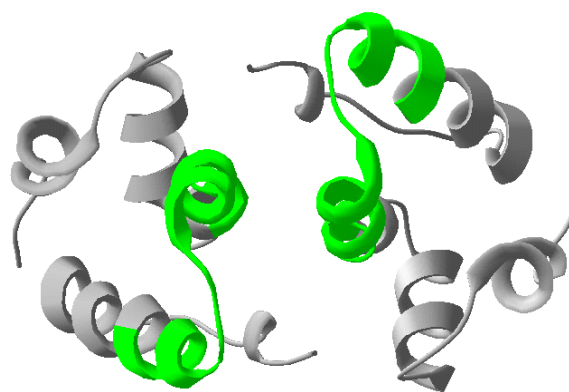
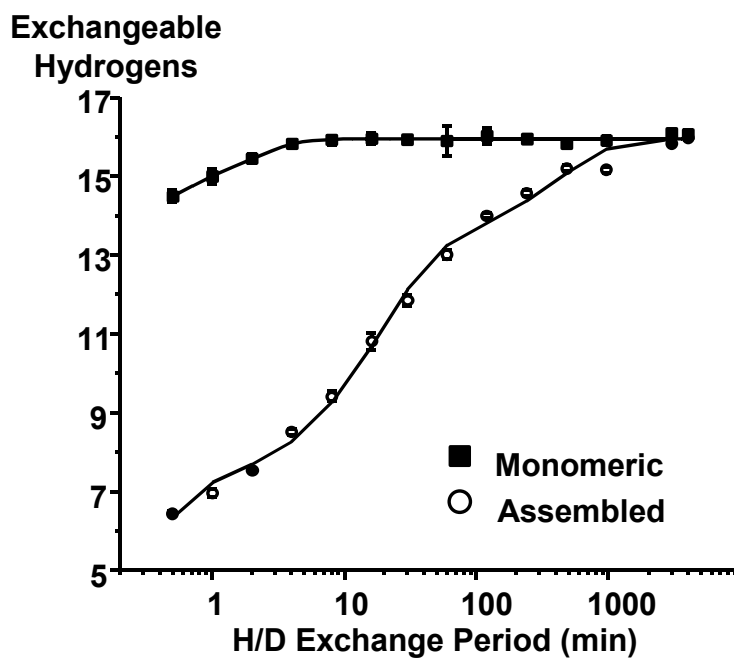


Figure 5.6d

Exchange Period (min)

- ◆ Uncomplexed fibrillarin
- Complexed fibrillarin

Figure 5.7 Automated Analysis of HIV Capsid Proteins by Liquid-phase H/D Exchange with HPLC FT-ICR Detection



was compiled. The regions of interaction were isolated and a previously unmapped interaction between individual hexamer subunit in HIV-1 CA capsid protein was discovered [11].

5.4 Conclusions

An automated method for the batch analysis of H/D exchange data was proposed. The primary benefit this algorithm afforded was speed. Several additional considerations were addressed regarding the minimization of interference and the accuracy of the data. Direct comparisons of the automated and manual yielded only slight variances of the results that were expected given variations of how the individual calculations were performed. Results from the automated analysis were then applied to two biological systems: the Nop5 – fibrillar complex and the HIV-1 CA capsid protein assembly. The results from these studies describe distinct regions of interaction within each system.

5.5 Acknowledgements

The author would like to thank Dr. TuKiet Lam at NHMFL for his expertise and assistance with HPLC, protein purification and discussions of the biological systems involved in these studies. The author would also like to recognize Jason Lanman and Dr Peter Prevelidge for their collaborative work in the HIV-1 CA capsid study. Finally, the author would recognize Dr. Hong Li for providing the Nop5-fibrillar complex for analysis.

5.6 References

1. Freitas, M.A., et al., *Gas-Phase Bovine Ubiquitin Cation Charge State Conformations Resolved by Gas-Phase Hydrogen/Deuterium Exchange Rate and Extent*. Int. J. Mass Spectrom, 1999. **185/186/187**: p. 565-575.
2. Howe, I. and F.W. McLafferty, *Hydrogen Scrambling in Organic Ions as a Function of Internal Energy. Extension of Energy Range*. J. Am. Chem. Soc., 1970. **92**(12): p. 3797-3799.
3. Johnson, R.S., D. Krylov, and K.A. Walsh, *Proton mobility within electrosprayed peptide ions*. J. Mass Spectrom., 1995. **30**(2): p. 386-387.
4. McLafferty, F.W., et al., *Electron Capture Dissociation of Gaseous Multiply Charged Ions by Fourier-Transform Ion Cyclotron Resonance*. J. Am. Soc. Mass Spectrom., 2001. **12**: p. 245-249.
5. Demmers, J.A.A., et al., *Factors Affecting Gas-Phase Deuterium Scrambling in Peptide Ions and Their Implications for Protein Structure Determination*. J. Am. Chem. Soc., 2002. **124**: p. 11191-11198.
6. Katta, V. and B.T. Chait, *Hydrogen/deuterium exchange electrospray ionization mass spectrometry: a method for probing protein conformational changes in solution*. J. Am. Chem. Soc., 1993. **115**(14): p. 6317-21.
7. Anderegg, R.J., et al., *The Mass Spectrometry of Helical Unfolding in Peptides*. J. Am. Soc. Mass Spectrom., 1994. **5**: p. 425-433.
8. Dharmasiri, K. and D.L. Smith, *Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins*. Anal. Chem., 1996. **68**(14): p. 2340-2344.
9. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the noncovalent structure of proteins by amide hydrogen exchange and mass spectrometry*. J. Mass Spectrom., 1997. **32**(2): p. 135-146.

10. Kim, M.-Y., et al., *Conformational changes in chemically modified Escherichia coli thioredoxin monitored by H²D exchange and electrospray mass spectrometry*. Protein Sci., 2002. **11**: p. 1320-1329.
11. Lanman, J., et al., *Identification of Novel Interactions in HIV-1 Capsid Protein Assembly by High Resolution Mass Spectrometry*. J. Mol. Biol., 2003. **325**: p. 759-772.
12. Lam, T.T., et al., *Mapping of protein:protein contact surfaces by hydrogen/deuterium exchange, followed by on-line high performance liquid chromatography - electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry*. J. Chromatography, 2002. **982**: p. 85-95.
13. Senko, M.W., et al., *Electrospray Ionization FT-ICR Mass Spectrometry at 9.4 Tesla*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1824-1828.
14. Senko, M.W., et al., *External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 1997. **8**: p. 970-976.
15. Senko, M.W., et al., *A High-Performance Modular Data System for FT-ICR Mass Spectrometry*. Rapid Commun. Mass Spectrom., 1996. **10**: p. 1839-1844.
16. Blakney, G.T., et al. *Further Improvements to the MIDAS Data Station for FT-ICR Mass Spectrometry*. in *49th Amer. Soc. Mass Spectrom. Conf. on Mass Spectrom. Conf. on Mass Spectrom. & Allied Topics*. 2001. Chicago, IL.
17. Konigsberg, W., J. Goldstein, and R.J. Hill, *The structure of human hemoglobin VII The digestion of the beta chain of human hemoglobin with pepsin*. J. Biol. Chem., 1963. **238**: p. 2028-2033.
18. Ryle, A.P., A.P. Methods Enzym. Anal., 1984. **5**: p. 228.
19. Pain, R.H. in *Proceedings of the 613th meeting of the Cardiff Biochem. Soc. Trans.* 1985.

Chapter 6. Improved ICR Data Station for HPLC FT-ICR MS of Liquid Phase H/D Exchange of Proteins

6.1 Introduction

Liquid phase H/D exchange has been shown to have unique capabilities for the determination of biological interactions [1-7]. Since the H/D exchange occurs in physiologically neutral conditions, the natural conformation of the biological complex is conserved. Liquid phase exchange followed by proteolysis has the ability to localize the exchange information to small fragments within a protein of interest. Mass spectrometry has been shown to be an invaluable partner to liquid phase exchange in many cases [7-18].

FT-ICR MS offers excellent sensitivity, unparalleled resolution and high mass accuracy [15, 19, 20]. The resolution of the FT-ICR allows for the accuracy afforded the isotope methods described previously. HPLC ESI is required to interface liquid phase exchange to FT-ICR to provide on-column desalting and to provide some degree of separation and concentration of the proteolysis products prior to introduction to the FT-ICR. Baseline separation of individual HPLC chromatographic peaks is not necessary as the resolution of the FT-ICR can easily resolve overlapping analytes.

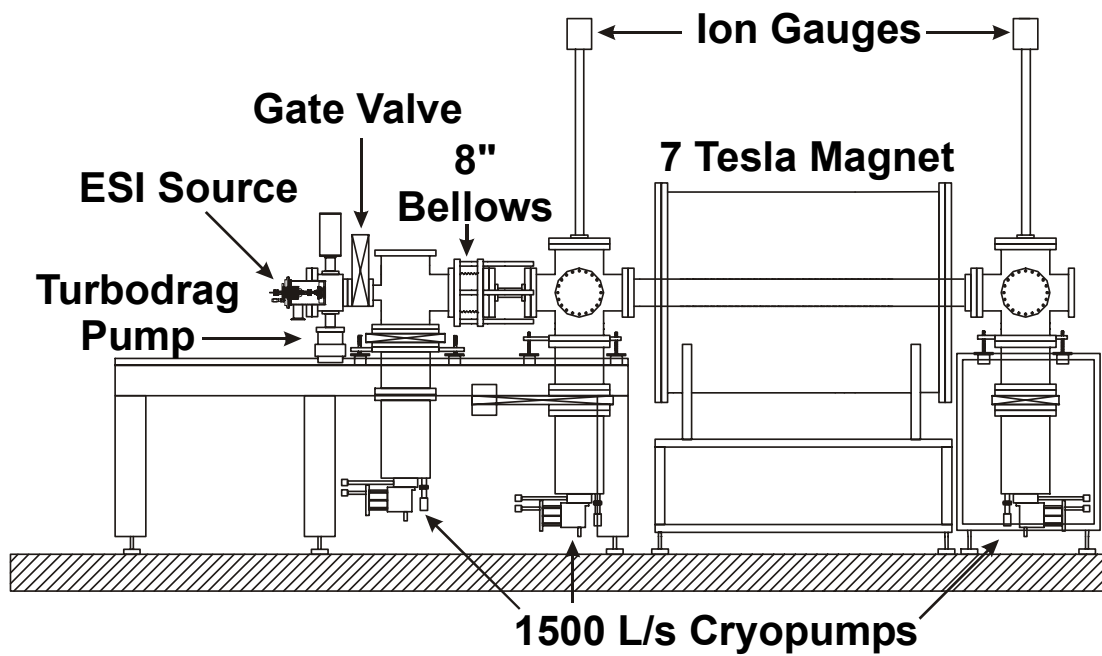
Although HPLC and FT-ICR have many complementary traits, there are also many obstacles that must be overcome to effectively couple these techniques. The primary differences are the fundamental nature of HPLC as a continuous source and the fundamental pulsed nature of FT-ICR detection. Electrospray ionization (ESI) has been shown to be the best choice for linking HPLC to FT-ICR [12, 15, 16, 19, 21], although because of mismatched flow rates, it is not uncommon to use a split-

flow arrangement to reduce the solvent load on the ESI source [22-24]. FT-ICR typically has a fairly slow scan rate of 4 s to 15 s; this slow scan rate degrades the perceived chromatographic resolution. While FT-ICR resolution of the peaks can answer many questions, there are additional benefits one can enjoy if the speed of the FT-ICR experiment can be increased.

The speed of the FT-ICR experiment is primarily determined by two primary factors: ion accumulation period and instrument data station performance. While improved external ion accumulation for the 7.0 T FT-ICR instrument at the National High Magnetic Field Laboratory (NHMFL) is described elsewhere [25, 26], improved data station performance is described in this work. A schematic of the 7T FT-ICR instrument is provided in figure 6.1. The Modular Ion cyclotron resonance Data Acquisition System (MIDAS) affords the user unprecedented experimental flexibility compared to commercial data stations. In the work presented here, the MIDAS is optimized for HPLC performance through the implementation of new hardware.

The MIDAS was developed as a powerful and inexpensive alternative to commercial FT-ICR data stations [27]. The original MIDAS used the industry standard VXIbus that was developed in 1987 as a high-speed instrument interface to replace GPIB communication. PXIbus is poised to replace VXIbus as the performance leader, due to its fast data transfer rate and the advances of the latest PXI cards. As such, PXI is a logical step in the evolution of MIDAS. Further, PXI leverages its technology from the PCI standard so improvements in personal computers can be more rapidly applied to PXI. For instance, increased data throughput when coupled with liquid chromatography (LC) FT-ICR not only enhances chromatographic resolution but also enables further time or data dependent manipulations.

Figure 6.1 7T FT-ICR MS at NHMFL



Nano-LC Micro ESI FT-ICR MS experiments highlight the performance enhancements afforded by the addition of PXI instruments to the MIDAS data station. An LC study of a mixture of bradykinin, des-Arg1 bradykinin and leutenizing hormone releasing hormone (LHRH) is presented. Functionality is further investigated with an LC ESI FT-ICR MS study of a tryptic digest of bovine serum albumin (BSA). Noted benefits include improved Total Ion Chromatogram (TIC) resolution, improved mass resolution, increased mass range, or a combination of the three depending on the conditions selected.

6.2 Methods

6.2.1 FT-ICR MS

Experiments were performed on a homebuilt FT-ICR instrument based on an unshielded 7.0 T Oxford magnet [25]. Analytes were separated by nano-LC, ionized by micro-ESI, and externally accumulated [28] in a modified octopole [26]. The application of the modified octopole was equally important to successfully reducing the effective FT-ICR scan rate since it reduced the period of time required for external ion accumulation by a factor of 10. Without fast ion accumulation, there is little to gain from faster electronics since the total experiment time would remain constant. Ions were then gated from the octopole through a transfer hexapole into an open cell. The accumulation octopole was operated in "continuous" accumulation mode (i.e., except during the ion transfer event --1 ms out of each second). Both the octopole and the hexapole ion guides were operated in rf-only mode.

All FT-ICR experiments were controlled by the MIDAS data station as first described by Senko *et. al.* [27] Of the many recent improvements to MIDAS data

station [29], the latest consists of the addition of PXI hardware. PXI hardware (figure 6.1) allows for rapid transfer of data from the digitizer and rapid and flexible implementation of custom waveforms during the FT-ICR experiment. Specifically, a 14-bit digitizer (model 5620, National Instruments, Austin, TX) and a 12-bit arbitrary waveform generator (AWG model 5411, National Instruments) have been incorporated. The digitizer is housed in a PXI chassis and managed by an embedded processor (model 8176, National Instruments). The original VXI timing cards (model DG600 and DG605 from Interface Technologies, Walnut, CA) control the experimental timing. Control software is written in Measurement Studio (National Instruments). Increased performance was noted only in the digitizer performance because all data was loaded to the AWG prior to initial data collection. Data throughput to the AWG will be critical in future efforts involving data-dependent processing during LC FT-ICR experiments. Figure 6.2 demonstrates the PXI chassis with the improved digitizer housed within.

6.2.2 HPLC Conditions

LC experiments were conducted with a fast-load capillary scale apparatus comprising a pre-column and a separating column, plumbed into 2 opposite sides of a low dead volume T-piece. A vent line was connected to the third port. Samples (10 fmol/ μL , dissolved in 0.1% formic acid) were loaded onto the 75 μm x 1 cm pre-column (Biobasic C18, New Objective inc.) at a flow rate of 1 $\mu\text{L min}^{-1}$ (5 μL sample loop). At this stage the vent line was open, and the flow through was directed to waste. Closing the vent line directed flow through the 75 μm x 5 cm analytical column (BioBasic C18, New Objective inc.) at a reduced flow rate of approximately 200 nL min^{-1} . A 6-minute linear gradient of 2% CH_3CN (0.05% FA), increasing to 72% CH_3CN (0.05% FA), was used to elute the peptides from

Figure 6.2 Image of the PXI Chassis Housing the Improved MIDAS Hardware



the pre-column, across the analytical column, where they were micro-electrosprayed directly from the column tip (15 μm) into the 7T FT-ICR instrument.

6.3 Results and Discussion

The primary focus of this work is to improve the speed of the FT-ICR experiment while maintaining mass range and resolution typical of high-field FT-ICR MS. Figure 6.3a represents a reconstructed ion chromatogram of a mixture of bradykinin, des - Arg1 bradykinin, and LHRH. Note that the single ion chromatogram (SIC) for each component is indicated in color while the TIC is overlaid in black. For this experiment, 5 μL of a 10 fmol/ μL solution of each analyte was loaded on the column (50 fmol total). 300 scans were collected with 1 MSample of data at 2.4 MSample/s with a total scan time of 1.25 s. In Figure 6.3a the centroids are 6 scans apart. Figure 6.3b represents an identical experiment using the same experimental script with the previous MIDAS configuration of all VXI hardware. For this version of MIDAS the average scan time was 3.5 seconds, and the bradykinin and LHRH primary chromatographic peaks were separated by only two FT-ICR scans. The 3.5 seconds scan includes general speed improvements incorporated during the development of the PXI MIDAS and realizes a 2 second advantage over our traditional LC FT-ICR experiment. These additional scans are critical to the success of future data-dependent LC FT-ICR experiments. Further at 1.25 s per scan, the scan resolution is more compatible with the LC resolution for this experiment.

Figure 6.4 provides a chronological sketch of the activities that occur in a 1.25 s HPLC FT-ICR scan. It is important to note that the 1.25 s includes 0.75 s of

Figure 6.3 Comparison of TIC Collected Using Identical Experimental Scripts with the VXI and PXI MIDAS Data Stations

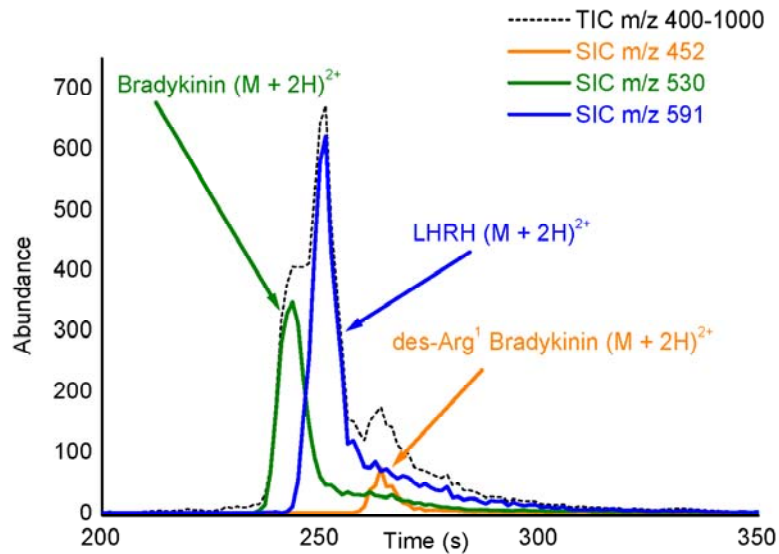


Figure 6.3a Reconstructed TIC with PXI enabled MIDAS

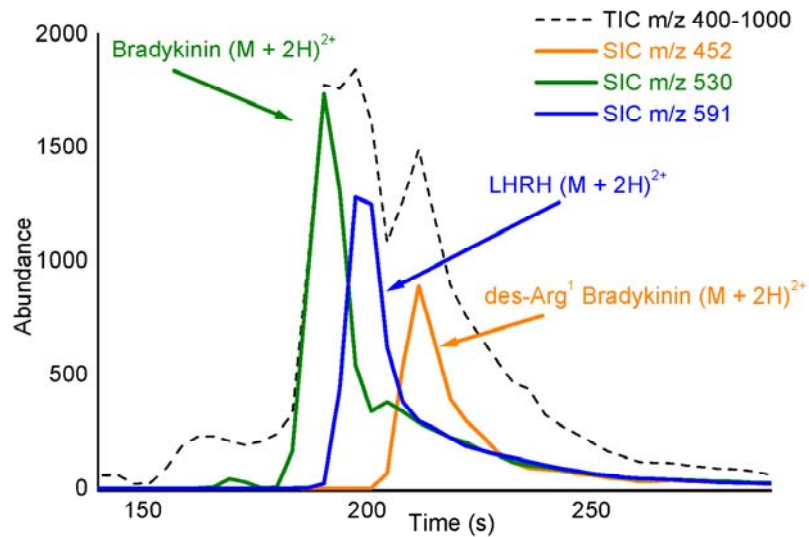
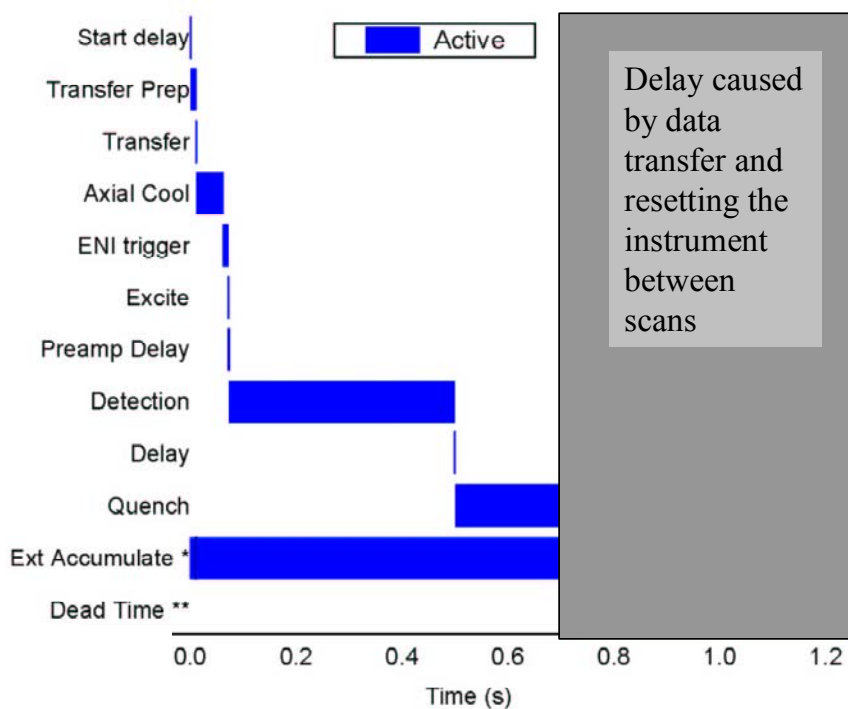


Figure 6.3b Reconstructed TIC collected with the classic MIDAS

Figure 6.4 ICR Experimental Timeline of a Single 1.25s HPLC FT-ICR MS Scan



* External Accumulation occurs throughout the experiment except during the ion transfer event. This equates to an effective duty cycle of 99.9% (1 ms downtime in a 1.25 s experiment)

** Dead Time refers a period in which voltages and triggers become static and are accessible to the user. Data transfer from the digitizer, storage of data, resetting the instrument for the next scan and updating the user interface are some of the procedures which are processed during this period.

time that is directly controlled by the timing card while the timing card is actively controlling the experiment. The additional 0.5 s is required to store the data and reset the timing card, AWG, and digitizer for the subsequent scan. Of this “dead” time, the time required for the data transfer of 1 MSample from the PXI digitizer was 50 ms. Using the same script with the VXI MIDAS result in a dead time of 2.75 s. Much of the dead time could be contributed to the 2 s that is required to download the VXI digitizer. Additional time of approximately 0.2 s is also required to reset the digitizer relative to the PXI digitizer.

Figure 6.5 demonstrates the quality of the spectra obtained in the fast LC experiment. Fourier limited mass resolving power for the spectra in Figure 6.5a, b, c (PXI digitizer at 1.25 s accumulation) averages 20,000 with signal to noise ratio (S/N) from 650 to 3300. Resolution in figures 6.5d, e, f (VXI digitizer with 3.5 s accumulation time) ranges from 30,000 to 48,000 with S/N of 8700 to 14000. Resolution was calculated as $m/\Delta m_{50\%}$ or full width at half-maximum (FWHM) [30]. This comparison of 1.25s and 3.5s scan rates was continued in figure 6.6. The SICs for each of the mass range remain scaled to their original abundance. The abundance for 3.5 s scans, shown in red, demonstrates significantly higher abundance the 1.25 s scans. While it is expected that increased accumulation would lead to increased FT-ICR signal, there seems to be some mass dependence on the abundance ratio for the 3.5 vs 1.25 s scans periods. This mass dependence is most noted in the m/z range near 452, the abundance ratio slipped to approximately 1:10 for accumulation time that only varied by approximately a factor of 3. There is a known time of flight effect during transfer from the external accumulation to the ICR cell that could account for this apparent discrepancy[26].

Figure 6.7 demonstrates an additional LC FT-ICR experiment using the PXI MIDAS with the preceding analyte system but with the total accumulation time

further reduced from 1.25 s to 940 ms. The 940 ms data demonstrates marked differences in the relative abundances of the analytes. While outside the scope of this work, further studies are planned to investigate the abundance inconsistencies at short vs. long accumulation times. Figure 6.7 was presented to demonstrate that the PXI MIDAS data station was capable of collecting sub-second LC FT-ICR with 1 MSample data sets.

The mass insets in Figure 6.7 demonstrate the signal-to-noise (S/N) and mass resolution from the scan of each analyte at its maximum abundance as determined from its SIC. SIC are present in color whereas the TIC is overlaid in black. For this experiment, mass-resolving powers of 35,000 to 50,000 FWHM were observed and S/N ranged from 360 to 3500.

Figure 6.8 represents an LC FT-ICR run of a tryptic digest of BSA. For this run 40 fmol of sample was injected on column. 600 scans were collected at 1.25 s per scan (1 MSample at 2.4 MSample/s). Data analysis was performed by MIDAS analysis and fragments were assigned by the THRASH algorithm [31]. Mass resolving power for the displayed peptides was ~20,000 FWHM and S/N ranges from 200 to 400. Obviously, the S/N for the more minor components is proportionally lower.

6.4 Conclusions

The study of biological interactions can be best characterized by liquid phase H/D exchange studies. The use of HPLC and ESI are required to facilitate the coupling of FT-ICR MS to liquid phase exchange. Typical FT-ICR instruments operate on a time period that is not ideally suited to optimum HPLC. The author has

Figure 6.5 Direct Comparison of the Spectral Quality of the 1.25s and 3.50s LC FT-ICR MS Experiments

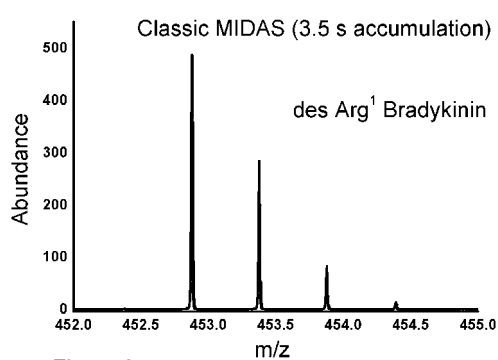


Figure 6.5a

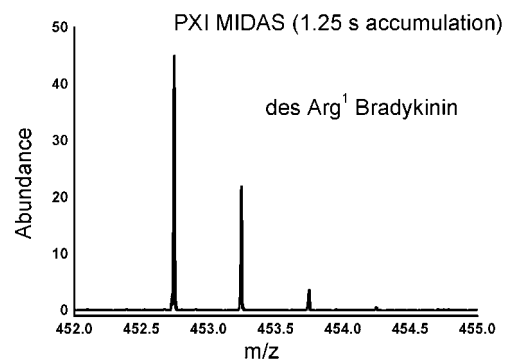


Figure 6.5d

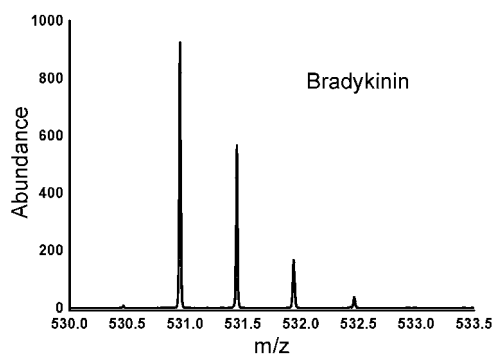


Figure 6.5b

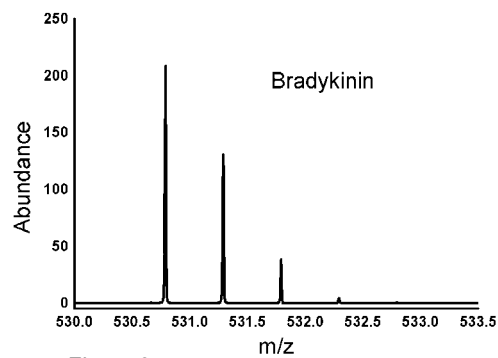
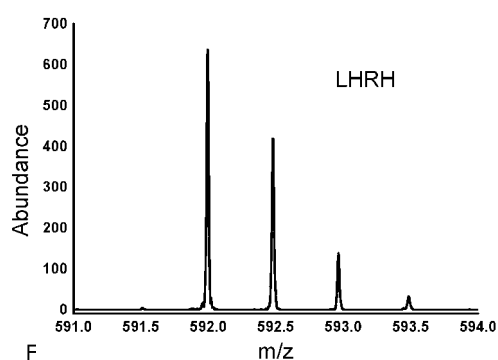


Figure 6.5e



F

Figure 6.5c

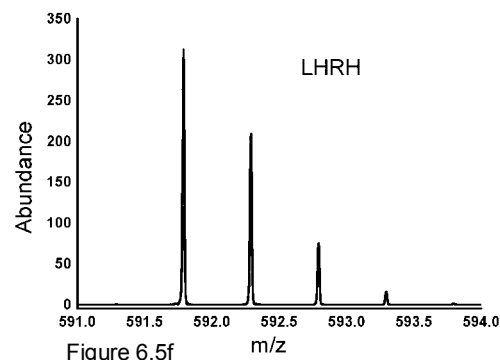


Figure 6.5f

Figure 6.6 Single Ion Chromatograms (SICs) of Three Peptide from the 1.2s and 3.5s FT-ICR Experiments

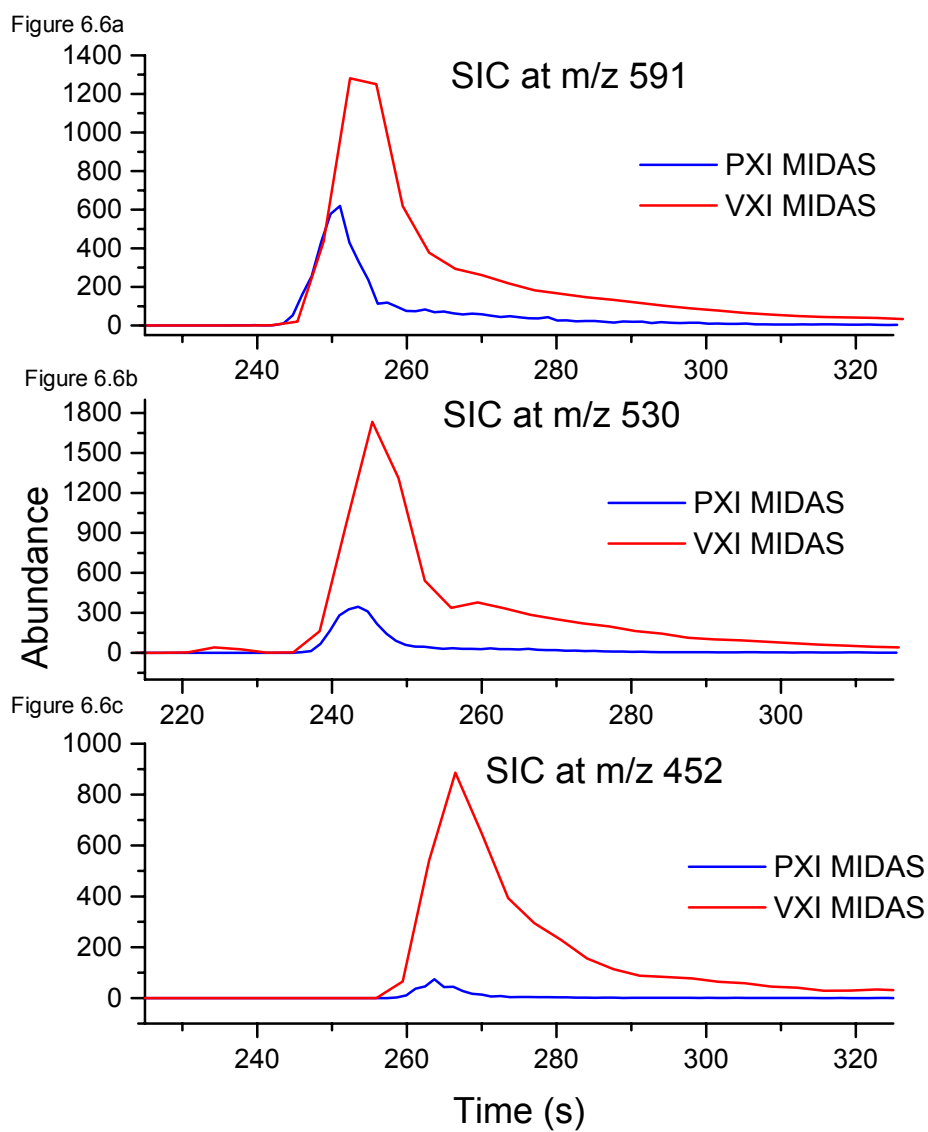


Figure 6.7 HPLC FT-ICR Experiment Using the PXI MIDAS Provides Further Reduction in ICR Scan Time to 940 ms per Scan

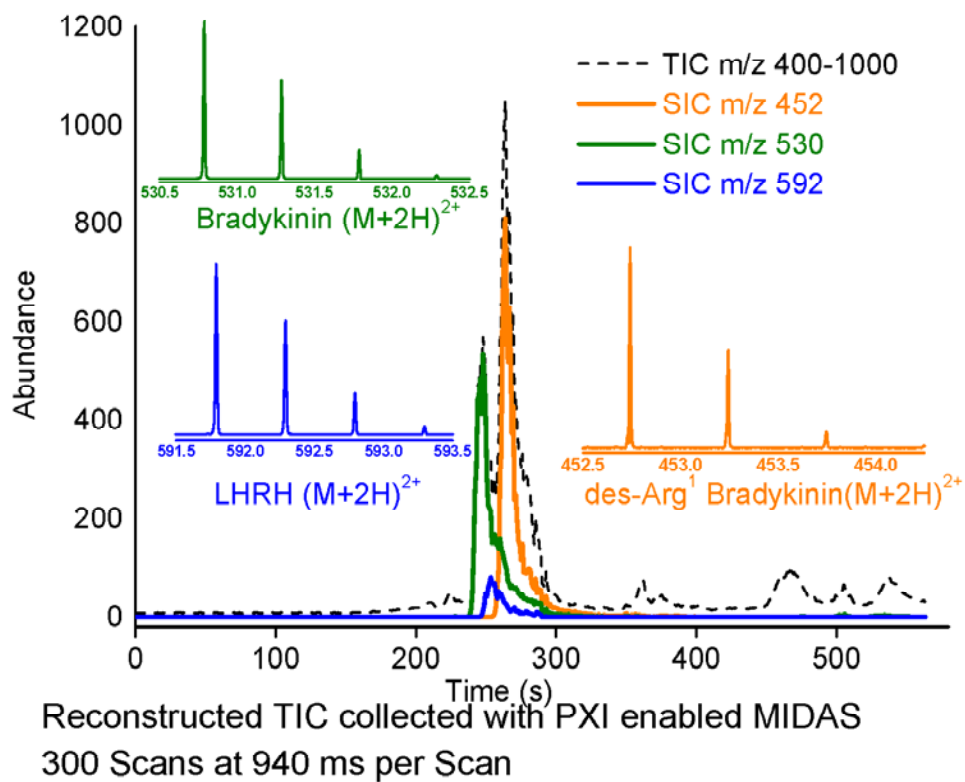
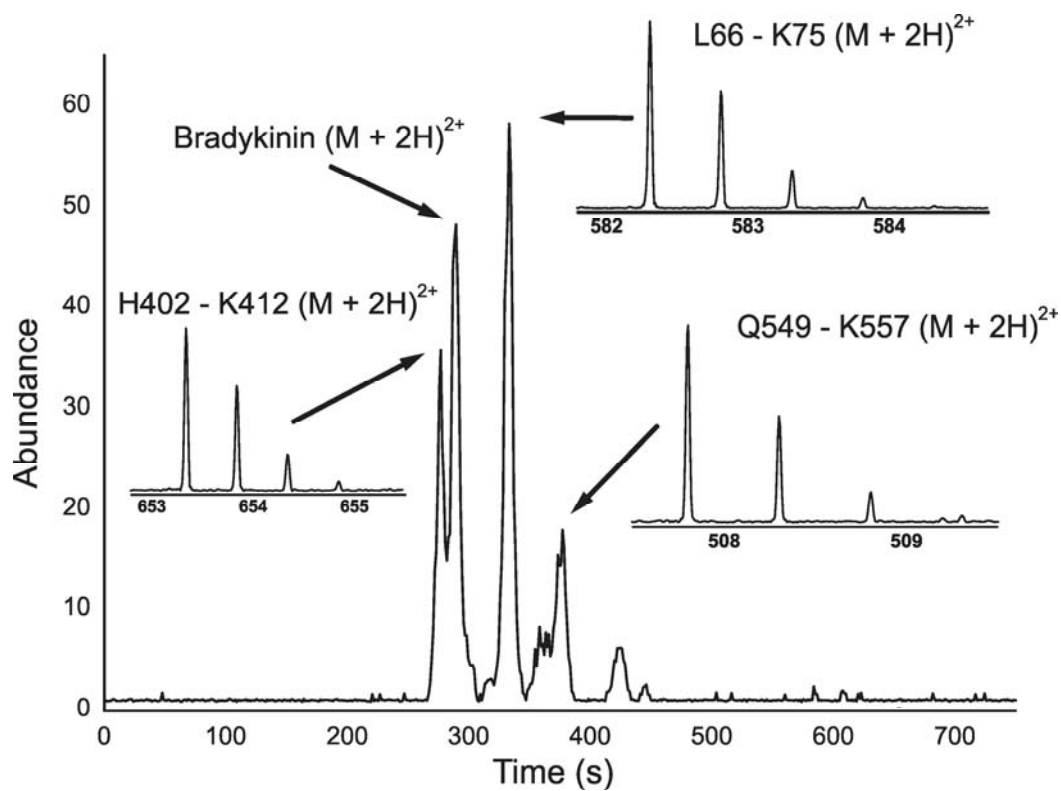


Figure 6.8 HPLC FT-ICR MS Analysis of a Tryptic Digest of Bovine Serum Albumin



introduced PXI hardware to an ICR data station that reduces the total period for an HPLC FT-ICR MS experiment to less than 1 second. This represents an improvement of more than a three and a half times the scan time compared to current FT-ICR technology. These improvements were shown to increase chromatographic resolution, increase number of target scans per chromatographic peak and maintain FT-ICR resolution. Data was presented that demonstrated improved performance for standard systems as well as tryptic digest representing more complex matrices expected from liquid phase exchange experiments.

References

1. Molday, R.S., S.W. Englander, and R.G. Kallen, *Primary Structure Effects on Peptide Hydrogen Exchange*. *Biochemistry*, 1972. **11**(2): p. 150-158.
2. Roder, R., G. Wagner, and K. Wuthrich, *Amide Proton Exchange*. *Biochemistry*, 1985. **24**: p. 7396-7407.
3. Englander, S.W., et al., *Mechanisms and uses of hydrogen exchange*. *Current Opinion in Structural Biology*, 1996. **6**(1): p. 18-23.
4. Li, R. and C. Woodward, *The hydrogen exchange core and protein folding*. *Protein Sci.*, 1999. **8**: p. 1571-1591.
5. Flaux, J., et al., *NMR Analysis of a 900K GroEl-GroES complex*. *Nature*, 2002. **418**: p. 207-211.
6. Hoang, L., et al., *Cytochrome c folding pathway: Kinetic native-state hydrogen exchange*. *PNAS*, 2002. **99**(19): p. 12173-12178.
7. Kim, M.-Y., et al., *Conformational changes in chemically modified Escherichia coli thioredoxin monitored by H²D exchange and electrospray mass spectrometry*. *Protein Sci.*, 2002. **11**: p. 1320-1329.
8. Katta, V. and B.T. Chait, *Hydrogen/deuterium exchange electrospray ionization mass spectrometry: a method for probing protein conformational changes in solution*. *J. Am. Chem. Soc.*, 1993. **115**(14): p. 6317-21.
9. Anderegg, R.J., et al., *The Mass Spectrometry of Helical Unfolding in Peptides*. *J. Am. Soc. Mass Spectrom.*, 1994. **5**: p. 425-433.
10. Kim, H.Y., T.C.L. Wang, and Y.C. Ma, *Liquid Chromatography/Mass Spectrometry of Phospholipids Using Electrospray Ionization*. *Anal. Chem.*, 1994. **66**: p. 3977-3982.
11. Johnson, R.S. and K.A. Walsh, *Mass spectrometric measurement of protein amide hydrogen exchange rates of apo- and holo-myoglobin*. *Protein Sci.*, 1994. **3**(12): p. 2411-2418.

12. Stacey, C.C., et al., *Reverse-phase Liquid Chromatography/Electrospray-Ionization Fourier-Transform Mass Spectrometry in the Analysis of Peptides*. Rapid Commun. Mass Spectrom., 1994. **8**: p. 513-516.
13. Dharmasiri, K. and D.L. Smith, *Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins*. Anal. Chem., 1996. **68**(14): p. 2340-2344.
14. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the Non-covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry*. J. Mass Spectrom., 1997. **32**: p. 135-146.
15. Emmett, M.R., et al., *Application of Micro-Electrospray Liquid Chromatography Techniques to FT-ICR MS to Enable High Sensitivity Biological Analysis*. J. Am. Soc. Mass Spectrom., 1998. **9**: p. 333-340.
16. Zhang, Z. and A.G. Marshall, *Method for Gradient Elution in Micro-Flow Liquid Chromatography*". J. High Resolution Chromatogr., 1998. **21**: p. 291-297.
17. Li, W., et al., *Identification of Intact Proteins in Mixtures by Alternated Capillary Liquid Chromatography Electrospray Ionization and LC ESI Infrared Multiphoton Dissociation Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. Anal. Chem., 1999. **71**: p. 4397-4402.
18. Anand, G.S., et al., *Amide H/H2 Exchange Reveals Communication Between the cAMP and Catalytic Subunit-binding sites in the R'alpha Subunit of Protein Kinase A*. J. Mol. Biol., 2002. **323**: p. 377-386.
19. Andren, P.E., M.R. Emmett, and R.M. Caprioli, *Micro-Electrospray: Zeptomole/Attomole per Microliter Sensitivity for Peptides*. J. Am. Soc. Mass Spectrom., 1994. **5**: p. 867-869.
20. Emmett, M.R. and R.M. Caprioli, *Micro-Electrospray Mass Spectrometry: Ultra-High-Sensitivity Analysis of Peptides and Proteins*. J. Am. Soc. Mass Spectrom., 1994. **5**: p. 605-613.

21. Marshall, A.G., C.L. Hendrickson, and M.R. Emmett, *Recent Advances in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*, E.J. Karjalainen, et al., Editors. 1998, Elsevier, B.V.: Amsterdam. p. 219-237.
22. Licklider, L.J., et al., *Automation of nanoscale microcapillary liquid chromatography - tandem mass spectrometry with a vented column*. *Analytical Chemistry*, 2002. **74**: p. 3076-3083.
23. Chalmers, M.J., et al. *Affinity Isolation and Tandem Mass Spectrometric Approaches for the Characterization of Phosphorylation within Mitogen Activated Protein Kinases*. in *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics*. 2001. Chicago.
24. Chalmers, M.J., et al., *Liquid Chromatography - Fourier Transform Ion Cyclotron Resonance Mass Spectrometric Characterization of Protein Kinase C Phosphorylation*. Submitted to J. Proteome Research, 2003.
25. White, F.M., J.A. Marto, and A.G. Marshall, *An External Source FT-ICR Mass Spectrometer with Electrostatic Ion Guide*. *Rapid Commun. Mass Spectrom.*, 1996. **10**: p. 1845-1849.
26. Wilcox, B.E., C.L. Hendrickson, and A.G. Marshall, *Improved Ion Extraction from a Linear Octopole Ion Trap: SIMION Analysis and Experimental Demonstration*. *J. Am. Soc. Mass Spectrom.*, 2002. **13**: p. 1304-1312.
27. Senko, M.W., et al., *A High-Performance Modular Data System for FT-ICR Mass Spectrometry*. *Rapid Commun. Mass Spectrom.*, 1996. **10**: p. 1839-1844.
28. Senko, M.W., et al., *External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. *J. Am. Soc. Mass Spectrom.*, 1997. **8**: p. 970-976.
29. Blakney, G.T., et al. *Further Improvements to the MIDAS Data Station for FT-ICR Mass Spectrometry*. in *49th Amer. Soc. Mass Spectrom. Conf. on*

Mass Spectrom. Conf. on Mass Spectrom. & Allied Topics. 2001. Chicago, IL.

30. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer*. *Mass Spectrom. Rev.*, 1998. **17**: p. 1-35.
31. Horn, D.M., R.A. Zubarev, and F.W. McLafferty, *Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules*. *J. Am. Soc. Mass Spectrom.*, 2000. **11**: p. 320-332.

Chapter 7: Conclusions and Future Directions

7.1 Conclusions

H/D exchange remains a powerful tool for the study of biological interaction and structural elucidation for large biomolecules. Work presented in this dissertation has shown that liquid phase H/D exchange followed by high performance liquid chromatography (HPLC) Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) is a robust technique capable of the determination of biological interactions of species of 250 kDa. H/D exchange by FT-ICR MS has multiple advantages over similar studies by NMR. Among these advantages are reduced sample and sample preparation time, and the ability to analyze a wider variety of samples [1-3].

Analytical techniques were described that improved data analysis time and accuracy. The automated center of mass calculations show a better fit to MEM studies than calculations based on the last isotope method. Additional refinements led to the development of automated batch analysis based upon the center of mass calculation, and the automation reduces the total analysis time for liquid exchange HPLC FT-ICRMS studies from two months to about two weeks. These automated batch routines allow the user to begin high-level analysis of the H/D exchange data directly from the output of the algorithm.

Automated analysis of spectra was also presented with the work required to transport the algorithm to the PC platform and improve the functionality of the thorough high resolution analysis of spectra by Horn (THRASH) algorithm [4]. Improvements include speed gains that were realized from the use of modern PC processors and the use of speed-optimized libraries for routines such as fast Fourier

transform and array-based algebra. THRASH was used for the analysis of a complex electron capture dissociation spectrum of bovine ubiquitin. The automated analysis resulted in 85% sequence determination of the peptide in a five minute experiment and a 3 minute analysis. Comparable analysis by THRASH on a Sun Microsystems platform required 19 minutes.

Liquid phase exchange HPLC FT-ICR MS experiments were performed to discover unique sites of biological interaction for both the HIV-1 capsid hexamer interaction and the Nop5 – fibrillar interaction [5, 6]. Additional experiments are underway that will continue to combine liquid phase exchange HPLC FT-ICR MS and automated batch analysis as an effective process for the better understanding of the biological interactions inherently to life itself.

7.2 Future Directions

The further hardware upgrades to the modular ICR data acquisition system (MIDAS) will be paramount to improvement of H/D exchange and HPLC data at the NHMFL. The reduction in HPLC scan time to under 2 seconds with the ability to generate unique timing sequences on the fly will allow the collection of data dependent MS/MS scan throughout an HPLC run. The software has been developed to allow such experiments while the hardware is being developed off-site.

Figure 7.1 demonstrates the user panel that allows one to perform data-dependent experiments. The upper right portion of the panel describes events that are to be performed upon the identification of a qualified peak. The lower right portion of the panel describes events that return the data collection to normal HPLC conditions. The user has the ability to choose multiple targets per selection period. If

Figure 7.1 Control Panel for Data Dependent Acquisition

Automatic Fragmentation Criteria

Selection Commands

Use Selection

Continuous Accumulation

Number of Peaks to Select: 3

Peak Picking Parameters

Lower m/z: 500.00 Upper m/z: 1400.00

Peak Threshold: 20.00 Noise Threshold: 10.00

Min Peak Width: 3.00

Relative Absolute

Exclusion List Parameters

Exclusion Error (m/z): 1.000 Number of Scans to Exclude: -1

Event Label	Command	Selection Formula	Loop	Table	Event	Delay
1 accumulate in o'pole 1	V8	mass/400	0	1	0	0
2 transfer to o'pole 2	V8	mass/400	0	1	1	0
3 accumulate in o'pole 1	T11		0	1	0	0
4 transfer to o'pole 2	T11		0	1	1	0

Return to Normal Commands

Event Label	Command	Selection Formula	Loop	Table	Event	Delay
1 accumulate in o'pole 1	V8	1.00	0	1	0	0
2 transfer to o'pole 2	V8	1.00	0	1	1	0
3 accumulate in o'pole 1	T11		0	1	0	0
4 transfer to o'pole 2	T11		0	1	1	0

Buttons: Add New to List, Remove from List, Delete All from List, Start Experiment, Delete All and Cancel

multiple targets are chosen, the instrument will remain in selection mode until all qualified targets have been isolated. It is the responsibility of the user to assure that the number of targets selected will not create a condition in which the time required to collect the preceding targets will not hinder the collection of the last target. In other words, the user must match the experiment time to the chromatographic peak widths if he or she expects to obtain satisfactory results.

The user input panel for the data-dependent acquisition parameters is shown in figure 7.2. If one is familiar with the MIDAS acquisition design, it is apparent that the commands for data dependent work closely mimic the commands for normal data acquisition. This familiarity allows one to immediately focus on the task of learning the new technique of data dependent data acquisition. Drop down menus in figure 7.2 are labeled identically to the primary commands. Since the data controls for data dependent acquisition process commands sequentially, each command must be separated. Therefore, the selection of a new command type automatically resets to remainder of the commands to “None”. The user can currently command the following parameters: voltages, triggers, excitation and event length. The addition of parameters such as the ability to disable and enable events is planned shortly. The “Formula commands” at the top of the panel is the location where the data value is stored for each event. The data value can be as simple as a 0 or 1 for the operation of TTL triggers or may include the use of mathematical operators of multiplication (*), division (/), addition (+) and subtraction (-). The use of parentheses is permitted as the execution order of the operators follows scientific methodology. There are two variables that allow true data-dependent acquisition to occur. These variables are labeled “mass”, which is in actuality m/z , and “amp” represents the amplitude or abundance of the qualified peak.

Figure 7.2 Selection Input Panel for Data Dependent Acquisition

The image shows a software dialog box titled "Add to List". At the top, there is a dropdown menu labeled "Add To Selection". Below this is a text field labeled "Formula Commands" containing the value "0".

The main area of the dialog is divided into several sections:

- Major Loop:** A dropdown menu set to "MainLoop".
- Event Tables:** A dropdown menu set to "q-c accumulate".
- Experiment Count:** A numeric input field set to "1".
- Count:** A numeric input field set to "1".
- Count:** A numeric input field set to "10".
- Disabled:** A radio button that is currently selected.
- Events:** A text field containing "transfer to o'pole 2".

On the right side of the dialog, there are several dropdown menus for selection criteria:

- Event Length for Selection:** Set to "None".
- Trigger for Selection:** Set to "Quad Selec".
- Voltage for Selection:** Set to "None".
- Excite for Selection:** Set to "None".

At the bottom of the dialog, there are two buttons: "ADD" and "EXIT". In the bottom right corner, there is a checkbox labeled "Delay until Next Experiment" which is currently unchecked.

Current operation of the data dependent algorithm allows the user to use standard peak picking parameters including a data range for the analysis to generate a list of peaks from each “normal” LC run. The peaks are ranked from most abundant to least. In order for a peak to qualify for selection, it must be contained in the peak list, not be on the exclusion list and have no more than the user defined number of qualified peaks of greater abundance already selected. To clarify the rules of qualification, the following scenario is described. The user selects 2 target peaks in the main data dependent panel; the largest two peaks will qualify for selection for this run. This assumes that there were at least two peaks and that neither is listed on the exclusion list. If the most abundant peak is excluded, the second and third most abundant peaks are qualified. If only one peak is qualified, then there is only selection scan followed by a normal scan. If there are no qualified peaks, then normal LC scans continue until there is at least one qualified peak.

The exclusion list is determined by user selection “Number of Scans to Exclude” on the main data dependent panel. Initially, there are no target masses on the exclusion list. As each target peak is qualified, it is placed on the exclusion list so that the same peak is not selected multiple times. The exclusion of larger peaks allows the program to probe deeper into the subsequent spectra and get better sequence coverage. The excluded peaks remain on the exclusion list for as many scans as the user described above. The value of “-1” for the exclusion command signifies that a peak that will remain on the exclusion once selected list for the entire experiment. The use of exclusion files will allow the user to permanently preclude the selection of certain peaks such as trypsin auto-digestion products from a tryptic digest analysis. Associated with the exclusion list is an error that allows one to alleviate small drifts in mass as well as excluding the isotope peaks of previous selection targets.

A mixture of ubiquitin, substance P, and mellitin was analyzed on the 9.4T instrument at NHMFL [7]. The results are presented in figure 7.3 with the ten most abundant species labeled. Using the selection parameters from figure 7.1, the ten most abundant peaks from this spectrum were selected in correct order. Figure 7.4 demonstrates the spectra for the six most abundant of the species in the order the spectra were collected (the interim normal LC scans were removed during analysis). The selection in this case was not complete since the only criteria for ion selection was the MIDAS control of the front-end quadrupole on the FT-ICR MS. In the future, stored waveform inverse Fourier transform (SWIFT) isolation will supplement the quadrupole isolation. While the data-dependent isolation of these species was demonstrated, the ability to do so on a HPLC timescale was not. The average scan for the data-dependent isolation required approximately 13 seconds to acquire.

The addition of new hardware is eagerly anticipated as the solution to data dependent data acquisition. The hardware currently under incorporation into the MIDAS at NHMFL promises to minimize the acquisition time for data dependent ICR data to 2 seconds or less. This hardware incorporates complete use of PCI and PXI components to replace their slower VXI counterparts. Additionally a robotic front end (LEAP Technologies, Carrboro, NC) for sample prep, H/D exchange control and sample introduction to the FT-ICR is expected to complement the data acquisition hardware.

Figure 7.3 Broadband Spectrum of Peptide Mixture with the Ten Most Abundant Species Listed

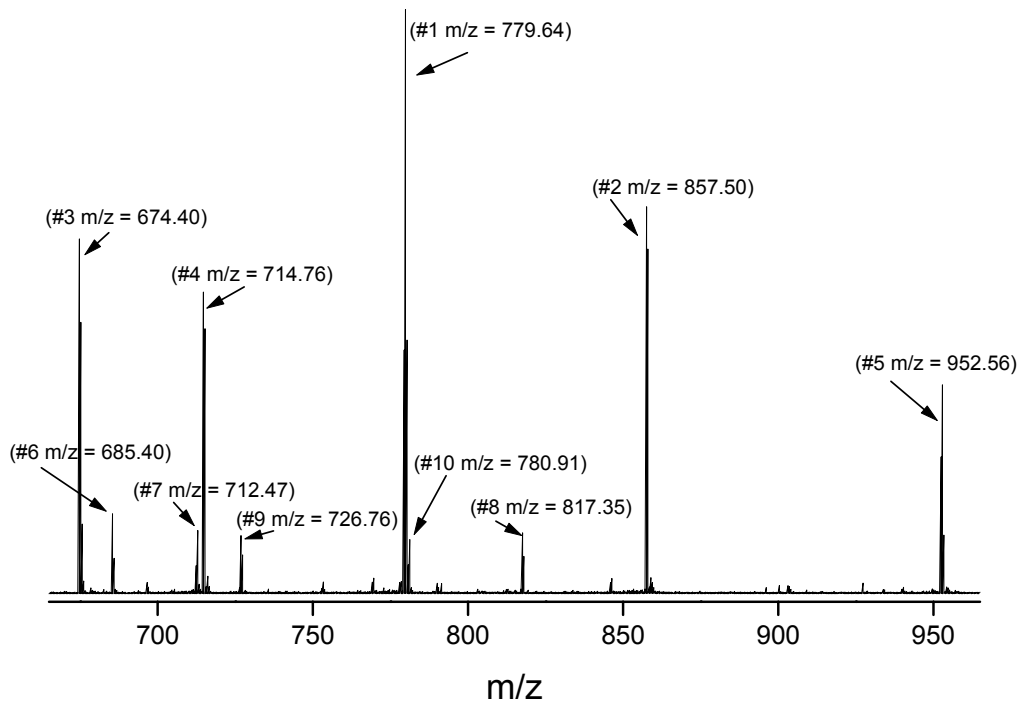


Figure 7.4 Data dependent Selection of the 6 Most Abundant Peaks in a Peptide Mixture

Figure 7.4a.

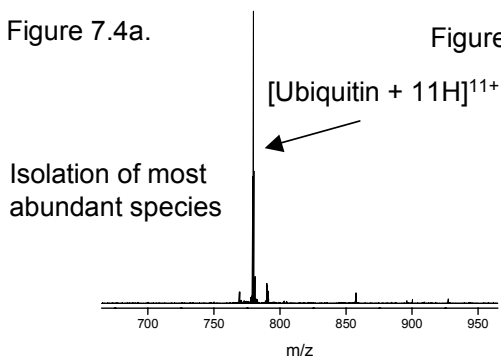


Figure 7.4d.

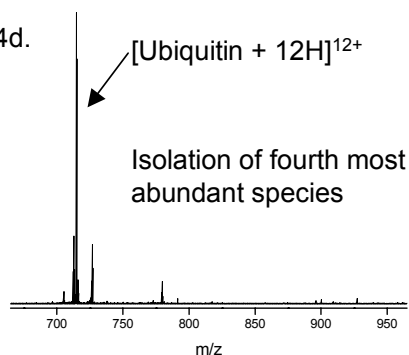


Figure 7.4b.

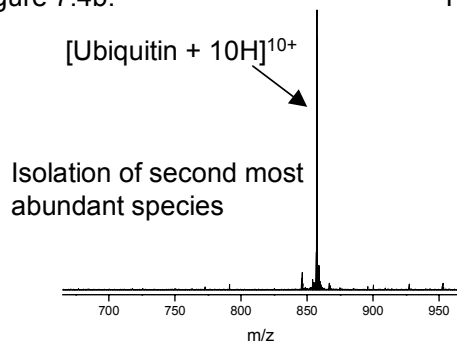


Figure 7.4e.

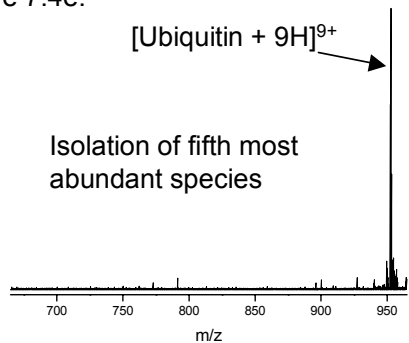


Figure 7.4c.

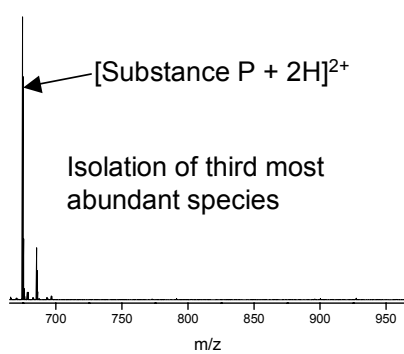
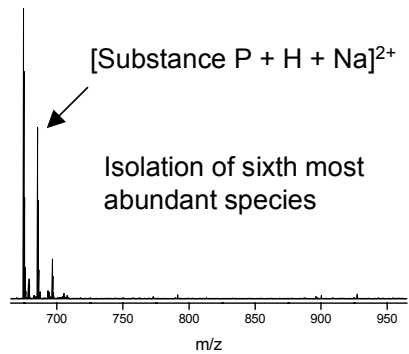


Figure 7.4f.



References

1. Akashi, S. and K. Takio, *Probing higher order structure of proteins by hydrogen/deuterium exchange and mass spectrometry*. J. Mass Spectrom. Soc. Japan, 2000. **48**(2): p. 94-100.
2. Dharmasiri, K. and D.L. Smith, *Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins*. Anal. Chem., 1996. **68**(14): p. 2340-2344.
3. Smith, D.L., Y. Deng, and Z. Zhang, *Probing the Non-covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry*. J. Mass Spectrom., 1997. **32**: p. 135-146.
4. Horn, D.M., R.A. Zubarev, and F.W. McLafferty, *Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules*. J. Am. Soc. Mass Spectrom., 2000. **11**: p. 320-332.
5. Lam, T.T., et al., *Mapping of protein:protein contact surfaces by hydrogen/deuterium exchange, followed by on-line high performance liquid chromatography - electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry*. J. Chromatography, 2002. **982**: p. 85-95.
6. Lanman, J., et al., *Identification of Novel Interactions in HIV-1 Capsid Protein Assembly by High Resolution Mass Spectrometry*. J. Mol. Biol., 2003. **325**: p. 759-772.
7. Senko, M.W., et al., *External Accumulation of Ions for Enhanced Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*. J. Am. Soc. Mass Spectrom., 1997. **8**: p. 970-976.

Bibliography

- Asamoto, B. and R. C. Dunbar (1991). Analytical Applications of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. New York, VCH.
- Cole, R. B., Ed. (1997). Electrospray Ionization Mass Spectrometry. New York, John Wiley & Sons, Inc.
- Horn, D. M., R. A. Zubarev, et al. (2000). "Automated Reduction and Interpretation of High Resolution Electrospray Mass Spectra of Large Molecules." J. Am. Soc. Mass Spectrom. **11**: 320-332.
- Marshall, A. G., C. L. Hendrickson, et al. (1998). "Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer." Mass Spectrom. Rev. **17**: 1-35.
- Marshall, A. G. and F. R. Verdun (1990). Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook. Amsterdam, Elsevier.
- Senko, M. W., J. D. Canterbury, et al. (1996). "A High-Performance Modular Data System for FT-ICR Mass Spectrometry." Rapid Commun. Mass Spectrom. **10**: 1839-1844.
- Voet, D. and J. G. Voet (1995). Biochemistry, Second ed. New York, John Wiley & Sons, Inc.

Vita

Gregory Terrell Blakney was born on September 2, 1969 in Fort Worth, Texas to Nancy Larue Blakney and Richard Douglas Blakney. After graduating as valedictorian from Granbury High School in 1988, he entered the University of North Texas in Denton, Texas. In the spring of 1993, he graduated with honors receiving a Bachelor of Science degree with a major in biochemistry. During the next year he was employed as an environmental field technician at National Environmental Testing laboratory. He entered the graduate school at the University of Texas at Austin in August of 1994.

Permanent Address: 2413 Ionic Ct., Tallahassee, FL 32303

This dissertation was typed by the author.