Hydrogen-deuterium exchange studies by mass spectrometry: A method for obtaining conformational structures of proteins in solution

Teerapat Rojsajjakul

West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation


https://researchrepository.wvu.edu/etd/443

This Dissertation is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Graduate Theses, Dissertations, and Problem Reports by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.
HYDROGEN-DEUTERIUM EXCHANGE STUDIES BY MASS SPECTROMETRY: A METHOD FOR OBTAINING CONFORMATIONAL STRUCTURES OF PROTEINS IN SOLUTION

Teerapat Rojsajjakul

Dissertation Submitted to the Eberly College of Arts and Sciences at West Virginia University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in
Chemistry

Committee Members
Fred L. King, Ph.D. Chair
Ronald B. Smart, Ph.D.
Kenneth C. Showalter, Ph.D.
Patrick S. Callery, Ph.D.
Stephen J. Valentine, Ph.D.

C. Eugene Bennett Department of Chemistry

Morgantown, West Virginia
May 2014

Keywords: Hydrogen-deuterium exchange, Protein structure, Interaction, Electron capture dissociation, Electrospray ionization, Mass Spectrometry

Copyright 2014 [Teerapat Rojsajjakul]
ABSTRACT


Teerapat Rojsajjakul

Hydrogen-deuterium exchange (HDX) coupled with electrospray ionization (ESI) in mass spectrometry (MS) has become an important method in the study of protein structure and dynamics. It provides an opportunity for obtaining conformational structures of protein motions in physiological conditions and offers unparalleled limits of detection. This method relies on the fact that the accessibility of exchangeable amide hydrogens of a protein to D₂O induces H/D isotopic exchange. The exchangeable amide hydrogens that lack hydrogen bonding when exposed to solvent are exchanged rapidly. Protected amide hydrogens from tightly folded elements and have hydrogen bonding exchange slow and can be detected by mass spectrometry. This dissertation discusses basic fundamentals of both the conventional HDX ESI MS method that is called bottom-up approach that utilizes an enzymatic digestion and gas phase fragmentation method or a top-down approach to obtain hydrogen isotopic exchange at specific amide hydrogens.

Insulin, substance P and ubiquitin were used as model peptide and proteins to investigate the utility of these approaches. Bovine insulin was used to study the reduction of disulfide bonds using TECP as a reducing agent after the hydrogen isotopic reaction. The results show that the peptic fragments obtained from this method are smaller than those from the previous study without the reduction of disulfide bonds. Therefore, this approach offers more structural details and ease in identifying the peptic fragments. However, the drawback from back exchange reaction during the reduction time and chromatographic separation requires further study. Substance P was used to investigate gas phase fragmentation of the deuterated peptide by electron capture dissociation (ECD). The results show that this top-down approach is feasible for HDX MS method. The ECD fragmentation of the doubly charged ion of deuterium labeled substance P after chromatographic separation method that eliminated the rapid exchanged deuteriums at the side chains and termini produces less
scrambling. Ubiquitin was used to further investigate gas phase fragmentation of deuterium labeled protein by ECD. The surprising results from direct infusion show that the higher charge states gain more deuteriums than the lower ones. On the contrary, the gas phase fragmentation of the deuterated protein from chromatographic separation that eliminated the rapid exchanged deuteriums show the same number of deuteriums for all charge states. Thus, the results from both experiments provide the labile exchanged deuterium locations. Human insulin was used to study protein conformational change in the presence of Congo red (CR) by pulsed HDX ESI MS. CR is known as an inhibitor of insulin fibrillization by binding at the C-terminus of the B-Chain of insulin. Remarkably, the ESI mass spectrum of labeled protein in the presence of CR shows bimodal pattern, suggesting that CR binds with the partial unfolded intermediate structure of the protein. Moreover, the local HDX resulted from peptic digestion show that the conformational structure of the insulin in the presence of CR is different from that of the native form.

Overall, the results presented highlight the ability of the methods to provide insight into the conformational changes of proteins in solution.
Acknowledgements

I first and foremost would like to thank my advisor, Prof. Fred L. King for his support during my Ph.D. study. His guidance and encourage has not only led me to completing my PhD degree but has also provided me with skills and abilities to pursue a chemistry career as an independent researcher.

I would also like to thank my committee members Prof. Ronald Smart, Prof. Kenneth Showalter, Prof. Patrick Callery and Prof. Stephen Valentine for taking time to be involved in my graduate study.
# Table of Contents

CHAPTER 1. INTRODUCTION ........................................................................................................... 1

1.1 Hydrogen-Deuterium Exchange (HDX) .................................................................................. 1
   1.1.1 Fundamental and kinetics of hydrogen-deuterium exchange ........................................... 2
   1.1.2 Hydrogen – deuterium exchange dynamics of intact protein and noncovalent interactions .......................................................................................................................... 5
   1.1.3 HDX and its developments ................................................................................................ 9
   1.1.4 Hydrogen-deuterium exchange couple with mass spectrometry (HDX MS) ............... 12
      1.1.4.1. Bottom-up approach ................................................................................................. 13
      1.1.4.2. Top-down approach ................................................................................................ 17

1.2 Protein: Structures and Functions .......................................................................................... 18
   1.2.1 Primary structure ............................................................................................................. 21
   1.2.2 Secondary structure ....................................................................................................... 22
   1.2.3 Tertiary structure .......................................................................................................... 22
   1.2.4 Quaternary structure ..................................................................................................... 23
   1.2.5 Insulin ............................................................................................................................. 23
   1.2.6 Ubiquitin ......................................................................................................................... 26

1.3 Mass Spectrometry (MS) ....................................................................................................... 28
   1.3.1 Electrospray Ionization (ESI) ....................................................................................... 29
      1.3.1.1 Charge Residue Model (CRM) ............................................................................. 32
      1.3.1.2 Ion Evaporation Model (IEM) ............................................................................ 34
   1.3.2 Quadrupole Ion Trap (QIT) ............................................................................................. 36
   1.3.3 Linear Trap Quadrupole (LTQ) ...................................................................................... 39
   1.3.4 Fourier Transform Ion Cyclotron Mass Spectrometry (FT ICR MS) ....................... 41

2.1 Introduction .......................................................................................................................... 53
2.2 Experimental Section ........................................................................................................... 56
  2.2.1 Sample preparation ......................................................................................................... 56
  2.2.2 Isotope analysis by HPLC ESI MS .................................................................................. 58
2.3 Results and Discussion......................................................................................................... 59
  2.3.1 Reduction of disulfide bonds by TECP after labeling. ....................................................... 59
  2.3.2 Hydrogen exchange of bovine insulin.............................................................................. 60
2.4 Conclusions ........................................................................................................................ 66
2.5 References .......................................................................................................................... 67


3.1 Introduction .......................................................................................................................... 69
3.2. Experimental Method.......................................................................................................... 72
  3.2.1. Direct infusion method .................................................................................................. 72
  3.2.2. Separation method ....................................................................................................... 73
3.3 Results and Discussion......................................................................................................... 73
  3.3.1. Fragmentations of Substance P by Top-down ECD........................................................ 73
  3.3.2. Analysis HDX of substance P by Top-down ECD using direct infusion method. 75
  3.3.3.. Analysis HDX of substance P by Top-down ECD using HPLC, equipped with C18 .......................................................................................................................... 77
3.4. Conclusions ....................................................................................................................... 82
3.5. References ........................................................................................................................ 83

CHAPTER 4. Examining the Charge State - Protein Conformation Relationship by H/D Exchange - Electrospray Ionization and Electron Capture Dissociation Mass Spectrometry 85
4.1 Introduction ........................................................................................................... 85
4.2 Experimental Section .......................................................................................... 89
  4.2.1 Materials ........................................................................................................ 89
  4.2.2. Deuterium exchange ..................................................................................... 89
  4.2.3. Mass Spectrometry ...................................................................................... 90
4.3 RESULTS AND DISCUSSION .......................................................................... 90
  4.3.1 Unfolding of Ubiquitin in Nanoelectrospray Ionization (nESI) ..................... 90
  4.3.2 Analysis of amide HDX for individual charge states of ubiquitin in the
      conventional ESI ................................................................................................. 94
  4.3.4. MS/MS of HDX Ubiquitin by Top-down ECD FTICR MS ......................... 99
  4.3.4.1 HDX and Top-down ECD FTICR MS from direct infusion ....................... 99
  4.3.4.2 HDX and Top-down ECD FTICR MS by chromatic separation ............... 103
  4.3.5 Combining direct infusion and chromatographic separation HDX experiments . 105
4.4 Conclusions ......................................................................................................... 106
4.5 References .......................................................................................................... 108

CHAPTER 5. Conformational Structure of Human Insulin in the Presence of Congo Red by
Pulsed Labeling Hydrogen-Deuterium Exchange Mass Spectrometry ..................... 114
  5.1 Introduction ......................................................................................................... 114
  5.2 Experimental Section ......................................................................................... 117
    5.2.1 Sample preparations ................................................................................... 118
    5.2.2 HDX analysis by HPLC ESI MS ................................................................. 119
  5.3 Results and Discussion ....................................................................................... 121
    5.3.1 Pulsed labeling HDX for global analysis .................................................... 122
    5.3.2 Pulsed labeling HDX for local analysis ....................................................... 124
  5.4 Conclusion .......................................................................................................... 133
  5.5 References .......................................................................................................... 134
List of Figures

Figure 1.1. General procedure for HDX MS experiment, established by Zhang and Smith . 15

Figure 1.2. A typical amino acid structure (with the exception of proline)....................... 19

Figure 1.3. Structures and properties of the 20 common amino acids based on polarity and acid and basic properties. ................................................................. 20

Figure 1.4. Sequence structure of bovine insulin ................................................................. 26

Figure 1.5. Structure of ubiquitin ..................................................................................... 28

Figure 1.6. Schematic diagram of an electrospray ion source . ............................................ 30

Figure 1.7. The ion formation process in ESI described by CRM....................................... 33

Figure 1.8. The ion formation process in ESI described by the ion evaporation model ...... 35

Figure 1.9. Schematic of a quadrupole ion trap. ................................................................. 36

Figure 1.10. Diagram of the stability regions of ions in a quadrupole ion trap based on voltage and frequency applied. ................................................................. 39

Figure 1.11. A segmented linear trap quadrupole................................................................. 41

Figure 1.12. A schematic representation of an ICR cell..................................................... 44

Figure 2.1. The primary sequence of bovine insulin with three disulfide bonds using standard 1- letter amino acid abbreviations......................................................... 56

Figure 2.2. Procedure for continuous labeling hydrogen-deuterium exchange. .............. 57

Figure 2.3. The instrumental setup for HDX MS experiments........................................ 59

Figure 2.4. The electrospray mass spectra of a 2 mM solution of bovine insulin at pH 2.0 following reduction of disulfide bonds by TCEP at 4 °C and 0, 10 and 60 minutes. ......... 60
Figure 2.5. The structure of the T state insulin monomer derived from the X-ray crystal analysis of the two Zn\textsuperscript{2+} hexamer of bovine insulin. Yellow, red, green and blue represent amide hydrogens protected against deuterium exchange.

Figure 3.1. Structure of Substance P, including the ECD fragmentation sites. The exchangeable hydrogens at amide backbone sites are circled in red and at side chain and termini sites are circled in green.

Figure 3.2. ECD mass spectrum of [M+2H]\textsuperscript{2+} of substance P.

Figure 3.3. ECD mass spectrum of [M+2H]\textsuperscript{2+} of labeled substance P after 3 hours incubation from direct infusion method.

Figure 3.4. A plot of percent deuterium incorporation at the amide backbone of substance P against time from direct infusion.

Figure 3.5. ECD mass spectrum of [M+2H]\textsuperscript{2+} of labeled substance P after 3 hours incubation from chromatographic separation technique.

Figure 3.6. A plot of deuterium incorporation at the amide backbone of substance P against time from separation method.

Figure 3.7. The result of the exponential curve fitting from Sigma plots (San Jose, CA) of individual fragmentations of c4 to c10.

Figure 4.1. NanoESI mass spectra of native ubiquitin in 50 mM ammonium acetate at various voltages ranging from 1.8 to 2.5 kV.

Figure 4.2. ESI mass spectrum of bovine ubiquitin in 50 mM ammonium acetate and 0.4% formic acid.

Figure 4.3. ESI mass spectrum from direct infusion of native bovine ubiquitin in deuterium oxide with 50 mM ammonium acetate and 0.4% formic acid.

Figure 4.4. ESI mass spectrum of native bovine ubiquitin in deuterium oxide with 50 mM ammonium acetate and 0.4% formic acid, introduced by chromatic separation, equipped with a C4 protein trap into the ESI source.
Figure 4.5. Mass spectra of selected ubiquitin ECD fragments of c24^{3+} and c31^{4+} after 60 min of D_{2}O labeling for charge +13, +12, +11 and +10 ions. Note the different mass shifts exhibit 5 exchangeable hydrogens for the +13 charge state ion.

Figure 4.6. Structure of ubiquitin from Protein Data Bank (PDB) and UCSF Chimera. The location of the different deuteriums that are incorporated into the amide backbone of the +13 vs +11 and +10 charge states are shown in pink.

Figure 4.7. Mass spectra of selected ubiquitin ECD fragments of c24^{3+} and c31^{4+} after 60 min of D_{2}O labeling, followed by eliminating deuterium oxide and the fast exchanged deuteriums at the side chains and termini prior to the ESI, for charge +13, +12, +11 and +10 ions.

Figure 4.8. Structure of ubiquitin from Protein data bank and UCSF Chimera shown locations of the labile amide hydrogens (the difference in deuteriums between the +13 charge state from direct infusion and the +13 charge state from HPLC/C4).

Figure 5.1. Molecular structure of Congo red.

Figure 5.2. The experimental procedure for pulsed labeling HDX for insulin binding Congo red.

Figure 5.3. The experimental procedure for pulsed labeling HDX for insulin binding Congo red.

Figure 5.4. Representative ESI mass spectra of + 5 charge state of human insulin, pulsed labeling hydrogen-deuterium exchange of human insulin and pulsed labeling hydrogen-deuterium exchange of human insulin bound with Congo red.

Figure 5.5. ESI mass spectrum of peptic fragments from pulsed labeling human insulin.

Figure 5.6. Mass spectra of the molecule and despentapeptide-(B26-30) of human insulin under identical conditions show the number of deuterium incorporation for both isotropic envelopes.

Figure 5.7. Amino acid sequence of human insulin in the presence of CR displaying hydrogen exchange properties at amide backbone. Pink and Red color indicate region where significant and the most significant exchange occurred, respectively.
Figure 5.8. Conformational structures of labeled insulin and labeled insulin in the presence of Congo red.
List of Tables

Table 2.1. Assignment of the peptides from peptic digestion followed by reduction of disulfide bonds……………………………………………………………………………….61

Table 4.1. Deconvolution of charge states to average molecular mass of native bovine ubiquitin at various voltages in nESI………………………………………………………..92

Table 4.2. Deconvolution of charge states to average molecular masses of bovine ubiquitin………………………………………………………………………95

Table 4.3. Deconvolution of charge states to average molecular mass of deuterated bovine ubiquitin from direct infusion method……………………………………………..97

Table 4.4. Deconvolution of charge states to average molecular mass of deuterated bovine ubiquitin with chromatographic separation method………………………….99

Table 5.1. Deuterium levels in peptic fragments of insulin and insulin in the presence of Congo red………………………………………………………………………………….128
CHAPTER 1. INTRODUCTION

1.1 Hydrogen-Deuterium Exchange (HDX)

Hydrogen – deuterium exchange (HDX) is a powerful method for the study of protein structure and dynamics. It is very sensitive to intra-molecular hydrogen bonding and solvent accessibility. This unique method depends on the principle that the exposure of a protein to D$_2$O induces H → D exchange, providing the ability to sense a wide range of structural changes. The basic theory of HDX was well described by Woodward [1] and Englander [2]. In general, proteins contain a number of hydrogens that are unique depending on their location and environment. Rapidly exchangeable hydrogens are bonded to atoms such as O, N and S that are located on side chains and termini of the proteins or peptides. Upon exposure, these hydrogens rapidly exchange with deuterium and immediately the proteins or peptides become deuterated. Because they are labile and exposed to the surrounding deuterium oxide, their exchange rates are too fast and usually cannot be distinguished by mass spectrometric analysis. Hydrogens covalently bonded on carbon atoms do not exhibit this exchange behavior. The most important information from the hydrogen – deuterium exchange detected by mass spectrometry is obtained from the characterization of exchange of backbone amide hydrogens. The exchange rates of these hydrogens are in a range that can be measured by mass spectrometry. Additionally, every amino acid within a polypeptide chain contains a backbone amide hydrogen (with the exception of proline, which has no amide hydrogen). Therefore, the deuterium exchange behavior of individual amide hydrogens can be monitored and taken together they can represent the entire protein. In addition, these
amide backbone hydrogens are commonly involved in hydrogen bonding, creating secondary structure elements such as α-helices and β-sheets. Thus, they are indispensable for probing insight into how proteins function.

1.1.1 Fundamental and kinetics of hydrogen-deuterium exchange

The deuterium exchange rates of amide hydrogens depend on several important parameters such as temperature, pH, the presence of denaturants, solvent accessibility and hydrogen bonding [2,3]. HDX is acid and base-catalyzed, which can be described by the following equation:

$$k_{\text{int}} = k_{\text{H}}[H^+] + k_{\text{OH}}[OH^-]$$  \hspace{1cm} (1)

where $k_{\text{int}}$ is the intrinsic exchange rate constant and $k_{\text{H}}$ and $k_{\text{OH}}$ are the acid- and base-catalyzed exchange rate constant, respectively. The hydrogen – deuterium exchange rates are strongly pH and temperature dependent. The minimum exchange rate for average backbone amide hydrogens occurs at approximately pH 2.5. However, the pH can be slightly different depending on environmental effects from neighboring amino acids. Therefore, the basic-catalysis is more prevalent under physiological pH where most proteins function. Temperature also plays an important role in the exchange rate. Lower temperatures significantly decrease the deuterium exchange rate. Exposed amide hydrogens typically
exchange on the millisecond time scale at room temperature and neutral pH; at pH 2.5 and 0°C nearly an hour is required to exchange the same amide hydrogens [4].

When a protein is exposed to the deuterium oxide solution, amide hydrogens that are not involved in hydrogen bonding but exposed to the solvent are rapidly exchanged. This is referred to as deuterium exchange–in. The rate constant of this rapid conversion from N–H to N–D is referred to as $k_{ch}$ (chemical rate constant or intrinsic rate of exchange). The value of $k_{ch}$ can be determined by flanking the side chains and measuring the rates of both acid- and base-catalyzed reactions [5,6]. Amide hydrogens involved in intra-molecular N–H---O=C hydrogen bonds are less accessible to the solvent and the overall rate constants $k_{	ext{HDX}}$ (exchange rate constant) is much smaller than $k_{ch}$. The ratio of $k_{ch}/k_{	ext{HDX}}$ is sometimes greater than $10^6$ for most intact proteins [7]. Protein dynamics in solution can create short-lived transitions, called “unfolding” and “refolding” conformations. The rate constants of the ongoing events of unfolding and refolding transitions are represented by $k_{op}$ and $k_{cl}$, respectively. These parameters are also used to describe the protection against the deuterium exchange reaction. The protection against deuterium exchange imposed by protein structure is often expressed by the protection factor, $P = k_{cl}/k_{op}$, derived by comparing the HDX rate measured in a protein with the rate expected in a random coil model [3]. In general, the overall exchange mechanism of the protein in solution can be described by

$$
\begin{align*}
N-H_{\text{close}} & \xrightleftharpoons[k_{cl}]{k_{op}} N-H_{\text{open}} \\
& \xrightleftharpoons[k_{cl}]{k_{op}} N-D_{\text{open}} \\
& \xrightleftharpoons[k_{op}]{k_{cl}} N-D_{\text{close}}
\end{align*}
$$

(2)
where the combination of $k_{ch}$, $k_{op}$ and $k_{cl}$ is used to differentiate two important mechanisms referred to as EX1 and EX2 [7,8]. The term EX1 is used to describe the intermediate (partial unfold) and occur when $k_{cl} \ll k_{ch}$. Thus, two distinct populations are observed as a result

$$k_{HDX} = k_{op}$$

(3)

Most proteins under physiological conditions follow EX2 mechanism. The term of EX2 is used when $k_{cl} \gg k_{ch}$. As a result, the exchange rate constant in EX2 can be expressed by

$$k_{HDX} = \left(\frac{k_{op}}{k_{cl}}\right)k_{ch}$$

(4)

where $k_{op}/k_{cl}$ is the equilibrium constant of the opening event. It can be noted that when the opening event and closing event rates are similar, the exchange rates of these amides hydrogen are rapid and can be expressed by

$$k_{HDX} = k_{ch}$$

(5)
These unprotected amides are normally located in flexible loops. It is noteworthy that even though the pH under physiological conditions is greater than the pH of minimum deuterium exchange, the protein conformation is typically favoring the EX2 mechanism of exchange.

1.1.2 Hydrogen – deuterium exchange dynamics of intact protein and noncovalent interactions

The thermodynamic hypothesis remains important in modern protein chemistry especially in the description of small globular species. It provides a better understanding of the stepwise structural transitions of proteins that occur during their folding as well as provides clues to their functions. The folding and dynamics of a protein in solution are complicated. To reduce the complexity, often a simple two-state model represented by equilibrium transitions between the native conformation (N) and the unfolded state (U) is used to describe the behavior of the protein and can be expressed by

\[ \text{N} \leftrightarrow \text{U} \] (6)

This two-state model is reasonable for small, single domain proteins [9]. The unfolding equilibrium constant, \( K_u \), from this model is expressed by

\[ K_u = \frac{[U]}{[N]} \] (7)

where \([N]\) and \([U]\) represent concentrations at equilibrium of the folded and unfolded forms of the protein. The standard free energy of unfolding at equilibrium, \( \Delta G_u \), is given by
\[ \Delta G_u = -RT \ln K_u \] (8)

Where \( R \) is the universal gas constant and \( T \) is absolute temperature.

Thermodynamics also plays an important role in Protein – Ligand binding interactions. Under physiological conditions, the population of the unfolded form of the protein \([U]\) is much smaller than the folded form \([N]\) which results in \( K_u \) being much smaller than 1 and \( \Delta G_u \) is positive. In the presence of Ligand (L), which is capable of binding with only \( N \), the binding results in the formation of the complex \( NL \). According to Le Chatelier’s Principle, the unfolding equilibrium equation (6) becomes

\[ NL + N = U \] (9)

By substituting \([N] \) total = \([N] + [NL]\), the equilibrium constant, \( K_u \) in equation (7) becomes an apparent equilibrium constant, \( K_{app}^u \) which can be written as

\[ K_{app}^u = \frac{[U]}{[NL] + [N]} \] (10)

and can be rearranged to yield

\[ K_{app}^u = K_u \cdot 1 + \frac{[NL]}{[N]}^{-1} \] (11).
In Equation 11 $K_u$ was modified from equation (7). The binding affinity of NL is governed by the dissociation constant $K_d$. This $K_d$ can be expressed as

$$K_d = \frac{[N][L]}{NL} \quad (12)$$

where $[L]$ is the concentration of free Ligand. By substitution the expression of $K_d$, the equation (11) becomes

$$K_u^{app} = K_u \cdot 1 + \frac{[L]}{K_d}^{-1} \quad (13).$$

As a result, the apparent free energy of unfolding for the protein-ligand binding can be written as

$$\Delta_u G^{app} = -RTln \ K_u \cdot 1 + \frac{[L]}{K_d}^{-1} \quad (14)$$

or

$$\Delta_u G^{app} = \Delta_u G + \Delta \Delta_u G \quad (15)$$

Where $\Delta \Delta_u G$ is written as

$$\Delta \Delta_u G = -RTln \ K_u \cdot 1 + \frac{[L]}{K_d}^{-1} + RTlnK_u$$

$$\Delta \Delta_u G = RTln \ 1 + \frac{[L]}{K_d} \quad (16)$$
The positive term $\Delta\Delta_{u}G$ in (16) represents the ligand-induced protein stabilization resulting from the ligand binding interaction. The concentration of ligand is also critical in shifting the direction of the equilibrium. Increasing the concentration of the ligand can reduce the concentration of the unfolded form of the protein. The equilibrium constant in equation (10) can also be arranged to obtain information from hydrogen – deuterium exchange behavior and ligand binding properties, resulting in

$$K_{u}^{app} = \exp \left( \frac{\Delta_{u}G + \Delta\Delta_{u}G}{RT} \right)$$  \hspace{1cm} (17).

This equation can be used to describe amide groups of NL and N residing in N-H closed conformations. Conversely, amides in the U state are open and posed to undergo deuteration. Therefore it can be concluded that the hydrogen exchange rate constant depends on $[L]$ and $K_{d}$ according to

$$K_{HDX} = K_{u}^{app}([L], K_{d}) \times k_{ch}$$  \hspace{1cm} (18).

This simple model describes the isotopic exchange behavior of the protein as it is directly affected by the occurrence of noncovalent interactions. For example, an increase in the free ligand concentration $[L]$, may result in the reduction of $k_{HDX}$. The basic chemistry described
here provides the foundation for the use of HDX MS in protein-ligand binding interaction studies.

1.1.3 HDX and its developments

Hydrogen-deuterium exchange began with the relatively simple idea that was to determine hydrogen-bonded structural elements within proteins. That simple idea has become a powerful method that enables insight into how proteins establish functional conformations. HDX was introduced by Linderstrøm-Lang in 1954 based on density differences of the deuterium exchange on NH₄Cl [8]. A year later, Lang and coworkers demonstrated that free amide hydrogens in short peptides and insulin could be counted correctly by this deuterium exchange method. The experimental results showed that the deuterium exchange kinetics for insulin were highly complex (far from being single-exponential) and extended over much longer time than those in short peptides [9]. They also found that short peptides were completely exchanged in less than a minute at pH 3.0 and 0°C. Insulin, on the other hand, required several hours. Increasing the temperature to 38.6°C or adding 5.2 m urea the exchange became rapid like the exchange behavior observed for short peptides. The chemistry of deuterium exchange reaction for individual amide protons in short peptides was first revealed in 1970 by 1D-¹H-NMR [10]. By 1972, Molday and his coworkers were successfully measuring the effects of neighboring peptide groups and the side chains [5,6]. However, the obstacle for measuring the deuterium exchange reaction for individual amide hydrogens in proteins remained for almost another decade due to the complexity of data interpretation and the lack of advanced measurement technology. In 1979, Rosa and Richards demonstrated that proteins labeled with tritium could be digested with
pepsin under conditions that minimized hydrogen exchange and detected by radioactivity assays [11]. Later in 1985, Englander substantially refined the approach by replacing tritium with deuterium for Nuclear Magnetic Resonance analysis [2,3,12]. Hydrogen deuterium exchange combined with mass spectrometry was first published by Brian Chait in 1991, demonstrating that hydrogen-deuterium exchange of labeled intact proteins could be detected by electrospray ionization mass spectrometry [13]. However, the early studies of hydrogen-deuterium exchange aimed at monitoring specific amide backbone sites within a protein was performed by Zhang and Smith in 1993 and showed that peptic digestion was compatible with mass spectrometric analysis using fast atom-bombardment FAB [14]. Subsequently, Walsh quickly adopted the method for electrospray ionization [15].

For many decades, the rate at which deuterium in the solvent exchanges with hydrogen located on peptide linkages has been used to detect structural changes of proteins. Protein conformational changes are part of the protein life cycle from its first expression to the final post translational modification stage [16]. These changes can be monitored by conventional biophysical and analytical methods, including circular dichroism (CD) [17], tryptophan fluorescence [18], and infrared spectroscopy [19]. The experimental data obtained from far-ultraviolet CD spectroscopy provides information mainly on the secondary structural elements of a protein. Near-ultraviolet CD is on the other hand sensitive to changes in the tertiary structure around amino acids containing aromatic side chains. Although the fluorescence techniques are very sensitive for detecting such changes they exhibit low-resolving power and limited ability to focus on specific locations where conformational changes occur. In contrast, the information obtained from X-ray diffraction offers high-
resolving power (atomic level) [20,21]. However, the structural details obtaining from protein crystals may not compatible with the physiological environment in which proteins properly function. The main challenge in X-ray crystallography is that the sample preparation can be difficult to obtain good crystals, and the protein crystallization process is often time consuming. Currently calorimetric techniques [22-24] and Nuclear Magnetic Resonance spectroscopy (NMR) are the methods of choice for studying protein structure in the solution phase. NMR has been very useful because of its ability to detect H-D exchange rates at specific amide hydrogens. High resolution NMR, common to two dimensional NMR, has the capability to separate the individual amide resonance signal allowing exchange determinations at specific amide linkages. Lower resolution, one dimensional NMR, however, is limited due to the signals of amide hydrogens which are observed over a narrow frequency band and cannot be completely resolved, even in small proteins. The combination of multi-dimensional NMR and hydrogen-deuterium exchange has become an important technique for investigating structure and dynamics of proteins in solution. Examples of successful pioneering investigations include protein-ligand binding [25], protein folding and unfolding [26], as well as structural elucidation of mutants and functional variants [27]. A major drawback of using NMR to detect amide hydrogen-deuterium exchange rates is that it is unlikely to detect the most rapidly exchanging hydrogens. Slow hydrogen exchange rates are commonly used in NMR to identify the exchanging amide hydrogens that are assumed to be hydrogen bonded within secondary structure. Another challenge with this method is that the resolving power is limited by the overall size of proteins and also requires a relatively high concentration of sample. Thus, finding a suitable solvent for proteins under physiological conditions is sometimes difficult. Mass spectrometry (MS) when combined
with amide hydrogen-deuterium exchange currently plays an important role in probing the structures of proteins in solution. This hyphenated method offers high sensitivity and speed for studying global and regional conformational changes of protein in solution, regardless of size and physical properties. Moreover, mass spectrometers have developed rapidly to meet the demands for high-throughput, high resolution, sensitivity and accuracy. MS is especially suited for structural elucidation when a soft ionization technique is used to gently transform the protein molecules in solution into their associated gas phase ions. For these reasons, modern mass spectrometry (MS) coupled with electrospray ionization mass spectrometry (ESI MS) and HDX has become a powerful analytical tool for structural elucidation.

1.1.4 Hydrogen-deuterium exchange couple with mass spectrometry (HDX MS)

Proteins in solution are highly dynamic and can undergo conformational changes at any time [28]. Although the native state is the most populated under physiological conditions, the unfolded state resulting occasionally from so called breathing motions, is thought to be more flexible and often considered to be a prerequisite for proteins' function [29]. These changes can be observed by HDX MS. The study of HDX by mass spectrometry avails itself of the basic principle of the mass difference between hydrogen and deuterium, which is 1.0 Da. Dobson and coworkers first demonstrated the utility of ESI MS for hydrogen – deuterium isotopic exchange characterization of the refolding pathway of hen egg – white lysozyme [30]. The protein conformational changes detected by this method often rely on either continuous or pulsed labeling strategies [31]. The continuous labeling strategy is employed to monitor the deuterium incorporation as a function of time. At the end of each labeling time, the solution is rapidly quenched to terminate further exchange reactions. The
continuous labeling is useful for kinetic studies where time points for HDX can range from seconds to hours or even days, depending on the exchangeable amide hydrogens in proteins. Amide hydrogens of proteins that have tight folded region or closed conformation normally exchange with deuterium slowly (over a longer timescale) because their solvent accessibility is more limited. Conversely, amide hydrogen sites in disordered regions and open conformations, incorporate deuterium fast and more dynamic. Pulsed labeling is often used to detect short-lived folding intermediates resulting from protein conformational transformations in solution [32-33]. Protein dynamics may result in co-exiting species, folded and partially folded intermediate states. In this strategy, the protein is exposed to the deuterium for a short time, followed by rapid quenching to slow down the reaction. Proteins that exhibit partially unfolded conformation may have some N-H sites that remain protected. This unique dynamic protein undergoing labeling often results in a bimodal pattern representing the EX1 unfolding mechanism in mass spectrometric analysis [34]. In addition, HDX MS utilizing these strategies also provides an opportunity to maintain intact non-covalent complexes as observed from the analysis by the characteristic charge state distribution originating from proteins in solution [35-38]. There are many techniques that provide insight into conformational changes of proteins in solution using HDX MS; many of these depend on methods that break the protein into small peptides. The locations of deuteriums on the amide backbones of these small peptides reflect the surface structure of the protein that is arranged in the solution. These can be performed by methods that are termed bottom-up and Top-down MS approaches.

1.1.4.1. Bottom-up approach
The bottom-up approach is widely used in HDX MS and compatible for most mass spectrometers. It involves the enzymatic digestion of proteins in solution, followed by mass analysis of the fragmented peptides. A significant improvement for the HDX MS method developed by Zhang and Smith, is focused on the exchange of the amide backbone [14]. This technique has not only extended the utility of proteolytic digestion of partially labeled protein that allows the determination of the deuterium content in specific regions of proteins but also has opened new possibilities for studying the conformational structures and dynamics of proteins in solution under a wide variety of conditions. The general procedure for HDX MS used in this technique is shown in Figure 1.1.
In short, the labeled intact protein was quenched to terminate the exchange reaction by decreasing the pH to 2.5 and temperature to 0 °C. Under these conditions, deuterium levels of the protein can be measured to obtain the global exchange profile. To obtain the local exchange behavior, the quenched labeled protein solution was immediately digested by pepsin followed by the determination of deuterium levels of peptic peptides. These peptic
peptides from the cleavages made it possible to detect protein conformational changes with a spatial resolution of 5 – 10 amino acid residues. The incorporated deuterium on each individual fragment peptide can be measured from the centroid of the isotopic cluster in the mass spectrum. Deuterium incorporation on the amide backbone measured by this method may encounter the reversible Deuterium-hydrogen (D-H) exchange reaction during the enzymatic digestion and chromatographic separations. Thus, determination of the deuterium levels must consider the possible back-exchange reaction. The back-exchange correction can be made by assuming that deuterium loss is constant throughout the process with the overall exchange in the control sample. The correction for back exchange deuterium is written as

\[ D = \frac{(m_t - m_0)(m_{100} - m_0)}{m_0} \times N \]  

(19)

where \( D \) is the deuterium content of a peptide, composed of \( N \) amide hydrogens and \( m_0, m_t, m_{100} \) are the centroid masses of the peptide at 0, \( t \) and fully exchanged, respectively. Failure to correct the back exchange may result in a faulty conformational assessment of the protein [39]. In addition to the bottom-up approach, the deuterium exchange rate of a peptide under mimic physiological conditions can be described as the sum of a series of first-order exponentials of rate and can be estimated by

\[ D = N - \sum_{i=1}^{N} \exp(-kt) \]  

(20)
where D is the deuterium content of a peptide, containing N amide hydrogens and k is the rate constant for exchange for the peptide [40]. Nonlinear least-square fitting is often used to obtain exchange rate constants for different groups of amide hydrogens [3].

1.1.4.2. Top-down approach

Unlike the bottom-up approach, Top-down uses a different technique to localize deuterium incorporation in proteins. That is, the intact protein ions in the gas phase are selected and energized in a specific type of mass analyzer in order to obtain fragmented peptides. The top-down approach can be achieved by different fragmentation methods. Early Top-down HDX experiments employed collision-induced dissociation (CID) tandem MS (MS/MS) on the labeled protein ions resulting in fragmentation to produce the b_n and y_n homologous series of ions. Although HDX of proteins analyzed by CID MS/MS were determined to be moderately successful [41], isotopic label scrambling has become a major concern due to proton migration from protonated amino acid side chains to amide nitrogens [42]. More recent versions of Top-down methods, has focused on electron-based fragmentation methods. Electron capture dissociation (ECD) and Electron transfer dissociation (ETD) have shown potential in obtaining protein conformational changes in the solution phase [43-45]. Unlike CID, both ECD and ETD use low energy electrons interacting with the multiply-charged protein ions in the mass spectrometer to produce unique fragmentation patterns (c and z-type ions). These fragmentation methods when combined with HDX experiments have been successfully conducted without significant hydrogen scrambling [46-47]. In addition, the back exchange resulting from deuterium lost or gained during the enzymatic digestion step in bottom up approaches and in
chromatographic separations is eliminated. Although it is possible to achieve high spatial resolution (single amino acid residues) for peptides and small proteins from these electron-based fragmentation processes, the methods are unlikely to provide such resolution for highly compact and high molecular mass proteins, especially those containing disulfide bonds. This limitation has resulted in a hybrid approach between top-down and bottom-up, called middle-down [48]. In this combination approach, the ETD is used to analyze a specific peptic peptide resulting from the bottom-up approach. Thus, deuterium exchange of individual amide hydrogen in the peptide was detected. The method however encountered back exchange from the enzymatic digestion and separation steps associated with the bottom-up approach.

1.2 Protein: Structures and Functions

Proteins are one of the indispensable biological molecules consisting of amino acids as building blocks of their structure. Protein structure is based on combinations of the 20 different amino acids called amino acid sequences [49,50]. These amino acids possess common structural features including the tetrahedral alpha carbon (C\text{\textalpha}), which is covalently bonded to both the amino group, the carboxyl group, and the alpha hydrogen (H\text{\textalpha}, with the exception of proline, which does not contain the H\text{\textalpha}). It is the side chain (bonded to the C\text{\textalpha}), also called the R group that gives each amino acid its identity and unique physical and chemical properties. These common structural features make the amino acid molecules asymmetric called chiral (with the exception of glycine) resulting in non-identical mirror image isomers. The structure of a single amino acid is shown in Figure 1.2.
In neutral aqueous solution (pH ~7), the carboxyl group is ionized resulting in the –COO\(^-\) ion. The amino group is, on the contrary, protonated, resulting in the –NH\(_3^+\) ion. Thus, the amino acid molecule contains both positive and negative charges and is called a Zwitterion. Differences in the atomic composition of the side chains also play an important role in the solubility and acid-base properties of these amino acids. Thus, they are often classified based on their polarities and acid-base properties as shown in Figure 1.3. The amino and carboxyl groups bonded to the alpha carbon of all these common amino acids show similar chemical reactivity. The side chains, on the other hand, exhibit specific chemical reactivity. Under favorable conditions, the amino and carboxyl groups of the amino acids can react in a head-to-tail fashion, eliminating a water molecule and forming a peptide bond. Thus, dipeptides have two amino acid residues, tripeptides have three, tetrapeptides have four and so on. The terminology becomes inconvenient after about 12 amino acid residues; the word residue denotes portions of the peptide after the release of the water molecule to form the peptide bond. To avoid the cumbersome nomenclature, oligopeptides is often used for the peptide chains that range in size between 12 – 20 amino acid residues and polypeptides is for those containing more than 20 amino acid residues. The
The peptide bond contains the peptide backbone, which consists of the –C–N–C\(_\alpha\), where the C represents the carbonyl carbon, the N is the amide nitrogen and the C\(_\alpha\) is the alpha carbon atom. For naturally occurring proteins, the geometry of the backbone is in the trans-configuration to reduce steric hindrance from the R groups and contains only L–amino acids.

Figure 1.3. Structures and properties of the 20 common amino acids based on polarity and acid and basic properties [51].
Proteins are the biological workhorses that carry out vital functions in nature. The unique characteristics (structure and function) of each protein result from the distinctive sequence of amino acid residues. It is the amino acid sequence of proteins that is encoded by the nucleotide sequence of DNA as it contains the genetic information. In cells, the amino acid sequence is synthesized from the N-terminal end to the C-terminus. Importantly these amino acid sequences must fold correctly into a complex three dimensional structure in order for proteins to achieve their functions in the cell. Examples include proteins that catalyze biochemical reactions, regulate, transport, storage, control senses, movements, support skeleton, digest food, defend against infections, process emotions and etc. Proteins come in all shapes and sizes, depending on how their three-dimensional structure is arranged. The process of establishing functional structures is termed protein folding, and is dictated by the interactions of amino acid residues with the solvent, as well as other environmental factors such as pH and temperature. Moreover, the stability of the protein structures can be influenced by several different kinds of noncovalent interactions including: hydrogen bonds, hydrophobic, electrostatic and Van der Waals interactions. These key components, including such parameters result in the four distinct levels of protein structures.

1.2.1 Primary structure

The amino acid sequence is the primary (1°) structure of a protein. It is linear holding together by covalent bonds called peptide bonds. For many proteins, each amino acid sequence is unique and defines the structure and function of the protein.
1.2.2 Secondary structure

The secondary structure (2°) of a protein is considered to be higher level than the primary structure resulting from hydrogen bonding interactions between amino acid residues. It represents the three-dimensional arrangement of the amino acid residues comprising the polypeptide in space. Two main elements of well-defined secondary structure are α-helix and β-strand (or β-sheets) which are both stabilized by the formation of hydrogen bonds between N–H and O=C functional groups along the polypeptide backbone. These structural elements may be connected with loops that may contain partial helix or sheet characteristics, and are called turns. The α-helix is stabilized by hydrogen bonds between the carbonyl oxygen of the amino acid residue at the position n in the polypeptide chain with the amide hydrogen at the residue n + 4. Unlike the hydrogen bonds in the α-helix, the hydrogen bonds between carbonyl oxygen and the amide hydrogen of adjacent polypeptides stabilizes the β-sheet structure. In addition, the directions (according to termini locations) of these polypeptides in the β-sheets can exist in both parallel and antiparallel forms.

1.2.3 Tertiary structure

The folding of a single polypeptide chain in three-dimensional space is referred to as its tertiary (3°) structure. Therefore, the α-helices and β-sheets can be folded into a compact globular form. A major driving force for the folding into the tertiary structure is hydrophobic interactions to minimize contact between hydrophobic amino acid residues and water. In some proteins, however, disulfide bonds and the interactions of the side chains assist the proteins to fold into their stable conformations. Proteins in solution are generally flexible and
often are likely to fold into their most favorable conformations. In fact, a globular conformation gives the lowest surface to volume ratio, minimizing interaction of the protein with the surrounding environment.

1.2.4 Quaternary structure

Quaternary structure is the three-dimensional structure consisting of two or more interacting polypeptide chains with their associated tertiary structures. Each of the tertiary structures is called subunit of the protein and these subunits are held together by noncovalent interactions. More specifically, if the protein contains two subunits it would be called a dimer; a protein trimer would contain three subunits, etc. These subunits are often related to one another and can be described by symmetry operations. For instance, a 2-fold axis is used in an identical dimer protein. A protein, containing quaternary structure may undergo reversible structural changes in performing its biological function. The changes in structures of these proteins are the same as those containing only tertiary structure and are referred to as conformational changes.

1.2.5 Insulin

Insulin is a hormone protein, produced by beta cells in the islets of Langerhans in the pancreas. It is a relatively small protein with a molecular mass of 5808 Da and consists of two polypeptide chains held together by two disulfide bonds, shown in Figure 1.4 [52]. Insulin is also known as regulatory protein where its primary role is that of regulating
glucose metabolism and the use of fat as an energy source. In conditions where insulin levels are low, glucose is not taken up by the organism cells and the body begins to use fat as an energy source. On the contrary when insulin levels are high, glucose is taken up from blood and stored in the form of glycogen or fat. In addition, it also regulates amino acid uptake by the organism's cells. Because insulin controls the central metabolic pathways, failure of insulin production leads to a condition called diabetes mellitus, which can be subdivided into diabetes-type 1 and type 2.

Type 1 diabetes occurs when the production of insulin from the pancreatic beta cells fails. Thus, patients with type 1 diabetes depend on external insulin for their survival.

Type 2 diabetes results from the production of insulin from the pancreatic beta cells is inadequate, resulted from impaired receptors for insulin. In this case patients with type 2 diabetes may be treated with alternative diets or medicine to reduce blood glucose levels. They may also require external insulin if diet and medication fails to control the blood glucose levels.

In the β-cells of the Islets of Langerhans, single chain preproinsulin is converted to proinsulin, in which the A and B chains fold spontaneously into the insulin structure and then form dimers. In the absence of free metal ions, insulin exists a mixture of monomer, dimer, tetramer, hexamer and other aggregated forms, depending on the concentration [53]. In the presence of zinc ions, however, hexamerization takes place resulting in mature insulin hexamers which have a tendency to precipitate in crystalline form at pH ~5.5 in the presence of zinc and calcium ions. Upon extracellular release, the hexamers dissociate into dimers and eventually into monomers which are the biologically active species. The aggregation of
insulin taken as medication can cause problems in diabetic patients because of the slow release of insulin monomers. Today, genetically engineered insulins, which are not aggregated after the injection and quickly released into the blood stream, are being produced.

The insulin dimer is stable in aqueous solution between pH 2 and 8. The major force thought to drive oligomerization is hydrophobicity resulting from nonpolar amino acid residues mainly in residue positions B22 to B30 at the C-terminus of the B-chains. Additionally, the arrangement of residues B24 to B26 in β-strand within each monomer becomes an antiparallel in the dimers where the B13 glutamic acid residues of both monomers are forced to be more proximal. In order for dimer dissociation to occur, the electrostatic repulsion from the ionizable amino acid residues must be sufficiently large to overcome the favorable hydrophobic interactions. The conformation of insulin dimers can be different depending on the folding pathways utilized and the arrangement of the N-terminus of the B-chain. Examples of distinctive conformations of the dimers in various crystal structures classified by symmetry operations are: orthorhombic (T2-state), rhombohedral (T6-state), rhombohedral (T3R3-state), and monoclinic (R6-state).

The insulin hexamer is more stable in aqueous solution between pH 5 to 8. Under these conditions, all six B13 glutamic acid residues from three of the monomers are close together in the core of the hexamer and are, stabilized by Zn$^{2+}$ ions. Although the electrostatic interactions between Zn$^{2+}$ ions and the amino acid residues B9, B10, and B13 found in the core are crucial for the hexamerization, the stability and structural conformations of the hexamers may be altered by other parameters such as hydrophobicity, hydrogen bonding and environmental conditions. For example, replacing the B13, which provides the
negative charges in the core of the hexamers, with the neutral residue glutamine, the mutant readily forms the hexamers without the presence of the zinc or other metal ions [54]. By modifying the amino acid residues, especially in the B-chain of the insulin monomers, the hexamers are observed to pack differently in their crystals. Examples of distinctive conformational structures of the hexamers are: the T6 – conformation with (2Zn$^{2+}$); the R6 – conformation with (4 Zn$^{2+}$); and, the T3R3 – conformation (4Zn$^{2+}$) [55].

![Sequence structure of bovine insulin](image)

Figure 1.4. Sequence structure of bovine insulin [52].

1.2.6 Ubiquitin

Found in all eukaryotes, ubiquitin is a small regulatory protein consisting of 76 amino acid residues. Ubiquitin has a molecular mass of 8565 Da and forms a globular protein from a single polypeptide chain without disulfide bond. The structure of ubiquitin possesses four beta sheets interconnected by a single $\alpha$-helix and a short $3_{10}$ helix as shown in Figure 1.5
Ubiquitin influences a broad spectrum of cellular processes. It contains seven lysine residues, each of which can be conjugated by another ubiquitin to form a polyubiquitin chain that can influence the fate of target proteins [57]. The basic functions of ubiquitin were discovered in the early 1980s by Aaron Ciechanover, Avram Hershko, and Irwin Rose, who shared the Nobel Prize in Chemistry in 2004. Ubiquitin regulates other cellular proteins through its covalent attachment in the process known as ubiquitination which is the most common mechanism to label a protein for proteasome degradation in eukaryotes [58,59].

There are three other proteins in addition to ubiquitin involved in this process. These include: ubiquitin-activating enzyme (E1); ubiquitin-carrier protein (E2); and, ubiquitin-protein ligase (E3). This enzymatic cascade involving ubiquitin begins with ubiquitin activation which is followed by transferring of the activated ubiquitin to an E2 and ends with conjugation of ubiquitin to a target protein. In the ubiquitin activation step, E1 binds to the glycine residue at the C-terminus of ubiquitin through ATP-driven formation of an activated ubiquitin-adenylate intermediate. Ubiquitin is then transferred from E1 to an SH group of E2. The ubiquitin-E2 has two possible fates. It can be further transferred directly to free amino acids of the target proteins or to the free amino acid of the target proteins that selected by E3. E3 plays an important role in recognizing and selecting the target proteins for degradation. It selects the target proteins by the nature of amino acids at their N-terminus. These N-terminal amino acids must have a free α-amino terminus. Upon binding a target protein, E3 catalyzes the transfer of the ubiquitin tag from the ubiquitin-E2 to the free Lysine-ε-NH₂ position of the target protein and the C-terminus of the ubiquitin via isopeptide bond formation.
1.3 Mass Spectrometry (MS)

Progress in proteomics and biology is driven by new analytical tools that are fast, sensitive and reliable. Among them, mass spectrometry has become the method of choice for rapidly identifying proteins as well as determining details of their structures. A mass spectrometer is primarily composed of three parts: an ionization source, a mass analyzer and a detector. Much of the successful research from mass spectrometric analysis of biological compounds depends on soft ionization techniques such as Electrospray ionization (ESI) and Matrix-assisted laser desorption/ionization (MALDI). These methods can gently introduce nonvolatile biomolecules such as proteins and nucleic acids into the gas phase in the form of molecular ions without significant dissociation of covalent bonds and their molecular masses can be determined based on their mass to charge ratios (m/z). In this dissertation, electrospray ionization was used as the ionization source to ionize the proteins and peptides.
Subsequently mass analysis in either a quadrupole ion trap or a Fourier transform ion cyclotron mass spectrometer was performed.

1.3.1 Electrospray Ionization (ESI)

The phenomenon of ESI had been known for nearly one hundred years before Malcom Dole, a physical chemist at Northwestern University, and his colleagues demonstrated its potential for ionizing polystyrene from a solution containing benzene and acetone in 1968 [61, 62]. Although the result was promising, the instrumental setup was limited. Dole’s work on electrospray received attention from Seymour Lipsky and Csaba Horvath at Yale medical school, who were using HPLC to separate biological molecules. After reading Dole’s articles, Lipsky was interested in ionizing proteins and other biological compounds [63]. They noticed that two of Dole’s publications referenced Fenn’s work, focused on molecular beams and their production by nozzle-skimmer systems [64]. From the references, Lipsky contacted Fenn to produce biomolecular ions in the gas phase by electrospray ionization mass spectrometry. Later in 1984, John Fenn and Masamichi Yamashita successfully demonstrated that nonvolatile biomolecules in solution can be transformed into their molecular ions in the gas phase. This was accomplished by applying high voltage on the tip of a capillary through which a low flow rate of the solution was passed; the ions were subsequently analyzed by mass spectrometry [64]. This breakthrough research brought ESI into the analytical world and has significantly impacted many research areas. For the development of this ionization method for the identification and structural
analyses of biological molecules, John Fenn was awarded a share of the Nobel Prize in chemistry in 2002. The general instrumental setup for ESI is shown in Figure 1.6.

**Figure 1.6.** Schematic diagram of an electrospray ion source [65].

The conventional ESI process generally occurs when the high voltage (3 – 10 kV) relative to ground is applied at the sharp edged conical tip of a stainless steel needle, placed 1 – 3 mm away from the ground ion sampling orifice of a mass spectrometer while a low flow rate of the protein solution passes through the needle. The ionization is assisted by dry nitrogen gas passed through the region to sweep out solvent vapor from the evaporation droplets. Ions are typically desolvated completely and become free ions prior to mass analysis. The electric field, used in performing ESI can be estimated by the following equation:
\[ E_C = \frac{2V_c}{r_c \ln\left(\frac{4d}{r_c}\right)} \]  

(21)

where, \( V_c \) is the applied voltage, \( r_c \) is the radius of the metal capillary and \( d \) is the distance from the tip to the orifice of the mass spectrometer [66-68]. For example, if a voltage of 4 kV is applied to the metal capillary which has a radius of \( 10^{-5} \) meter and the distance from the tip to the orifice is 0.02 meter, an electric field of \( 10^8 \) volts/m is generated.

The applied voltage can be either positive or negative polarity with respect to the orifice of the mass spectrometer. The polarity is selected according to the nature of the analyte (i.e., the requirement to produce positive or negative ions). In positive mode, when the spraying nozzle is positive, protonation or sometimes metalation of the analyte molecule occurs. In negative mode, the spraying nozzle is negative and negative analyte charging is typically the product of deprotonation. For biological samples such as most proteins, both positive and negative modes are compatible but positive mode is often used due to the prevalence of basic amino acid residues in proteins and the ease with which they are protonated. Nucleic acids and carbohydrates are typically analyzed in negative ion mode because of the negative charge originating from the phosphate groups in nucleic acids and the hydroxyl groups in carbohydrate molecules. The unique characteristic of ESI mass spectra observed in both modes is the generation of intact analyte ions containing multiple charges. Multiply-charged ions from ESI can be beneficial for MS analysis. For example, the phenomenon can extend the mass range for large molecules as well as facilitate ion dissociation in tandem MS experiments.
Many assumptions and hypotheses have been made for the production of multiply charged ions in ESI [69-72]. As the solution of an analyte is transferred to the gas phase by ESI, the analyte is believed to undergo three major processes: the formation of charged droplets, solvent evaporation that leads to shrinkage of the droplets, and finally the mechanism by which the ion is formed. The first two processes are well described by the electrostatic dispersion of liquids and the creation of aerosol [73-74]. However, process involved in the formation of the multiply-charged ions is sometimes controversial. This is because a single model cannot explain all types of ion formation. The most popular models of the gas-phase ion formation in ESI described here are: the charge residue model (CRM) and the ion evaporation model (IEM).

1.3.1.1 Charge Residue Model (CRM)

Initially proposed by Dole [61] and further investigated by Schmelzeisen-Redeker [75], the charge residue model invokes a successive series of solvent evaporation and coulombic fission at the Rayleigh stability limit. The residual droplet size during the last stages of the ESI process is very small (R ∼ 1 nm), containing essentially only one molecular ion as shown in Figure 1.7 [76]. This mechanistic model works well with large molecular ions like proteins, containing multiple charging sites.
The correlation between the analyte molecular mass (M) and the average charge state $(Z_{av})$ for starburst dendrimers in ESI mass spectra results in:

$$Z_{av} = aM^b$$

(22)

where $a$ and $b$ are constants and $b = 0.53$ [76]. Based on the assumption that the density of the globular protein is the same as that of water, the radius of the protein ion in the gas-phase can be calculated by

$$\frac{4}{3}\pi R^3 \rho N_A = M$$

(23)

where $R$ is the radius of the water droplet, which is the same as that of the analyte, $\rho$ is the density of the protein and $N_A$ is Avogadro’s number [77]. A further study also found that the
combination of equations (22) and (23), including the elementary charge (e), and the permittivity of the surrounding medium ($\varepsilon_0$) yielded the Rayleigh limit charging ($Z_R$), which requires the radius of the water droplet be the same as that of the protein ion.

The new equations become:

$$Z_R \cdot e = 8\pi \varepsilon_0 \gamma R^2 \frac{1}{2}$$  \hspace{1cm} (24)$$

and

$$Z_R = 4 \frac{\pi \gamma e^2}{\rho e^2 N_A} \frac{1}{2} M^{1/2} = 0.078 M^{1/2} \hspace{1cm} (25)$$

1.3.1.2 Ion Evaporation Model (IEM)

Based on Transition state theory, the ion evaporation model was originally developed by Iribarne and Thomson to explain the production of atomic ions from randomly charged droplets produced by a spray atomizer [78,79]. Unlike CRM, the IEM does not require the production of the small droplet containing only one analyte molecule. After a series of solvent evaporation and coulombic fission, the charges on the surface of small droplets (<20 nm in diameter) contain concentrated analytes. The charge repulsion on the surface of these droplets is sufficiently strong to attain field desorption of solvated ions, yielding free analyte ions as shown in Figure 1.8.
According to the Transition state theory, the input energy from the strong electric field not only provides charges on the droplets' surface but also enlarges surface area of the droplet rapidly, causing the ions to be expelled from the droplet. The factors that influence the formation of the multiply-charged ions may result from the geometry of the analyte molecule, the reaction rate that is heavily dependent upon the chemical reaction, and the surface charge density of the droplets at the Rayleigh stability limit. This model, however, cannot be used to explain the observed evaporation rates of different ions.
1.3.2 Quadrupole Ion Trap (QIT)

Introduced by Wolfgang Paul in 1953 and first commercialized in 1984, the quadrupole ion trap mass analyzer consists of two hyperbolic metal electrodes (end caps) facing opposite one another. Located between these electrodes is a toroidal-shaped ring electrode to complete the three dimensional trap as shown in Figure 1.9 [80,81]. The end caps are electrically connected and the DC and RF potentials are applied between them and the ring electrode. The ions are confined in a three dimensional quadrupole field generated by the application of AC (oscillating, non-static) and DC (non-oscillating, static) voltages to the electrodes. The ions move axially (z-dimension) up and down while being pushed in radially (r-dimension) — and move radially back and forth while being pushed axially. If the ion excitation is needed, the AC voltage is applied to the end caps. Each end cap has a hole in the center for transferring the ions in and out.

Figure 1.9. Schematic of a quadrupole ion trap.
To select and control the movements of the ions, the fundamental RF with a fixed frequency is applied to the ring electrode while the electric fields at the endcaps are being generated. The strength of the electric field is linearly related to the distance from the center of the QIT. Under the condition that the center of the trap is the weakest electric field and the field strength increases toward the edges. The ions at the center experience a weaker stabilizing force while the ions at the edges experience a much stronger force that pushes the ions toward the center. The complex motion of ions in the QIT is described by the solutions of the equations of motion termed the Mathieu equations [82], written as:

\[ q_x = -2q_r = \frac{8eV}{m(r^2+2z^2)\Omega^2} \]  

where \( m \) and \( e \) is mass and charge of the ion and \( r \) and \( z \) are the distance from the center of the trap to the ring electrode and the endcaps along the radial and axial direction, respectively. \( \Omega \) and \( V \) represent the frequency and the amplitude of the fundamental RF voltage. The trapping of ions can be described in terms of stability regions in \( a_u \) and \( q_u \) space as shown in Figure 1.10 [82]. In most modern QIT, the parameters \( r, z \) and \( \Omega \) are fixed. Therefore, the stability of the ions is determined by \( V \) and the \( m/z \) of the ions.

The ejection of trapped ions in a QIT by increasing the trapping RF voltage on the ring electrode is widely recognized. However, the relatively high kinetic energy of the
ejected ions causes the ions to spread in different directions. This problem is resolved by adding 1 mtorr helium as a damping gas [81] to reduce the kinetic energy of the injected ions. The collisional cooling with this bath gas also improves the resolution of ions of different m/z values and the overall stability of the ions in the QIT. The mass resolution and stability of the ions are also limited by the number of ions due to space-charge effects [83]. Space-charge effects in the QIT occur when the trap is over-filled with ions, causing significant electrostatic interactions between the ions. That is ions of the same charge repel each other causing a spreading out of the ions packet which decreases mass accuracy and resolution. An insufficient number of ions, however, results in weak mass spectral signals. To improve the signal-to-noise ratio (S/N) of lower abundance ions, an accumulation scan mode is used for longer time periods prior to ejection of the ions for mass analysis.

The ejection of the ions can be performed by either mass selective instability [84] or resonance ejection [85]. By mass selective instability technique, the ions are sequentially selected and ejected out of the trap from low to high mass to charge ratio by ramping the ring electrode RF amplitude. Thus, the ions with larger m/z values require larger RF voltage to destabilize them from the stability region. In practice, the upper limit m/z to destabilizing by the RF voltage is ~650. Operating this mode for m/z values greater than 650 may result in electrical breakdown within the system. Ions having m/z ratios greater than 650 are ejected from the stability region by the resonance ejection technique. According to the solutions to the equations of motion, each ion with a specific m/z ratio has a specific secular frequency, describing the periodicity of its motion and can be tuned by the amplitude of the supplementary RF (applied to the endcaps). Thus, the ion will get excited out of the stability
region, if its secular frequency matches that of the supplementary RF. The ejection of the ions by the latter technique also improves the mass range of the QIT.

Figure 1.10. Diagram of the stability regions of ions in a quadrupole ion trap based on voltage and frequency applied.

1.3.3 Linear Trap Quadrupole (LTQ)

The linear trap quadrupole (LTQ) is similar to the QIT, originally developed by Paul and co-worker to produce and confine a plasma in 1964 [86,87]. Later in 1969, Church reconstructed the design by bending linear quadurpoles into closed-circle and racetrack geometries to store ions [88]. The LTQ can either used as a stand-alone mass analyzer or can
be combined with another mass analyzer as a mass filter, an ion guide. It is composed of a quadrupole with hyperbolic rods that cut into three segments, illustrated in Figure 1.11. It is a two dimensional (2D) trapping device; that is, the ions are confined radially by a two-dimensional radio frequency (RF) field, and axially by stopping potentials applied to end electrodes. The basis for ion trapping including collision cooling, ion excitation and ion detection in the LTQ is similar to that of the OIT. However, there are no well-defined stability diagrams of ions in a LTQ. Typically, ions are trapped at low amplitude of the fundamental RF potential and ejected to the detector by increasing the fundamental RF potential. To operate as analyzer, the assembly of rods requires a DC voltage supply to create the axial trapping field in z-direction, an RF voltage supply to deliver the radial quadrupolar trapping field in x- and y- direction and two phases of supplemental AC voltage applied across the x-rod for ion isolation, activation and mass-selective ejection. Helium gas is used as a bath gas to reduce the kinetic energies of the ions and help focus the ions to the center of the trap. In MS/MS experiments, the ions are activated and forced to collide with the target gas, He, by applying a supplementary RF voltage to induce ion dissociation. One of the advantages of linear traps is that a larger number of ions can be stored compared to the 3D trap [89]. This is because the larger volume provides higher injection efficiency and storage capacity which results in decreased space charge effects. Furthermore, linear traps can interface with other mass analyzers such as 3D ion traps [90], TOF [91], FTICR [92] and Orbitrap [93]. These hybrid MS instruments combine the sensitivity, speed, and robustness of the linear ion trap with the unique capabilities of other mass analyzers and have become popular technological platforms in proteomics.
1.3.4 Fourier Transform Ion Cyclotron Mass Spectrometry (FT ICR MS)

Introduced in 1974 by Marshall and Comisarow, FT ICR MS provides unparalleled mass accuracy and resolution among all mass analyzers [94,95]. The high mass accuracy and high resolution of this unique mass analyzer arise from the fundamental basis of FTMS. Briefly, the movement of an ion under the influence of the Lorentz force at a constant speed and perpendicular to the magnetic field is a circular trajectory. The circular motion of the ion is balanced by the centrifugal force and the centripetal force. The balance between these two forces is described by the following equation:

\[
\frac{m\nu^2}{r} = q\nu B
\]  

(28)

In Equation 28, \( m \) is the mass of the ion, \( q \) is charge of the ion, \( B \) is the magnetic field strength (tesla), \( \nu \) is the frequency (Hz) and \( r \) is the radius of the circular trajectory. The angular velocity, \( \omega \) (in rad/s) defined by...
\[ \omega = \frac{v}{r} \quad (29). \]

Substitution of \( \omega \) into (28) results in:

\[ \omega = \frac{qB}{m} \quad (30). \]

Since \( \omega = 2\pi v \)

\[ v = \frac{\omega}{2\pi} = \frac{1.535611 B}{m/z} \quad (31). \]

Equation (31) shows the relationship of the m/z ratio and the ICR orbital frequency \( (v) \). The unperturbed ion cyclotron frequency is denoted as \( \omega \) [96]. In a fixed magnetic field, the ICR orbital frequency is inversely proportional to the m/z of a given ion. In FT ICR mass spectrometry, the resolution is defined as the full width of a spectral peak at half-maximum peak height, \( \Delta\omega_{50\%} \) or \( \Delta m_{50\%} \). It is derived from the first derivative of equation (30) with respect to \( m \) as the following equation indicates:

\[ \frac{d\omega}{dm} = \frac{-qB_0}{m^2} = \frac{-\omega^3}{m} \quad (32), \]

or
\[
\frac{\omega}{d\omega} = -\frac{m}{dm}
\]  

(33).

The resolving power for frequency-domain is defined as \( \frac{\omega}{\Delta \omega_{50\%}} \) or \( \frac{m}{\Delta m_{50\%}} \) for the mass-domain. A schematic of an FT ICR cell, or Penning trap is shown in Figure 1.12. Initially, the charge on the ions of different m/z ratios experiences the magnetic field running through the center of the ICR cell and cycle in small orbits. The speeds of these ions depend upon their m/z values and the lighter more highly charged ions move faster than the heavier ones. These ions will be separated by their m/z ratios using radio frequency pulses emitted through electrodes at the excitation plates. The series of oscillation radio frequency pulses, corresponding to a specific cyclotron frequency is a function of the m/z ratio of an ion. It starts with low frequency and continues increasing. Therefore, the heavier ions respond first. A package of the ions with the same m/z ratio absorbs the energy, causing them to increase the orbit until reaching the detector plates. A stream of electrons that is proportional to the charge from the ions flow to the detector’s electrode back and forth. The Fourier Transform process begins after the radio frequency is turned off and the ions begin decreasing the energy orbit eventually reaching the original orbits.
The tremendous pressure difference between the atmospheric ionization sources and the FT ICR mass analyzer makes it difficult to directly interface atmospheric pressure ionization sources with this technique. The interface of the ESI source with the FT ICR MS instrument is achieved by using several stages of differential pumping. This is often achieved with the combination of other mass analyzers such as the linear ion trap (LTQ – Thermo Scientific). The coupling of the LTQ and FT ICR MS also enables accumulation, collision cooling and fragmentation of ions in the LTQ prior to entering the FT ICR MS [97].
1.4 References


27. Shi, J., Lua, S., Tong, J.S., Song, J. Elimination of the native structure and solubility of the hVAPB MSP domain by the Pro56Ser mutation that causes amyotrophic lateral sclerosis. *Biochemistry* **2010**, 49(18), 3887-3897.


2.1 Introduction

Hydrogen-deuterium exchange mass spectrometry (HDX MS) exploits the facility of hydrogen exchange of individual amide hydrogens to provide a powerful tool for the investigation of protein stability, dynamics, and function [1]. Some amide hydrogens on the protein backbone are protected from exchange reactions with deuterium either by intramolecular hydrogen bonding or by burial in the interior of the protein due to folding. HDX coupled with proteolytic fragmentation provides information regarding conformational structures and related properties of proteins [2,3]. Analysis of the relative kinetics for hydrogen-deuterium exchange at peptide amide linkages provides insight into the three-dimensional structures of proteins under a variety of environments [4].

In many proteins, disulfide bonds play a critical role in determining three-dimensional structure, which in turn determines biological function at the active site [5]. Most HX-MS experiments involve reduction of disulfide bonds with Dithiothreitol (DTT) or Tris(2-Carboxyethyl) phosphine hydrochloride (TCEP) before deuterium oxide (D₂O) labeling of the protein [6-7]. Unlike DTT, TCEP is a stable, water-soluble reducing agent for disulfide bonds. TCEP is widely used for reduction of disulfide bonds in peptides and proteins. The reduction of disulfides by TCEP occurs readily at low pH and inhibits thiolate-disulfide reversibility at low levels of thiolate. As a result, disulfide bonds will not reform after they are cleaved because disulfide bond reduction with TCEP occurs via an S₈2 mechanism [8].
The reduction of disulfide bonds in some proteins prior to deuterium labeling may not provide structural information relevant to their intact structures. In insulin for example, the reduction of disulfide bonds yields a peptide A-chain, composed of 21 amino acids and a B-chain, composed of 30 amino acids [9]. Reduction of disulfide bonds after deuterium labeling, however, can provide relevant structural information. Although the latter method is more appropriate for the reduction of disulfide bonds by DTT at low pH of the labeling condition is impossible. TCEP affords faster reduction of disulfide bonds than DTT over a wide range of pH and is especially effective at low pH [10]. Moreover, TCEP reduces disulfide bonds quickly and quantitatively in water at pH 5 [11] and can be used in HDX MS experiments to reduce disulfide bonds after deuterium labeling under aqueous conditions at lower pH. Bovine insulin can be studied by HDX MS under conditions that are known to lead to its aggregation to form amyloids without the reduction of disulfide bonds [12]. Such studies reveal that after peptic digestion disulfide bonds link two or three peptic fragments together, resulting in relatively large peptides and making it difficult to locate the deuterium exchange sites. To obtain the details of amide proton exchange sites in the native structure of the insulin, these disulfide bonds must be reduced after labeling with deuterium oxide. This approach provides smaller peptic fragments, for which it is easier to identify amides that undergo proton exchange.

The reduction of disulfide bonds after labeling with deuterium in HDX MS was first reported for homodimeric recombinant human macrophage colony-stimulating factor beta (rhM-CSF), a protein containing nine disulfide bonds [13]. These disulfide bonds stabilize the three dimensional structure and biological activity of rhM-CSF [5]. The reduction of the
disulfide bonds in rhM-CSF was performed by adding TCEP into the solution after deuterium labeling but before pepsin digestion to reduce the disulfide bonds and obtain smaller peptic fragments. In bovine insulin, the complete reduction of disulfide bonds at pH 2.5 by TCEP and at low temperature can take longer than 10 minutes. Therefore, the reduction of disulfide bonds prior to the digestion requires a significant amount of time and may lead to the back exchange of deuterium label. To reduce the time, and minimize the back exchange, digestion followed by TCEP reduction after the labeling can be considered as an alternative option. Adding TCEP at the same time as enzymatic digestion was recently carried out for studies of Myoglobin, Hepatitis-C Virus and sRAGE [14]. The results show that this method increases proteolytic sequence coverage for proteins with disulfide bonds. However, the activity of the proteolytic enzyme can be degraded. For example, the study shows that the activity of protease type XIII is slightly degraded in the presence of denaturing and reducing agents because it has one disulfide bond, which is far from the active site. Pepsin activity can, unfortunately, be degraded more because it has three disulfide bonds. The work presented here extends the information obtained from HDX MS of bovine insulin by adding TCEP to cleave disulfide bonds after peptic digestion. Peptic digestion was carried out by adding porcine pepsin solution into the insulin solution after labeling with deuterium to cleave peptide bonds and increase surface accessibility for TCEP to perform its function.
2.2 Experimental Section

Bovine insulin, porcine pepsin, acetonitrile (HPLC grade), formic acid (95%), and deuterium oxide (99%) were purchased from Sigma (St. Louis, MO). Hydrochloric acid and Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) were purchased from Fisher Scientific (Pittsburgh, PA). TCEP solution (1.70 M in 50% H$_2$O/D$_2$O) was stored in an ice bath prior to use.

2.2.1 Sample preparation

The general procedure used in this experiment was described in a previous study [10]. Briefly, HDX of insulin was initiated by dissolution of lyophilized insulin (1 mg/mL in deuterium oxide) at pH 2.0 and 22 °C. Peptic digestion was carried out by adding 10 μL of pepsin solution (1 mg/mL in deuterium oxide) at 4 °C into 10 μL of the insulin solution at 22 °C (20 mM at pH 2.0). Following incubation of this solution at 4 °C for one minute, 20 μL of TCEP (10 mg/20 μL of 50% H$_2$O/D$_2$O) was added to give a final concentration of 0.85 M.
TCEP. This solution was incubated at 4 °C for 3 minutes. The time course for peptic digestion was 0, 1, 2, 5, 10, 20, 30, 60, 120, 180, 240, 300 and 1440 minutes. All samples were stored in liquid nitrogen until analyzed by HPLC MS. The procedure for continuous labeling hydrogen-deuterium exchange in this experiment is shown in Figure 2.2.

Figure 2.2. Procedure for continuous labeling hydrogen-deuterium exchange.
2.2.2 Isotope analysis by HPLC ESI MS

Deuterium levels in peptic fragments were determined by HPLC electrospray ionization mass spectrometry (ESI MS). For analysis of the peptic fragment of bovine insulin, the HPLC system consisted of two pumps (Hitachi, L-7100), an injector (Rheodyne 8125, Altech, Deerfield, IL) equipped with a protein trap (Michrom BioResources, Inc., CA) and a C-18 reversed phase column (Discovery Bio Wide Pore, C18, 3 μm particle size, 2.1 mm x 100 mm). The instrument setup for this experiment is shown in Figure 2.3. Following injection, the protein digest was concentrated in the peptide trap, and subjected to a ten minute gradient elution (solvent B, acetonitrile, 0-85%) at a flow rate of 150 μL/min. Both solvents included 0.3% (v/v) formic acid. All components of the HPLC system from the injector and column were submerged in ice. The ESI MS data presented here were acquired using an ion trap mass spectrometer (LCQ, Thermo, San Jose, CA). Peptides of bovine insulin were identified by their exact masses. The average mass of each peptic fragment was determined using the software package MagTran [15]. The deuterium content of the peptides was calculated by using the difference between the average signal from deuterated and undeuterated peptides. Analysis of the rates of exchange was carried out using Sigma Plot (San Jose, CA).
2.3 Results and Discussion

2.3.1 Reduction of disulfide bonds by TCEP after labeling.

The mass spectra obtained for a 2 mM solution of insulin at pH 2.0 following reduction of disulfide bonds by TCEP at 4 °C and 0, 10 and 60 minutes are shown in Figure 2.4. Even though bovine insulin is a relatively small protein, the reduction of disulfide bonds in bovine insulin under standard HDX conditions takes longer than 10 minutes. It is possible to reduce the time of analysis by inducing disulfide bond reduction after enzymatic digestion. The reason for this improvement is two fold. First it enables pepsin digestion to occur in the absence of the reduction of pepsin itself arising from the presence of TCEP. Second, the interaction of TCEP with disulfide bonds is more facile for the digested fragments than for the intact protein because of more surface accessibility.
2.3.2 Hydrogen exchange of bovine insulin.

In this study, the chromatographic separation was used to remove fast exchanged deuteriums at the side chains, N and C-termini as well as to separate peptide fragments from the peptic digestion solution. The advantage of this method is that only exchanged amide hydrogens at protein backbone sites are observed. Unfortunately, the additional time and solvents required for this approach may significantly affect the exchange rates of the amide hydrogens of interest. In a previous study, hydrogen exchange coupled with electrospray ionization mass spectrometry provided the elucidation of bovine insulin structure under amyloid fibril formation conditions [12]. That work indicated the influence of three disulfide bonds in bovine insulin, two inter-molecular links between A7 - B7 and A20-B20 residues, and one inter-molecular link from residues A6-A11 using two or three peptic peptides. For
example, peptic digest fragment D, comprised of residues A14-21 and B17-25, has one disulfide bond, linked between A20 and B20, resulting in a total of 17 amino acids in this fragment. Similarly, peptic digest fragment P consists of residues A1-21, B1-11 and B17-25 resulting in a total of 41 amino acids. In HDX MS, the optimal size of peptides for the location of exchanged amide protons is ~5-6 amino acids [16,17].

In the present study, the disulfide bonds of bovine insulin are cleaved by TCEP solution after deuterium oxide labeling and pepsin digestion. Thus, the number of amino acids in each peptic fragments is closer to the optimal number cited above. For example, the peptic digest fragment D from the previous study is now separated into two fragments, A14-21 and B17-25, which have 8 and 9 amino acids, respectively. Similarly, the peptic digest fragment P is separated into three fragments, A1-21, B1-11 and B17-25 that have 21, 11 and 9 amino acids, respectively. The resulting assignments and kinetic information obtained with peptic digestion following TCEP reduction are shown in Table 2.1.

Table 2.1. Assignment of the peptides from peptic digestion followed by reduction of disulfide bonds.

<table>
<thead>
<tr>
<th>Peptic digest fragment No.</th>
<th>Residue</th>
<th>Assignment</th>
<th>Exchange rate Intermediate/min</th>
<th>Exchange rate Slow/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A1-8</td>
<td>GIVEQCCA</td>
<td>1.51x10^{-1} (2.48)</td>
<td>1.2x10^{-2} (2.54)</td>
</tr>
<tr>
<td>2</td>
<td>A9-13</td>
<td>SVCSL</td>
<td>1.03</td>
<td>4.50x10^{-3}</td>
</tr>
<tr>
<td>3</td>
<td>A1-13</td>
<td>GIVEQCCASVCSL</td>
<td>(0.89)</td>
<td>(0.97)</td>
</tr>
<tr>
<td>4</td>
<td>A14-21</td>
<td>YQLENYCN</td>
<td>2.20</td>
<td>3.33x10^{-2}</td>
</tr>
<tr>
<td>5</td>
<td>A15-19</td>
<td>QLENY</td>
<td>(4.21)</td>
<td>(0.82)</td>
</tr>
<tr>
<td>6</td>
<td>B1-11</td>
<td>FVNQHLCGSHL</td>
<td>1.20x10^{-1}</td>
<td>3.00x10^{-4}</td>
</tr>
<tr>
<td>7</td>
<td>B1-13</td>
<td>FVNQHLCGSHLVE</td>
<td>2.18</td>
<td>8.00x10^{-3}</td>
</tr>
<tr>
<td>8</td>
<td>B7-14</td>
<td>CGSHLVEA</td>
<td>9.8x10^{-1}</td>
<td>6.20x10^{-3}</td>
</tr>
<tr>
<td>9</td>
<td>B17-25</td>
<td>LVCGERGFF</td>
<td>1.62</td>
<td>7.40x10^{-3}</td>
</tr>
<tr>
<td>10</td>
<td>B17-26</td>
<td>LVCGERGFFY</td>
<td>1.28</td>
<td>4.10x10^{-3}</td>
</tr>
<tr>
<td>11</td>
<td>B25-30</td>
<td>FYTPKA</td>
<td>(2.95)</td>
<td>(2.36)</td>
</tr>
<tr>
<td>12</td>
<td>B26-30</td>
<td>YTPKA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The peptic digest fragments in Table 2.1. are generally 2 to 3 times smaller than those in the previous study because the disulfide bonds are reduced. The exceptions, peptic fragments 11 and 12, arise because these peptides do not have disulfide bonds. The smaller digest fragments numbered 1 to 10 in Table 1 facilitate the location of the amide sites subject to exchange. This study also provides kinetic information for the exchange of amide protons at different sites. The present results are based on the double exponential decay function described in Robinson’s previous work [12]. The data reveal that the average rate constants are fast > 10 min\(^{-1}\) in the initial phase, in agreement with Robinson’s results. Because the number of amino acids contained in peptic fragments is smaller than those utilized in the previous study, the exchange rates of amide protons also differs from those observed for the longer peptic fragments that are linked by disulfide bond(s). The overall results for the smaller fragments reveal intermediate phase and slow phase rates ranging from 2.18 to 1.20 \times 10^{-1} \text{ min}^{-1} \text{ and } 3.00 \times 10^{-4} \text{ to } 1.20 \times 10^{-2} \text{ min}^{-1}, respectively. These average rates of amide hydrogen exchange contribute to the hydrogen exchange protection from solvent accessibility of the peptides in both the A-chain and B-chain of bovine insulin.

There are 9 slowly exchanging amide hydrogens in the A-chain. The assignments for 5 of these amide hydrogens is at the A2-7 sites of the peptic peptide comprised of residues A1-8 which is in agreement with Robinson’s results [12]. A previous NMR study [18] also indicates that approximately 5 (±1) of the sites in helix A2-A8 are protected amide hydrogens, again consistent with the present results. The present studies indicate that 2 additional protected amide hydrogens are located in the A10-12 region of the peptic peptide comprised of residues A9-13. This result is different from both Robinson’s results and the
previous NMR results. In fact, there are two different rates of exchanging amide hydrogens in this A8-12 turn consisting of one intermediate and one slow exchanging amide hydrogen. The remaining 2 protected amide hydrogen sites are in the A15-20 region of the peptic peptide comprised of residues A14-21. This assignment is in agreement with both the Robinson and NMR studies. There are also 9 (+1) protected amide hydrogens located in the B-Chain. These protected amide hydrogens can be assigned as 4 exchanging amide hydrogens in the B2-10 region, 1 exchanging amide hydrogens in the B11-12 region, and 2 exchanging amide hydrogens in the B8-10 region or one at residue B13. The other assignments for 2 and 1 exchange amide hydrogens are in the B18-24 region and at residue B25, respectively. The fast exchange amide hydrogens in peptic fragments B25-30 and B26-30 are consistent with Robinson’s results, indicating that this region of insulin is readily accessible. The results for these protected amide protons under amyloid fibril forming conditions of bovine insulin are shown in Figure 2.5.
Figure 2.5. The structure of the T state insulin monomer derived from the X-ray crystal analysis of the two Zn$^{2+}$ hexamer of bovine insulin using Chimera, UCSF, CA. Red, yellow, green and blue represent amide hydrogens protected against deuterium exchange.
2.4 Conclusions

HDX MS is extended to the study of bovine insulin under amyloid fibril forming conditions by adding TCEP solution to reduce disulfide bonds of bovine insulin after labeling with deuterium oxide and digesting with porcine pepsin solution. This method decreases peptic segment lengths and improves the residue-level information of HX-MS for bovine insulin. This approach is also beneficial for proteins with disulfide bonds that cannot be reduced prior to labeling with deuterium oxide. The identification and characterization of amide sites subject to HDX on the A and B chains of bovine insulin achieved by this method is generally consistent with previous reports of the structure of insulin. The technique did enable the identification of two additional protected sites on the A-chain at the residue positions of A10-12.
2.5 References


3.1 Introduction

Amide hydrogen-deuterium exchange (HDX) coupled with mass spectrometry (MS) is widely used for study of protein dynamics [1-6]. This is because the deuterium exchange on the amide backbone reflects conformational structure of the protein in solution. The analysis of HDX by MS is based on the mass differences between those of deuterated and undeuterated peptides or proteins. These differences in degree of incorporative deuteriums on individual amide backbone describe solvent accessibilities at the surface of peptides or proteins and provide insight into conformational structures of the proteins in solution. Many HDX studies were successfully obtained not only information of protein folding and dynamics but also protein – ligand interactions [7-8]. In the conventional HDX method, monitoring exchange rates of local amide protons in solution depends upon enzymatic fragmentation (bottom - up) that allows to localize the exchanged deuterium on the individual amide backbone of the protein. This can be done by either kinetic study where the deuterium exchanged rates are detected over time or equilibrium study where the experiment is carried out at a given condition to study intermediate folding and refolding of the protein. In general, rapidly exchanged protons are bonded to atoms, such as O, N and S., often located at the side chains and termini of the protein. Both experiments involve incubation of protein in deuterium oxide solution for period of time, immediately followed by the addition of a quench solution at pH around 2.5 and decreasing temperature to 0 °C to decrease the reaction rate. To obtain the exchange of individual amide backbone, proteolytic enzyme,
often pepsin because it works well in this condition, is immediately added to the quenched sample solution followed by chromatographic separation HPLC equipped with reversed-phase column (C18) to wash out fast exchanged deuterium at the side chains and N-termini of the peptides, remaining only deuterium at the amide backbones [1]. This approach offers not only focusing on the amide backbones but also avoiding influences from fast exchanged deuteriums at the amide side chains and N-termini, also called deuterium scrambling. Typically, the peptic fragments, containing 5 – 10 amino acid residues, including overlapping peptides, are well suited for the HDX method. These overlapping peptides can be used to obtain the exchanged rate of individual amide proton. The resolution of the bottom-up HDX information obtained depends upon the size and numbers of the overlapping peptides. This approach, however, may be hindered by deuterium back exchange during the separations. In addition, the activity of proteolytic enzyme may not meet the criteria, especially for proteins, containing disulfide bonds. The disulfide bond plays an important role in producing small peptides [9]. Most HDX experiments were carried out by reducing disulfide bonds prior to labeling with deuterium oxide. The information from these experiments may not correctly provide higher order structure of protein dynamics due to the role of disulfide bond. Another drawback of the bottom-up approach is deuterium back exchange that may occur throughout the digestion and separations steps [10]. Often HDX by bottom-up approach are carried out under restricted conditions to minimize the back exchange that may affect the deuterium level on the amide backbone.

Top-down approach in HDX has emerged as an alternative method that showed potential for obtaining HDX at single amino acid residue for peptides and small proteins.
Top-down uses a different approach to localize deuterium incorporations in peptide or protein. Namely, the intact protein ion in gas phase is selected and fragmented in the mass analyzer to obtain fragmented peptides. Top-down approach can be achieved by different fragmentation methods. The early Top-down HDX experiment employed collision induced dissociation (CID) tandem MS (MS/MS) on the labeled protein ions, resulted in fragmentations, \( b_n \) and \( y_n \) ions. Although HDX of proteins analyzed by CID MS/MS were employed but hydrogen scrambling has become a major concern due to proton migration from a protonated amino acid side chain to amide nitrogen [12,13]. The later version of Top-down, focused on electron based methods, Electron capture dissociation (ECD) and Electron transfer dissociation (ETD), have shown potential in obtaining protein conformational changes in the solution phase [14-16]. The top-down approach may serve several advantages, including high spatial resolution, less deuterium back exchange, low deuterium scrambling and can also reduce disulfide bond.

ECD is an ion fragmentation technique in FTICR mass spectrometry that has emerged as a technique for protein identification and characterization. The ECD fragmentation offers spatial resolution at single amino acid, obtained from fragmentations of c or z ions. These ions were generated by colliding electrons and protein ions in the ICR cell. The radical fragmentation reaction at C-N bond along a peptide chain results in backbone dissociation of covalent bonds without significant dissociation of tertiary noncovalent bonds. The deuterium back exchange is minimized in most HDX ECD because it requires less time to separate fragments of proteolytic peptides in chromatographic process. In addition, the exchanged deuteriums at the amide backbone are less perturbed from solution into the gas-phase.
However, many research studies show the back exchange is minimized but scrambling may occur in the elucidation of protein structure and dynamics. Here we employed substance P as a model peptide to compare results from HDX ECD FTICR MS with and without chromatographic separation. The aim of the chromatographic separation method is to replace the fast exchanged deuteriums at side chain and N-termini back with hydrogens. These experimental results provide useful information for HDX Top-down methods.

3.2. Experimental Method

Substance P, Acetonitrile (HPLC grade), Formic acid (95%), and Deuterium oxide (99%) were purchased from Sigma, St. Louis, MO.

*Sample Preparation:* Triplicate samples of Substance P for HDX were prepared by adding 10 µL of Substance P from a stock solution (1 mg/mL) in deionized water at pH 2.5 and 4 °C into 1 mL of deuterium oxide (D$_2$O) at pH 2.5 and 0 °C. Sampling of this solution was conducted at 0°C for 1, 5, 10, 20, 30, 60, 120, 180, 240, and 360 minutes. The experiments composed of two methods.

3.2.1. Direct infusion method

The samples were analyzed by direct infusion into LTQ FT (Thermo, San Jose, CA.), equipped with ESI and ECD. The capillary voltage was 3.5 kV and the capillary temperature was 120 °C. The ECD heater current was 0.96 A. and the isolation width was 20.
3.2.2. Separation method

The samples were eliminated the deuteriums at side chains and N-termini by HPLC, utilizing a C18 trap (Michrom BioResources, Inc.) prior to the LTQ FT. Following injection, Substance P was concentrated in the peptide trap, and subjected to a 500 µL of H$_2$O at pH 2.5 and 4°C washed, followed by isocratic gradient elution (solvent B, acetonitrile, 70 %) at a flow rate of 150 µL/min into ESI. Both solvents included 0.1% (v/v) formic acid. All components of the HPLC system from the injector and the peptide trap were submerged in ice. ESI-MS and ECD parameters were the same as in 3.2.1. The mass spectra of Substance P were identified by their exact masses and centroid masses. The labeled peptides were used to determine the deuterium contents of the peptides from c$_4$ to c$_{10}$. Using doubly charged intact peptide ions, the average number of deuterium incorporation of amide proton N was determined by (average number of deuterium incorporation in c$_{n+1}$) – (average number of deuterium incorporation in c$_n$). In the experiments, including deuteriums at the side chains and termini, residue one did not have a backbone amide but carried 6 exchanged hydrogen sites (three at the side chains and 2 at N-terminus ) and the first c$_1$ is in residue 2, shown in Figure 1. The experiments excluding deuterium on side chains and N-termini in 3.2.2, focused only amide hydrogen on the backbone. All measurements were triplicated with error bars representing standard deviations.

3.3 Results and Discussion

3.3.1. Fragmentations of Substance P by Top-down ECD
Substance P is a neuropeptide with molecular mass of 1346.74 Da. It composes of eleven amino acid sequence; Arg Pro Lys Pro Gln Gln Phe Phe Gly Leu Met (RPKPQQFFGLM) commonly used for the calibration of ECD fragmentation technique in LTQ FT mass spectrometer. ECD fragmentations for c and z' positions resulted from dissociations by electron are shown in Figure 3.1 and the result of ECD mass spectrum for the doubly charged ion of Substance P is shown in Figure 3.2. Substance P contains 22 exchangeable hydrogens (8 at backbone amides, 10 at side chain sites, 4 on the termini). The doubly charge ion, however, contains two additional hydrogens on charge carriers at the side chains).

Figure 3.1. Structure of Substance P, including the ECD fragmentation sites. The exchangeable hydrogens at amide backbone sites are circled in red and at side chain and termini sites are circled in green.
3.3.2. Analysis HDX of substance P by Top-down ECD using direct infusion method.

In this study, substance P was subjected to continuous labeling with deuterium oxide and the labeled peptide solution was introduced into the ESI source by direct infusion method. Fourier Transform mass spectrometer for HDX by Top-down ECD of the doubly charged ion of substance P was subsequent mass analysis and an example is shown in Figure 3.3. The results from the triplicate experiments indicate that maximum $22 \pm 2$ deuteriums incorporated on the peptide and $7.6 \pm 1.0$ deuteriums (up to c10) incorporated on the amide backbone of the substance P. This result was calculated based on the assumption that all fast exchanged hydrogens at the side chains and N-termini of substance P were completely exchanged with deuterium.
The HDX kinetic analysis of substance P from 0 min to 3 hours by the direct infusion method shows that the deuterium incorporation increases over the time interval and completely deuterated in three hours. The plot between the percent of deuterium incorporations against time from triplicate data sets is shown in Figure 3.4. The result shows large standard deviations for each time points, suggesting that hydrogen scrambling occurred in the process.
Figure 3.4. A plot of percent deuterium incorporation at the amide backbone of substance P against time from direct infusion.

3.3.3. Analysis HDX of substance P by Top-down ECD using HPLC, equipped with C18

In this technique, chromatographic separation was used to eliminate fast exchanged deuteriums and replace back with hydrogens at the side chains and termini, remaining only slow exchanged deuteriums at the amide backbone of substance P. These exchanged deuteriums are slow and can be detected and observed in the ESI mass spectra. Fourier Transform mass spectrometer for HDX by Top-down ECD of the doubly charged ion of substance P was subsequent mass analysis. The results from ESI triplicate mass spectral analysis of [M+2H]^{2+} of the labeled substance P show the maximum of 7.5 ± 0.4 deuteriums
(up to c10) incorporated into the amide backbone of the peptide shown in Figure 3.5. The results were calculated based on the assumption that all fast exchanged deuteriums at the side chains and N-termini of substance P are replaced by hydrogens, remaining only the deuteriums at the amide backbone.

Figure 3.5. ECD mass spectrum of [M+2H]^{2+} of labeled substance P after 3 hours incubation from chromatographic separation technique.

The kinetic analysis of the labeled substance P from 0 min to 3 hours resulted from the separation method shows that deuterium incorporation increases over the time interval and completely deuterated in three hours. The plot between the percent of deuterium incorporations against time from triplicate data sets is shown in Figure 3.6. The plot shows
small standard deviations for each time points, suggesting that there is insignificant hydrogen scrambling occurring in the process.

Figure 3.6. A plot of deuterium incorporation at the amide backbone of substance P against time from separation method.

The results from both direct infusion and separation methods indicate that the HDX from the separation method provides small standard deviation and better exponential curve fitting. Thus, this method was chosen for further Top-down ECD analysis. Top-down ECD was performed on the HDX to localize the deuterium at the individual amide backbone. In HDX ECD, the spatial resolution at single amino acid can be obtained from fragmentations
of c or z ions. In this study, HDX ECD of the doubly charged ion, [M+2H]\(^{2+}\), of substance P, generated by ESI showed the maximum of exchange occurs respectively for 9 and 22 proton sites. The maximum number of amide protons exchanged with deuterium up to c10 for both experiments was 8. The results obtained from the separation method that excluded the fast exchange deuteriums at the side chains and termini yielded better exponential curve fitting. The plot of the measurements without the side chain influence exhibited an exponential rise to the maximum and less deuterium migration within the peptide chain. The mass spectrometric data showed different solvent accessibilities of cleavage sites on Substance P, suggesting that the structural information of this peptide can be studied by Top-Down ECD FTICR MS. The incorporation of HPLC separation, using a C18 peptide trap, provided less H/D scrambling and minimal amide back exchange. The results from exponential curve fitting using Sigma plots (San Jose, CA) of individual fragmentations of c4 to c10 are shown in Figure 3.7.
Figure 3.7. The result of the exponential curve fitting from Sigma plots (San Jose, CA) of individual fragmentations of c4 to c10.
3.4. Conclusions

In top-down approach, intact peptide ions are introduced into gas phase by ESI and subsequently fragmented in the mass spectrometer, yielding the molecular masses of the fragment ions. In the case of HDX of substance P, hydrogen scrambling is more likely to occur in the ESI process. Therefore, the deuterium incorporations of the substance P from direct infusion show more deuterium scrambling. The results from the chromatographic separation show less deuterium scrambling from the ESI process. Moreover, the ECD fragmentation also shows less deuterium scrambling.
3.5. References


4.1 Introduction

Hydrogen-deuterium exchange (HDX) coupled with ESI Mass Spectrometry (MS) has become a powerful method for the study of protein structure and dynamics [1-3]. However, the relationship between the charge state and protein conformation continues to be debated [4-13]. Early studies of electrospray ionization indicated that noncovalent binding was preserved in the gas phase and the deuterium exchange rates of backbone amides in solution, obtained from ESI MS, was extremely sensitive to conformational changes of proteins. This information was used to provide insight into conformational changes and dynamics of proteins in solution [14-17]. In addition, the charge state distribution (CSD) of a protein produced from the ESI is dependent on the nature of the protein and factors that affect the degree of ionization on the protein surface, such as pH, temperature, and solvent composition. When a protein denatures, buried ionizable groups become accessible to protonation, leading the CSD to shift to higher charge states. It is not always clear, however, the extent to which the variation originates from the conformation changes in solution or from the ESI process or from a combination of the two. This makes it difficult to correlate the conformations of a protein in the solution and gas phases. Often in HDX experiments it is assumed that the exchange takes place in solution [18,19]. Many studies suggest that the protonation responsible for the creation of higher charge states may continue during the ESI process and not be related to the solution phase conformation [15,20]. The evaporation of
solvent from the droplets in the ESI process increases surface charge density, leading to smaller charged droplets. These smaller charge droplets later repeat the same process, driving off surface ions to finally yield solvent-free ions in the gas phase [3]. This mechanism implies that a labeled deuterated protein may gain or lose deuteriums during the electrospray process.

Ion mobility - mass spectrometry (IM-MS) studies of specific ubiquitin charge states in the gas phase present conflicting results. Clemmer’s group found that the +7 charge state underwent extensive unfolding in the gas phase [21]. A subsequent report from Bowers’ group concluded that extended unfolding of the native state was not observed during the ESI process [22] for low charge states. This discrepancy resulted in further investigations by IM-MS, Molecular dynamics (MD) simulations and, ECD [23-25]. The results from the additional IM-MS and MD simulations from Robinson’s group show partial collapse of the GroEL complex in the gas phase. It appears that the balance between hydrophobic and electrostatic interactions is also crucial for the stability of complexes in the gas phase [23,24,26]. These results are in agreement with the ECD experimental results, demonstrating that ubiquitin is unstable in the gas phase and protein unfolding can occur during solvent loss in the ESI process [25]. Overall, the results from the multiple experiments suggest that the lower charge states of a protein are compact and more likely to preserve the native-like solution structure. Conversely, the higher charge states, may be subjected to further unfolding during the ESI process. Notably, chemical reactions that take place in the electrospray droplet, which can cause extensive unfolding of a protein in the ESI process, are also possible.
It is believed that the rate of deuterium incorporation within various proteins in HDX MS studies is sensitive to protein conformational changes as well as solution conditions [14-15]. An early HDX study by direct infusion and ESI MS analysis found that the rates of deuterium incorporation for CSD’s of ubiquitin and cytochrome c are different. The conformation structures of higher charge states, especially the highest, exchange faster than those of the lower charge states [14,19]. Although deuterium incorporation during unfolding in the ESI process is possible, the evaluation of such reactions in electrospray droplets remains difficult [4]. As a result, HDX combined with ESI is performed based on the assumption that all the charge states have the same exchange rate [27-35]. HDX experiments employing enzymatic digestion to localize the exchange rates of hydrogens at the amide backbone were often subjected to back exchange. Here, as the protein was cleaved to smaller fragments, the fragments did not have as many charges as the original protein [35-40]. The use of chromatographic separations to eliminate the influence of excess deuterium oxide and fast deuterium incorporation at the side chains and termini extends the utility of HDX MS [16]. The method enables the same deuterium incorporation for all observed charge states in ESI because of the slow exchange under quenched conditions of deuteriums on the amide backbone. More importantly, each of these charged ions should present evidence for the same solution conformation when analyzed by top-down methods.

Top-down HDX is considered as an alternative structural analysis method in which spatial resolution can be achieved at the single amino acid level. Single amide HDX experiments for peptides and small proteins have been reported for top-down MS/MS methods such as electron capture dissociation (ECD) [41-45] or electron transfer dissociation
(ETD) [46-50]. These top-down methods afford several advantages including, high spatial resolution, minimized deuterium back exchange, and low deuterium scrambling. ECD is a powerful method for top-down protein identification and structural characterization and HDX ECD provides spatial resolution at the single amino acid level from ion fragmentation to form c- or z-type ions [51-52]. Because these ions arise from low energy electrons, electron capture promotes a radical directed fragmentation at C-N bonds along a peptide chain, yielding backbone dissociation of covalent bonds without significant dissociation of tertiary noncovalent bonds. Under these conditions, the deuterium back exchange is minimized and the exchanged deuteriums at the amide backbone are less perturbed as the protein goes from solution into the gas-phase.

The studies reported here employed HDX coupled with nanoelectrospray ionization under native conditions (nESI) and the conventional FTICR MS under quenched conditions typically used in HDX experiments, to investigate unfolding of ubiquitin. Specifically, the deuterium incorporation of different charge states of ubiquitin produced by the ESI process is examined. Deuterium incorporation was examined, both with and without chromatographic separation to eliminate the influence of excess deuterium oxide together with examining the fast exchanging hydrogens at amino acid side chains and the N-termini prior to ESI MS analysis. In addition, top-down ECD fragmentation enables localization of deuteriums along the amide backbone for charge states +13, +12, +11 and +10. This approach provides not only insight into individual charge states of ubiquitin but also a better understanding of the HDX process occurring in ESI droplets.
4.2 Experimental Section

4.2.1 Materials.

Bovine ubiquitin, acetonitrile (HPLC grade), formic acid (95%), deuterium oxide (99%) and ammonium acetate (anhydrous) were purchased from Sigma-Aldrich Co. (St. Louis, MO).

4.2.2. Deuterium exchange.

HDX samples for nESI were prepared by dissolving 1 mg bovine ubiquitin into 10 mL of ammonium acetate in deuterium oxide (D$_2$O) at 22 °C for 60 minutes. Immediately a set of 40 samples were prepared by transferring 250 µL aliquots of the solution into a vial and diluting to 1.5 mL for each sample. The samples were frozen with liquid nitrogen and maintained at -80 °C. The samples were thawed and analyzed at 0 °C by direct infusion using a flow rate of 0.5 µL/min into the nESI source of the mass spectrometer (LTQ FT Ultra MS, Thermo, San Jose, CA.).

HDX samples for experiments employing conventional ESI were prepared by dissolving 1 mg bovine ubiquitin into 10 mL of 50 mM ammonium acetate in D$_2$O at 22 °C for 1.0 hr. The solution was quenched by adding 40 µL of formic acid. Immediately a set of 40 samples were prepared by transferring 250 µL aliquots of the quenched solution into a 1.5 mL tube for each sample. The samples were frozen with liquid nitrogen and maintained at -80 °C. The samples were thawed, mixed with 70% acetonitrile that was acidified with 0.4% formic acid and then analyzed at 0 °C by direct infusion into the ESI source of the mass spectrometer.
An identical set of samples were subjected to HPLC, utilizing a C4 protein trap prior to ionization. Following injection, ubiquitin was concentrated in the protein trap, washed with 3 mL of H2O at pH 2.5 and 0 °C, and eluted under isocratic conditions (solvent B, acetonitrile, 70 %) at a flow rate of 150 μL/min and the flow rate was split (~1:10) prior to ESI. Both solvents included 0.4% (v/v) formic acid. All components of the HPLC system, from the injector to the peptide trap, were submerged in ice.

4.2.3. Mass Spectrometry.

The capillary tip used for nESI was 20 μm in diamter. The voltages applied to the capillary were varied from 1.8 to 2.5 kV and the ion transfer capillary temperature was maintained at 150 °C. For experiments employing conventional ESI, the capillary voltage was operated at 3.5 kV and the ion transfer capillary temperature was maintained at 150 °C. ECD experiments were performed with precursor selection for each charge state within the ICR cell. The heater current through the dispenser cathode filament was 1.0 A. The resulting c- and z-type ion fragments were identified by their exact masses with the average mass of each individual charge state used to determine their deuterium content.

4.3 RESULTS AND DISCUSSION

4.3.1 Unfolding of ubiquitin in nanoelectrospray ionization (nESI)

Bovine ubiquitin is relatively small, globular protein (8565 Da) protein containing 76 amino acid residues. The native structure of this protein includes five β-strands, a three-turn α-helix, a single turn of 310 helix, six short loops, a highly flexible C-terminal region from
residues 71-76, and 144 exchangeable hydrogen atoms. Seventy-two of the exchangeable hydrogens are located along the amide backbone and 72 sites available for rapid deuterium exchange are located on the side chains and termini [53]. Although previous HDX studies by both NMR and MS show the utility of the technique for the study of protein dynamics in the solution phase [54-56], information for each individual charged ion is still not available. Here, we demonstrated that ubiquitin sprayed from native solution conditions was unfolded in nESI at various voltages applied to the tip of the nESI. The nanoelectrospray mass spectra of the ubiquitin obtained under native solution conditions, containing 50 mM ammonium acetate, incubated for 60 minutes and directly infused into the nESI MS, are presented in Figure 4.1. The nESI mass spectra for ubiquitin in 50 mM ammonium acetate show the presence of the +5, +6, +7 and +8 charge states which were presented at voltages of 1.8 and 2.3 kV. These protein ions incorporated relatively the same numbers of deuteriums. The mass spectrum obtained by using a needle voltage of 2.5 kV, however, shows the +5, +6, +7, +8, +9, +10, +11, +12 and +13 charge states, suggesting that the protein unfolded during the ESI process, resulting in the production of more highly charged ions. The more highly charged ions at this voltage incorporated different numbers of deuteriums compared to those of the low charge ions, shown in Table 4.1.
Table 4.1. Deconvolution of charge states to average molecular mass of native bovine ubiquitin at various voltages in nESI

<table>
<thead>
<tr>
<th>Applied Voltages (kV)</th>
<th>+13</th>
<th>+12</th>
<th>+11</th>
<th>+10</th>
<th>+9</th>
<th>+8</th>
<th>+7</th>
<th>+6</th>
<th>+5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8656</td>
<td>8656</td>
<td>8655</td>
<td>8656</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8660</td>
<td>8659</td>
<td>8659</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8661</td>
<td>8660</td>
<td>8660</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>8676</td>
<td>8668</td>
<td>8660</td>
<td>8656</td>
<td>8655</td>
<td>8659</td>
<td>8658</td>
<td>8659</td>
<td>8659</td>
</tr>
</tbody>
</table>
Figure 4.1. NanoESI mass spectra of native ubiquitin in 50 mM ammonium acetate at various voltages ranging from 1.8 to 2.5 kV.
4.3.2 Analysis of amide HDX for individual charge states of ubiquitin in the conventional ESI

The electrospray mass spectrum obtained from ubiquitin solutions containing 50 mM ammonium acetate, incubated for 60 minutes and quenched with 0.4% formic acid, by direct infusion into the MS, was used as a control experiment, presented in Figure 4.2. This figure shows the same molecular mass of 8565 Da for all charge states, based on the deconvolution of m/z from each charge state; \( M_w = [(m/z) \times z] - [z \times \text{mass of H}] \), where \( z \) is the number of charge states, presented in Table 4.2 [1].

Figure 4.2. ESI mass spectrum of ubiquitin in 50 mM ammonium acetate and 0.4% formic acid.
Table 4.2. Deconvolution of charge states to average molecular mass of intact bovine ubiquitin

<table>
<thead>
<tr>
<th>Charge State of Ubiquitin</th>
<th>Average Molecular Mass ((M_w))</th>
<th>Calculated Average Molecular Mass ((M_w))</th>
</tr>
</thead>
<tbody>
<tr>
<td>+13</td>
<td>8564.6 ± 0.1</td>
<td>8564.6</td>
</tr>
<tr>
<td>+12</td>
<td>8564.7 ± 0.1</td>
<td>8564.6</td>
</tr>
<tr>
<td>+11</td>
<td>8564.6 ± 0.0</td>
<td>8564.6</td>
</tr>
<tr>
<td>+10</td>
<td>8564.7 ± 0.1</td>
<td>8564.6</td>
</tr>
<tr>
<td>+9</td>
<td>8564.6 ± 0.1</td>
<td>8564.6</td>
</tr>
<tr>
<td>+8</td>
<td>8564.6 ± 0.1</td>
<td>8564.6</td>
</tr>
<tr>
<td>+7</td>
<td>8564.6 ± 0.1</td>
<td>8564.6</td>
</tr>
</tbody>
</table>

The deconvolution results from Table 4.2 demonstrate that average molecular masses of the multiple charge states at high resolution are the same. These results alone, however, do not differentiate nor provide structural information for the individual charge states. It is the combination of HDX in solution, followed with ECD MS that provides the structural information for the individual charge states. The introduction of the samples into the ESI MS was carried out by two different methods; direct infusion and chromatographic separation. Direct infusion of the native ubiquitin, subjected to continuous labeling with deuterium oxide in 50 mM ammonium acetate, incubated for 60 minutes and quenched with 0.4% formic acid, yields mass spectra that reveal differences in the deuterium incorporation of the various charges states of this protein as shown in Figure 4.3. The average molecular masses of deuterated bovine ubiquitin are shown in Table 2. Deuterium incorporation along the amide
backbone for each individual charge state can be obtained from the deconvolution of m/z from each charge state; $M_w = [(m/z) \times z] – [z \times (mass\ of\ D)]$, and subtraction from the molecular mass of ubiquitin, 8565 and 72 fast exchanged dertiurns at the side chains and termini. Table 4.3 shows the differences in deuterium incorporation among the charge states and that charge state +13 shows more deuterium incorporation than the lower charge states. These variations may arise from two possibilities - deuterium exchange in the solution or deuterium as the protein continues to unfold in the spray droplet during ESI.

![Figure 4.3. ESI mass spectrum from direct infusion of bovine ubiquitin in deuterium oxide with 50 mM ammonium acetate and 0.4% formic acid.]

Figure 4.3. ESI mass spectrum from direct infusion of bovine ubiquitin in deuterium oxide with 50 mM ammonium acetate and 0.4% formic acid.
Table 4.3. Deconvolution of charge states to average molecular masses of deuterated bovine ubiquitin, including fast exchanged deuteriums at the side chains and termini.

<table>
<thead>
<tr>
<th>Charge State of Ubiquitin</th>
<th>Average Molecular Mass (Mₚ)</th>
<th>Observed Number of Incorporated Amide Deuteriums</th>
</tr>
</thead>
<tbody>
<tr>
<td>+13</td>
<td>8682.4 ± 0.1</td>
<td>47</td>
</tr>
<tr>
<td>+12</td>
<td>8675.4 ± 0.0</td>
<td>40</td>
</tr>
<tr>
<td>+11</td>
<td>8670.3 ± 0.6</td>
<td>35</td>
</tr>
<tr>
<td>+10</td>
<td>8666.3 ± 0.9</td>
<td>31</td>
</tr>
<tr>
<td>+9</td>
<td>8665.6 ± 0.4</td>
<td>30</td>
</tr>
<tr>
<td>+8</td>
<td>8664.4 ± 0.3</td>
<td>29</td>
</tr>
<tr>
<td>+7</td>
<td>8665.7 ± 0.5</td>
<td>30</td>
</tr>
</tbody>
</table>

The results from Table 4.3 show that the lower charges have less deuteriums. These results agree with the studies that demonstrated that the lower charge states are compact allowing for the preservation of structure that can survive in the gas phase [4]. In the case of ubiquitin, the lower charge states, such as + 7 to +10, show a similar, low number of deuteriums, suggesting that these charge states are compact and more protected. On the other hand, the higher charged ions, like +13 to +11, are a result of the unfolding of lower charge states and exchange with available deuterium oxide in the ESI droplets. Therefore, the higher charge states reflect a higher degree of unfolding and incorporate more deuteriums during ESI.

The same set of continuously labeled samples was used with HDX experiments that incorporated chromatographic separation. A protein trap was employed to wash out excess
deuterium oxide and the deuteriums at the side chains and termini. For the results, shown in Figure 4.4, deuterium incorporation along the amide backbone for each individual charge state can be obtained by subtraction with the molecular mass of ubiquitin, 8565. The data in Table 4.4 show that the highest charge state does not gain more deuterium than the lower charge states. This result presumably arises because there was no deuterium oxide in the droplets that would allow deuterium incorporation into the protein during the unfolding process, and the exchanged rates of those deuteriums at the amide backbones are slow. Therefore, there is only a small possibility that further deuterium exchange reaction occurs in the droplet during electrospray.

Figure 4.4. ESI mass spectrum, of bovine ubiquitin in deuterium oxide with 50 mM ammonium acetate and 0.4% formic acid, introduced by chromatic separation, equipped with a C4 protein trap into the ESI source.
Table 4.4. Deconvolution of charge states to average molecular mass of deuterated bovine ubiquitin with chromatographic separation.

<table>
<thead>
<tr>
<th>Charge State of Ubiquitin</th>
<th>Average Molecular Mass ((M_w))</th>
<th>Number of Incorporated Amide Deuteriums</th>
</tr>
</thead>
<tbody>
<tr>
<td>+13</td>
<td>8591.5 ± 0.5</td>
<td>26</td>
</tr>
<tr>
<td>+12</td>
<td>8591.4 ± 0.4</td>
<td>26</td>
</tr>
<tr>
<td>+11</td>
<td>8591.4 ± 0.4</td>
<td>26</td>
</tr>
<tr>
<td>+10</td>
<td>8591.3 ± 0.2</td>
<td>26</td>
</tr>
<tr>
<td>+9</td>
<td>8591.5 ± 0.5</td>
<td>26</td>
</tr>
<tr>
<td>+8</td>
<td>8591.5 ± 0.3</td>
<td>26</td>
</tr>
<tr>
<td>+7</td>
<td>8591.2 ± 0.3</td>
<td>26</td>
</tr>
</tbody>
</table>

4.3.4. MS/MS of HDX Ubiquitin by Top-down ECD FTICR MS

4.3.4.1 HDX and Top-down ECD FTICR MS from direct infusion

The same method of continuous labeling of samples was used with HDX measurement employing direct infusion was further studied by top-down ECD FTICR MS. The results from this can provide important information for HDX MS experiments because many HDX experiments that were carried out by top-down ECD or ETD simply used the higher or most abundant charge states as precursor ions [35,44,56-58]. The reason for this is that the higher charge states yield more fragmentation and a higher percent recovery in HDX ECD MS. In this study, the analysis of amide HDX of charge states +13, +12, +11 and +10 of ubiquitin by MS/MS employing ECD enabled the identification of the precise locations of amide hydrogens for those individual charge states. MS measurements of the ECD
fragments from c- and z-type· ions provide the number of deuteriums incorporated. The percent deuterium incorporation, D, of each fragment was determined by the equation \( D = \frac{(R - R_0)}{N} \), where \( R \) is the centroid mass of the isotope distribution, \( R_0 \) is the centroid of the unlabeled control, and \( N \) is the number of exchangeable hydrogens [59-60]. The results show that the +13, +12, +11 and +10 charge states yield fragment ions from c3 to z-31 and z-3 to z-23 that had high S/N ratios and are reliable for HDX determination. Comparison of the deuterium content for the same c and z· fragment ions with differing charge states revealed a significant change in the deuterium content between charge state +13 and charge states +11 and 10 for the fragments c24 to c31. For fragment c24, the number of deuteriums incorporated for charge states +13, +11 and +10 are 16, 13 and 13, respectively. The fragment c31 shows that the number of deuteriums incorporated at the +13, +11 and +10 charge states are 20, 15 and 14, respectively. In this finding, charge state +13 exhibited a significant change in the number of deuteriums incorporated, 5, while the charge states +11 and +10 showed little change. Mass spectra of fragments c24 and c31 of charge states +13, +11 and +10 are shown in Figure 4.5 and the differences in deuterium incorporation, found in the long helices between E24 and Q31 in bovine ubiquitin are shown in Figure 4.6.
Figure 4.5. Mass spectra of selected ubiquitin ECD fragments of c24\(^{3+}\) and c31\(^{4+}\) after 60 min of D\(_2\)O labeling for charge +13, +12, +11 and +10 ions. Note the different mass shifts exhibit 5 exchangable hydrogens for the +13 charge state ion.
The results for the +13 charged ion of ubiquitin indicate that the hydrogen bonds of the 24-31 $\alpha$-helix were disrupted and rapidly incorporated available deuterium during ESI, gaining additional deuteriums. The unfolding of ubiquitin ions during the transition from the solution phase to the gas phase allows the possibility of additional deuterium incorporation.
during ESI. Upon unfolding, the deuterium oxide in the droplets can engage in deuterium exchange with labile hydrogens that were exposed from the protein core. These results are in good agreement with the previous study of ubiquitin refolding in solution phase by HDX NMR, in which the 23-34 α-helix was found to fold back into the native form on the millisecond time scale [61]. The evidence from gas phase studies by HDX and ECD, show that the +13 charge state ions of ubiquitin have less tertiary noncovalent bonding and the ECD bond cleavages are substantially different from those of other charge states [62-63]. In addition, the unfolding of a protein during the ESI process was proposed to arise from coulombic repulsion between the charged residues and liquid/droplet surface charge [64].

4.3.4.2 HDX and Top-down ECD FTICR MS by chromatic separation

When chromatographic separation was employed to remove the deuterium oxide and the fast exchanged deuterium at the side chains and termini, the deuterium incorporations on the amide backbone were constant for all of the charge states observed in the experiment. Each charge state was further selected and fragmented by ECD to measure the deuterium incorporation on the amide backbone, especially at the 24-31 α-helix residue. The results show that there are the same number of deuteriums incorporated on the amide backbone of the 24-31 α-helix residue for all charge states. Mass spectra of fragments c24 and c31 of charge states +13, +11 and +10 from this approach are shown in Figure 4.7.
Figure 4.7. Mass spectra of selected ubiquitin ECD fragments of c24$^{3+}$ and c31$^{4+}$ after 60 min of D$_2$O labeling, followed by eliminating deuterium oxide and the fast exchanged deuteriums at the side chains and termini prior to the ESI, for charge +13, +12, +11 and +10 ions.
4.3.5 Combining direct infusion and chromatographic separation HDX experiments

Considering the advantages and drawbacks from both experiments, the fast exchanged deuterium is preserved while the additional deuterium may associate in the direct infusion experiment. In the chromatographic separation experiment, on the opposite, the fast exchanged deuterium may wash out throughout the chromatographic separation but the addition of deuterium during the electrospray process is eliminated. Therefore, the results from both experiments provide information regarding the labile exchangeable sites, as shown in Figure 4.8.

Figure 4.8. Structure of ubiquitin from Protein data bank and UCSF Chimera shown locations of the labile amide hydrogens (the difference in deuteriums between the +13 charge state from direct infusion and the +13 charge state from HPLC/C4).
4.4 Conclusions

The results from the unfolding of ubiquitin under native condition by nESI MS indicate that the protein undergoes unfolding during the ESI process and the degree of the unfolding is dependent on many parameters. For example, the voltage of the ESI needle influences the degree of unfolding. Unfolding of the protein during ESI appears to be associated with the deuterium exchange reaction. This affects the deuteriums on the higher changed ions more than the lower ones. The results from the conventional HDX and the ESI MS analysis of ubiquitin under quenched condition show the CSD distribution of ubiquitin from +7 to +13. The higher charge states, especially the highest charge state, incorporated more deuteriums than the lower ones in the experiment that included the exchange sites on the side chains and the termini. The results of this ESI-MS analysis were compared to that of an ESI-MS analysis that used chromatographic separation, equipped with C4 protein trap that eliminated excess deuterium oxide and the fast exchanged deuteriums at the side chains and termini prior to the electrospray source. The individual charge states of ubiquitin from the latter experiment exhibited the same numbers of deuteriums, especially at the 24-31 helix location. The results from both experiments suggest that the higher charge state may be produced from the unfolding of the lower compacted protein ions and incorporating with the available deuterium in the droplets. The results from amide HDX and Top-Down ECD FTICR MS of ubiquitin show that the charged +13 ion incorporated more deuterium than those of +12, +11 and +10 ions at the helix position, suggesting that the conformation of each charge state in solution can be slightly different from others, depending on the degree of unfolding in ESI. Eliminating the deuterium oxide and fast exchange deuteriums at the side chains and termini approach can be benefit for Top – down HDX ECD of higher charge
states that can achieve single amino acid residue resolution, similar to those of the lower charge ions and represent the native structure of the protein in solution. The combination of both experimental results from the higher charge ions provides information of labile exchangeable hydrogen, and exchangeable amide hydrogen. Although these results supports that the lower charge ions are compact and more likely to preserve a native-like structure from the solution-phase, the Top-down ECD of the lower charge ions does not provide much structural information.
4.5 References


CHAPTER 5. Conformational Structure of Human Insulin in the Presence of Congo Red by Pulsed Labeling Hydrogen-Deuterium Exchange Mass Spectrometry

5.1 Introduction

Congo red (CR) is a small linear molecule, consisting of hydrophilic part from two amino groups and sulfate groups — and hydrophobic part from a biphenyl group shown in Figure 5.1[1]. It is water soluble and has propensity to aggregate in aqueous and organic solutions, resulting from hydrophobic interactions between the aromatic rings in the molecule. Early studies, conducted at various concentrations, pH and temperatures found that CR aggregation was depending on its concentration, temperature, ionic strength and pH of the solution [2-4]. At high concentration, CR forms self-aggregation in form of ribbon-like micelles and has been reported to inhibit amyloid formation of proteins [5,6]. On the contrary, CR at very low concentration promotes fibril formations in some proteins [7]. Increasing temperature, pH and decreasing the ionic strength of the solution decrease the extent of micelle formation. CR is often used for staining amyloid fibrils because it binds specifically to $\beta$-sheet structures in amyloids resulting in yellow-green birefringence under crossed polarization and has been used as the analytical “gold standard” for amyloid characterization and diagnostic since the beginning of the 1920s [8,9]. CR was also reported as interference in processes of protein misfolding and aggregation and as an inhibitor of the neurotoxicity of amyloid of Alzheimer’s diseases by binding to the fibrils. Despite its unique binding to amyloid, CR was used as a molecular amyloid probe in vivo and has recently been reported to reduce a$\beta$ neurotoxicity [10,11]. It is however concerned for having high toxicity and poor pharmacokinetics. The binding mechanism between CR and amyloid is believed to
arise from the interaction of both hydrophobic and hydrophilic parts of the CR molecule. The negatively charged sulfate groups of Congo red were proposed to interact with positive charged position in amyloid [12]. Hydrogen bonds and ionic interactions have also been studied and assumed to have contributions to the binding of CR to amyloid fibrils [9]. However, most studies have focused on hydrophobic interactions, distributing from the biphenyl group of CR and the β-sheets of amyloid. This is because amyloids are β-sheets rich proteinaceous aggregates that are linked to a numerous human neurological diseases [13]. Binding interactions of CR and amyloids are highly specific and β-sheet structure of amyloid is the important factor in CR binding [14].

Insulin is one of most widely study hormone proteins, consisting of 51 amino acid residues that adopts a primarily helical structure in the native state, in which residues A1-A8 and A13-20 form two helices in the chain, while residues B11-B19 form a helix in the B
chain. It is believed that the fibrilogenesis of insulin is initiated by the partial unfolding of the monomers, followed by further association forms of oligomers and aggregation, containing β-sheet rich element [15,16]. Insulin aggregation involves specific processes, displays a wide variety of morphologies and has long been known to form amyloid fibrils under specific conditions [16-19]. An early study by X-ray crystallography, focused on the interaction between Congo red and insulin, proposed that CR binds selectively to anti-parallel monomers forming dimer between the B-chains of pig insulin molecules [15]. Later, some studies found that the insulin fibrillation was resulted from the formation of the despentapeptide and deoctapeptide in insulin monomers rather than dimers [20-22] and the formation of the fibril involves a nucleation step in which three or four insulin monomers simultaneously form through hydrophobic interaction [25,26]. Therefore, Congo red when bound the insulin monomer can prevent the nucleation process.

Pulsed labeling hydrogen-deuterium exchange coupled with electrospray mass spectrometry (HDX MS) has been employed to study the secondary structures of partial folded intermediate that appear during folding reactions as well as of stable equilibrium intermediates [30-40]. This approach has been very useful when coupled to HPLC, equipped with reversed phase trap and column at minimum exchange conditions, producing peptic digested peptides that can be separated and assigned to specific regions of the peptide backbone [41]. The method focuses on the protection of backbone amide hydrogen against exchange with deuterium exchange reaction when the protein is incubated in the deuterium oxide solution. Typically, the exchange rate of amide hydrogen is rapid for those amide
groups located in the unfolded or unstructured regions. On the contrary, the protected amide hydrogen is exchange very slowly, especially for those that participate in hydrogen bonding or located in the protein interior. In addition, Congo red is eliminated before entering to the ESI source.

There is currently no hydrogen-deuterium exchange experimental structure of CR bound to an amyloid fibril available. In addition, the binding modes, and the mechanisms governing its inhibitory and optical properties are not fully understood. In this study, we demonstrate the use of pulsed labeling hydrogen-deuterium exchange mass spectrometry to study the binding interaction between human insulin and Congo red. Because Congo red is known for binding to critical intermediate structural forms of amyloid protein, it has been suggested as a potential therapeutic agent against neurodegenerative disorders, such as Alzheimer’s disease. The goal of this study is to apply this method to map the partial folded intermediate in human insulin in the presence of CR. The results may provide further insights into the binding of CR to inhibit amyloid toxicity in neurodegenerative diseases.

5.2 Experimental Section

Human insulin solution (5 mg/mL), porcine pepsin, acetonitrile (HPLC grade), formic acid (95%), and deuterium oxide (99%) were purchased from Sigma, St. Louis, MO. Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) were purchased from Fisher Scientific, Pittsburgh, PA. Congo red was purchased from Sigma.
5.2.1 Sample preparations

Pulsed labeling hydrogen-deuterium exchange of human insulin was initiated by adding 5 µL of CR solution (2 mg/mL) into a solution of 5 µL of human insulin (5 mg/mL) at pH 2.0 at 22 °C, incubated for 30 minutes at 22 °C. Immediately, the aliquot was submerged in ice bath for 5 minutes to bring the temperature down to 0 °C, followed by adding 500 µL of deuterium oxide buffered with 10 mM ammonium acetate at 0 °C into the solution, incubated for 10 seconds, and immediately quenched by adding 5 µL of cold formic acid to adjust pH down to 2.5. The HDX sample was subjected for global ESI MS analysis. For local analysis, after adjusting the pH down to 2.5, 10 µL of TCEP solution at pH 2.5 and 0°C and 1 µL of pepsin solution (1.00 mg/mL at pH 2.5 and 0°C) were added into the solution. The sample then subjected to separations and ESI MS analysis. The schematic procedure of pulsed labeling HDX HPLC ESI MS is shown in Figure 5.2.
Figure 5.2. The experimental procedure for pulsed labeling HDX for insulin binding Congo red.

5.2.2 HDX analysis by HPLC ESI MS

The deuterium incorporations at the amide backbone of the human insulin for both global and local ESI MS analyses were determined by using HPLC equipped with C18 reversed phase trap at the injector and C18 reversed phase column prior to ESI MS analysis. The HPLC system consisted of two pumps (Hitachi, L-7100), an injector (Rheodyne 8125, Altech, Deerfield, IL) equipped with a protein trap (Michrom BioResources, Inc., CA) and a C-18 reversed phase column (Discovery Bio Wide Pore, C18, 3 μm particle size, 2.1 mm x 50 mm). The instrument setup for this experiment is shown in Figure 5.3. Following injection, the HDX sample was concentrated in the peptide trap, washed with 500 μL cold quenched solution (H₂O, pH 2.5, 0 °C) to eliminate fast exchanged deuteriums and Congo
red. The sample then subjected to ten minute gradient elution (solvent B, acetonitrile, 5-85%) at a flow rate of 150 μL/min. Spit flow was employed before the solution reached the ESI source. Both solvents included 0.3% (v/v) formic acid. All components of the HPLC system from the injector and column were submerged in ice. The ESI MS data presented here were acquired using an ion trap mass spectrometer (LTQ FT, Thermo, San Jose, CA). Peptides of bovine insulin were identified by their exact masses. The centroid mass of each peptic fragment was determined using the Xcalibur program (Thermo, San Jose, CA). The deuterium content of the peptides was calculated using the difference between the centroid masses of deuterated and undeuterated protein and peptides.

Figure 5.3. The experimental procedure for pulsed labeling HDX for insulin binding Congo red.
To obtain an unambiguous assignment, each undeuterated peptic fragment was sequenced using a collision-induced dissociation fragmentation method. Individual peptic digestion fragments were selected and subjected to low-energy electron fragmentation in the ion trap mass spectrometer. The fragment ions were subsequently detected by Fourier transform ion cyclotron mass analyzer. The monoisotopic peaks were analyzed and assigned by Protein prospector, San Francisco, CA.

5.3 Results and Discussion

The results presented here combine the use of peptic digestion and pulsed labeling hydrogen-deuterium exchange chromatographic separation and mass spectrometry to investigate the binding sites between human insulin and Congo red. A general procedure commonly used to study the hydrogen exchange properties of proteins is illustrated in Figure 5.2. The solution of protein samples were diluted 50 fold with deuterated solution to initiate isotopic exchange. After 10 second incubation, the labeled solution was quenched to minimize the exchange reaction. The deuterium level in labeled proteins was determined by HPLC and ESI MS. The instrumental setup for the separations is shown in the figure 5.3. The purpose is to eliminate fast exchanged deuteriums at the side chains and termini, remaining the deuteriums incorporation at the amide backbone. This method allows the analysis to focus on the amide backbone and to reduce the complexity of the data interpretations. More importantly, Congo red and water soluble salts are eliminated during the washing step. Thus, it does not interfere with the mass analysis or contaminate the mass spectrometer.
5.3.1 Pulsed labeling HDX for global analysis

Typically, most probable structures of proteins exist in the native fold forms at ground state. The protein can, however, undergo dynamic excursions from its ground state into entirely different structural forms, referred as excited states [40,41]. Pulsed labeling HDX of the intact insulin and insulin in the presence of CR provide global view of the hydrogen exchange properties. Binding interactions of CR and amyloids are highly specific. In fact, β-sheet conformation of amyloid is the important factor in CR binding. ESI mass spectra of human insulin monomer, pulsed labeling hydrogen-deuterium exchange of human insulin monomer and human insulin monomer in the presence of CR are shown in Figure 5.4.
Figure 5.4. Representative ESI mass spectra of +5 charge state of human insulin, pulsed labeling hydrogen-deuterium exchange of human insulin and pulsed labeling hydrogen-deuterium exchange of human insulin bound with Congo red.

Human insulin consists of two poly peptide chains (A and B), constrained by one intramolecular and two intermolecular disulfide bonds and contains of 49 amide hydrogens. Charge state +5 is chosen in this study to determine the deuterium incorporation at the amide backbone for global analysis. The ten second pulsed labeling with deuterium oxide of this
monomeric peak in human insulin shows one envelope peak at average m/z 1166.03 or 5825 Da, resulting in 17 deuteriums incorporated at the amide backbone for monomeric insulin. The result from the pulsed labeling of the insulin in the presence of CR shows a bimodal pattern, indicating that pulsed labeling is very sensitive to the partial fold structure of the insulin bound with Congo red. The deuterium incorporations at the amide backbone for the first envelope, which is close to that of the monomeric insulin, and the second envelope are 18 and 31, respectively. It is clear that the binding increases the amount of incorporated deuteriums, suggesting that the binding between the insulin and Congo red, opens the structure of insulin for deuterium to exchange with amide hydrogens at that position. It is possible that CR binds to the partial fold intermediate of the insulin and this fold intermediate structure may hold the key to unlock how the protein folds and functions. To identify the binding site between Congo red and insulin, the insulin solution was subjected to enzymatic digestion followed by the same separations and mass spectrometric analysis.

5.3.2 Pulsed labeling HDX for local analysis

The hydrogen exchange properties of the insulin bound with CR in specific regions of polypeptide backbone was determined by similar analysis to the labeled protein. The results presented here combine the use of hydrogen exchange and peptic digestion in the presence of TCEP to identify the regions that are partially folded and bound with CR. The labeled sample was digested with pepsin solution in the presence of TCEP at pH 2.5 and 0°C. The peptic digested samples were immediately subjected to HPLC equipped with a peptide trap and a reversed phase column followed by ESI MS analysis. The chromatographic separation in this
experiment is mainly to eliminate the fast exchanged deuterium incorporated in the insulin, deuterium oxide and more importantly Congo red prior to MS analysis. The peptic digestion products were eluted in one minute interval. Therefore, the effect from the back exchange should be similar for all peptic fragments. The ESI mass spectrum resulted from pulsed labeling hydrogen deuterium exchange, followed by the pepsin digestion is shown in Figure 5.5. From the pattern of peptic digestion products it is evident that undergo proteolysis more readily at C-terminal of the B-chain. In addition, B-chain is basic and more likely to be protonated and more abundant in the mass spectrum.
Figure 5.5. ESI mass spectrum of peptic fragments from pulsed labeling human insulin.

The mass spectrum of the peptic digestion of human insulin in the presence of Congo red shows the peptic fragments, containing molecular mass of insulin. Because mass spectra of the insulin bound with CR, resulted in native and partial folded forms, it is expected that mass spectra of the peptic peptides that bound with CR, exhibit the pattern similar to that of the insulin in the presence of CR, representing the two populations. These experiments remain the analyses of the peptic digestions of intact insulin, labeled insulin and insulin bound with CR in the same mass spectrum. The digestion of the intact insulin is served as the
reference sample, 0 % and this sample was not exposed to the deuterium oxide solution. The digestion of the labeled insulin provides the deuterium levels of peptic peptides that are used in comparison with those of the insulin bound with CR. The deuterium incorporations of the peptic fragments analysis are shown in Table 5.1.
Table 5.1. Deuterium levels in peptic fragments of insulin and insulin in the presence of CR

<table>
<thead>
<tr>
<th>Peptic digest residue</th>
<th>Assignment</th>
<th>% D incorporation in Insulin</th>
<th>% D incorporation in Insulin + Congo red</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2-11 B2-10</td>
<td>IVEQCCTSIC VNQHLCGSH</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>A4-8</td>
<td>IVEQ</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>A11-13</td>
<td>CSL</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>A12-14</td>
<td>SLY</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>A14-21</td>
<td>YQLENYCN</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>A19-21</td>
<td>YCN</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>B1-5</td>
<td>FVNQH</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>B1-11</td>
<td>FVNQHLCGSHL</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>B8-14</td>
<td>GSHLVEA</td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td>B12-15</td>
<td>VEAL</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>B14-22</td>
<td>ALYLVCGER</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>B17-25</td>
<td>LVCGERGFF</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>B25-30</td>
<td>FFYTPKT</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>B26-30</td>
<td>FYTPKT</td>
<td>70</td>
<td>60</td>
</tr>
</tbody>
</table>
The result from the figure 5.4 shows the mass spectrum containing the molecular mass of the labeled insulin that exhibits folded and partial unfolded envelopes of charge state +5 at m/z 1166.38 and 1169.05, corresponding to 19 and 32 deuteriums, respectively. The peptic fragment, resulted from removal of C-terminal residues from the B-Chain to form despentapeptide-(B26-30) also show a similar pattern, containing fold and partial unfolded envelopes of the same charge state at m/z 1047.31 and 1049.27, corresponding to 15 and 24 deuteriums, respectively, illustrated in Figure 5.6. Thus the difference of deuterium between the labeled molecule and the fragment is 4 deuteriums, suggesting that the amide hydrogen of five amino acids at the C-terminus of B-chain are exchanged with deuterium and involved in the binding.

Figure 5.6. Mass spectra of the molecule and despentapeptide-(B26-30) of human insulin under identical conditions show the number of deuterium incorporation for both isotropic envelopes.
The peptic fragments B25-30 and B26-30 also show similar behavior. However, this region is known for rapid exchange with deuterium, compared to the other parts of the molecule [42]. The result from binding between insulin and Congo red has partial effect for protection from the deuterium exchange reaction in this region. This is because the hydrophobic part of Congo red prevents deuterium getting contact to the surface. After the removal of B25-30, the rest of the labeled fragment still show the bimodal like character, suggesting that other parts of the structure were also affected by Congo red. The results from Table 5.1 show level of deuterium incorporations of peptic digestion fragments. The deuterium contents in fragments A2-11- B2-10, A4-8, 11-13, A12-14, A19-21, and B1-11 are the same for both labeled insulin and labeled insulin in the presence of CR, suggesting that these parts of the molecule do not change after the binding occurred. The results for the labeled peptic digestion analysis show that the CR binding contributes to the difference in hydrogen exchange protection, observed in fragments A14-21, B8-14, B12-15, B14-22, B17-25, B25-30 and B26-30. Thus, amides involved in the changes are located at A14-19, which is the second helix element in the A-chain and B13-29 in the B-chain. Remarkably, the most significant change is at residues B18-24 in the B-chain. The summary of the hydrogen exchange properties for amide hydrogens of insulin in the presence of CR is illustrated in Figure 5.7.
Figure 5.7. Amino acid sequence of human insulin in the presence of CR displaying hydrogen exchange properties at amide backbone. Pink and Red color indicate region where significant and the most significant exchange occurred, respectively.

Based upon the fragmentation analysis, the regions where the most significant exchange took place resulted from the greatest change in structural details of the insulin that transformed from significantly protected to less protected from the deuterated solvent and allowed the deuterium exchange reaction taking place. Thus, the amide backbone of amino acids A14-19, B13-29 and B18-24 may play an important part to the partial unfolded intermediate structure of the insulin in the presence of CR and may initiate the conversion to β-sheet structure. The conformational structures, representing the deuterium exchange levels of peptic fragments for human insulin and human insulin in the presence of Congo red are shown in Figure 5.8.
Figure 5.8. Conformational structures of labeled insulin and labeled insulin in the presence of Congo red, using Chimera, UCSF, CA.
5.4 Conclusion

Human insulin can sample conformational states, also called partial unfolded intermediates, which are critical for function but are seldom detected directly because of their low occupancies and short lifetimes. The results from this experiment show that human insulin in the presence of CR possesses a partial folded intermediate that involves amide backbone at amino acid positions A14-19 in the A-chain and B13-29 in the B-chain. The difference of the exchange rate between human insulin and human insulin the presence of CR becomes large at the amide backbone residues B18-24, suggesting that these amino acid residues involve in the partial folded intermediate structure that may provide to insight into how the protein aggregates.
5.5 References


