9-21-2017 2:00 PM

Investigations into the Role of Conformational Dynamics in Protein Function: Insights From Hydrogen/Deuterium Exchange-Mass Spectrometry

Courtney S. Fast
The University of Western Ontario

Supervisor
Konermann, Lars
The University of Western Ontario

Graduate Program in Chemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

Abstract

Deciphering protein structure and dynamics is a key prerequisite for understanding biological function. The current work aims to apply HDX-MS to improve the understanding of protein structure and dynamics for systems that remain challenging for other techniques. Following a general overview of the field (Chapter 1), Chapter 2 investigates the relationship between enzyme dynamics and catalysis. By conducting comparative HDX-MS measurements on *r*M1-PK during substrate turnover and in the resting state, catalytically active *r*M1-PK undergoes significant rigidification of the active site. However, virtually the same rigidification was seen upon exposing *r*M1-PK to substrate or product in the absence of turnover. These findings demonstrate that comparative experiments on enzyme dynamics by HDX-MS (and other bioanalytical techniques) should be interpreted with caution. In Chapter 3, HDX-MS is used to probe the intrinsically disordered protein Nrf2. HDX-MS is used to investigate the structure and dynamics of the full-length Nrf2 and its interaction with the Kelch domain. The data obtained demonstrate the highly-disordered nature of Nrf2. Its interaction with Kelch causes protection of the binding sites on Nrf2, while the rest of the protein becomes slightly more dynamic. This works highlights the limitations of using truncated protein constructs when investigating their structure and dynamic properties using biophysical techniques.

Keywords: protein structure | conformational dynamics | electrospray ionization | hydrogen/deuterium exchange | mass spectrometry | rabbit muscle pyruvate kinase | catalysis | intrinsically disordered proteins | Nrf2 | Kelch

Co-Authorship Statement

The work in Chapter 2 was submitted in the following articles:

> Fast, C. S., Vahidi, S., Konermann, L. (2017) Changes in Enzyme Structural Dynamics Studied by Hydrogen Exchange-Mass Spectrometry: Ligand Binding Effects or Catalytically Relevant Motions? *Anal. Chem.* (Submitted)

The original draft of this manuscript was prepared by the author (C. S. F). Subsequent revisions were performed by the author, Dr. Siavash Vahidi, and Dr. Lars Konermann. All experimental work was performed by the author under the supervision of Dr. Lars Konermann.

Acknowledgments

First, I would like express my sincerest appreciation to my supervisor, Dr. Lars Konermann. Lars is the most patient, charismatic, and hard-working person and his support and inspiration is more than I could ask for. He has provided me with many opportunities throughout my time in the Konermann Lab and has been a great mentor whom I owe my future successes in the protein mass spectrometry field to.

I would like to thank the past and present members of the Konermann Lab for their support. To the current members of the lab; Haidy, Ming, Victor, Maryam, Angela, and Katja. Thank you for the many useful discussions and thought provoking ideas that have help me through the challenges of my graduate studies. I would like to thank Dr. Siavash Vahidi for his mentoring in the early stages of my time in the Konermann Lab.

I would like to thank Dr. Stanley Dunn (UWO Biochemistry) for his help in the early stages of Chapter 2. As well, Nadun Karunatilleke and Dr. James Choy (UWO Biochemistry) for their work and insights in Chapter 3.

A big thank you to my committee members and examiners: Dr. Martin Stillman, Dr. Len Luyt, Dr. Brent Sinclair, Dr. Johanna Blacquiere, and Dr. Michael Kerr.

And last but definitely not least, I would like to acknowledge my friends and family. Thank you for your love and support through my M.Sc. thesis over the past two years.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Symbols and Abbreviations

ADP – adenosine diphosphate

ATP – adenosine triphosphate

BTB – Broad complex, Tramtrack, and Bric-a-Brac

bZIP – basic leucine zipper

CCS – collision cross section

CD – circular dichroism

CEM – chain ejection model

CID – collisional induced dissociation

CRM – charged residue model

DC – direct current

ESI – electrospray ionization

FA – formic acid

FBP- fructose-1,6-bisphosphate

FT-ICR – Fourier transform- ion cyclotron resonance

HDX – hydrogen/deuterium exchange

IDP – intrinsically disordered protein

IEM – ion ejection model

IMS – ion mobility spectrometry

IVR – central intervening region

Keap1 – Kelch-like ECH-associated protein 1

LDH – lactate dehydrogenase

LC – liquid chromatography

Maf – musculoaponeurotic fibrosarcoma

MALDI – matrix-assisted laser desorption/ionization

MD – molecular dynamics

MS – mass spectrometry

m/z – mass-to-charge ratio

NMR – nuclear magnetic resonance

Nrf2 – nuclear factor erythroid 2-related factor 2

PEP - phosphoenolpyruvate

PDB – protein data bank

RF – radio frequency

ROS – reactive oxygen species

*r*M1-PK – rabbit skeletal muscle pyruvate kinase

TOF – time-of-flight

UV-Vis- ultraviolet-visible

Chapter 1 – Introduction

1.1 Proteins: Structure, Dynamics and Function

1.1.1 Protein Structure

Proteins are macromolecules that carry out essential physiological functions. Their roles in biological systems include transport, catalysis, signaling, and cellular defense. The biologically active structure of a protein is called their *native* structure. Historically, proteins follow a classic structure-function paradigm. This model implies that the three-dimensional structure of a protein is associated with a specific function.[1] Therefore, visualizing and understanding a protein's three-dimensional native structure is key to understanding the specific function they carry out.

Proteins are made up of amino acids. Amino acids are organic compounds that contain an amine ($-NH_2$), a carboxylic acid (-COOH), and a variable R-group (side-chain), all of which are bound to a central $C_\alpha H$ moiety shown in Figure 1.1a.[2] Amino acids are linked together by an amide bond (-OC-NH-), forming a polypeptide backbone (Figure 1.1b, shown in red). The linear sequence of amino acids constitutes the protein's primary structure. The chemical composition of the polyamide backbone permits hydrogen bond (H-bond) formation between CO and NH groups. These H-bonds give rise to secondary structures such as α-helices and/or β-sheets. The three-dimensional structure of a protein is known as its tertiary structure. A protein's native tertiary structure represents the lowest free energy conformation, as governed by numerous bonding interactions between the side-chain groups and main chain moieties.[3] In addition to H-bonds, these interactions include hydrophobic interactions, disulfide bridges, salt bridges, and/or van der Waals contacts. In

the native tertiary structure, nonpolar side chains form a hydrophobic core whereas the hydrophilic side chains are found on the surface of the protein. Some proteins are made up of multiple polypeptide chains. Depending on their subunit composition, these complexes can be homo- or hetero-multimers. The quaternary structure refers to how the individual subunits are arranged within the complex.

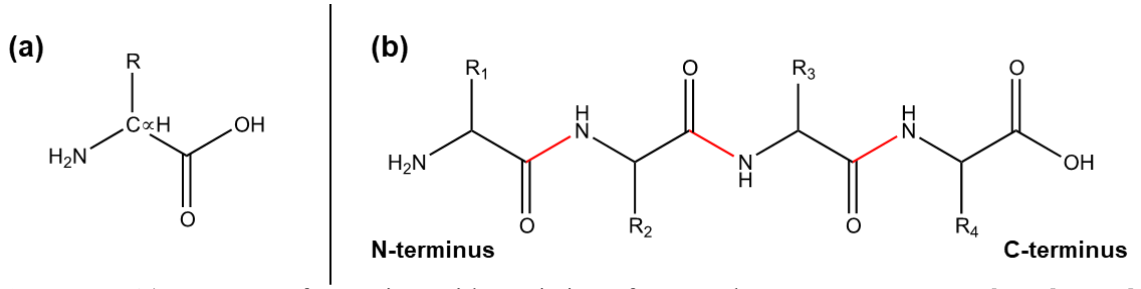

Figure 1.1. (a) Structure of an amino acid consisting of a central CαH, an amine, a carboxylic acid, and a variable R group. (b) A polypeptide chain consisting of an N/C-terminus. The red bond signifies the peptide bond formed between two amino acids.

1.1.2 Protein Folding

Protein folding is based on the classic principle that all the information required for a protein to adopt the correct three-dimensional conformation is encoded in its amino acid sequence. This concept was first discovered by Christian B. Anfinsen who demonstrated the spontaneous refolding of RNase A into its active conformation with full return of catalytic activity.[4] Ever since, understanding the nature of protein folding has been a compelling area of research and is commonly investigated by molecular dynamic (MD) simulations[5] and spectroscopic techniques.[6] For a protein to fold, their sequence must satisfy two requirements; a thermodynamic and a kinetic one.[3, 7] The thermodynamic requirement is that the protein must adopt a unique folded conformation which is stable under physiological conditions i.e. their native structure. The kinetic requirement is that

the protein folding process to the native state must occur within a reasonable amount of time i.e. typically within milliseconds to seconds.[8-9]

In order to describe the folding mechanism of proteins, the folding funnel model is used. The energy landscape of a protein is in the form of a multidimensional "funnel" where the vertical axis (depth of the funnel) represents the relative free energy of the conformational space, and the horizontal axis corresponds to the conformational entropy of the system.[10] The native state of a protein resides in a local minima typically at the bottom of the folding funnel with a low entropy and a global free energy minimum.[11] This model is based on the idea that there are many thermodynamic conformations in which the free-energy landscape of a protein exhibits several local minima separated by barriers that are associated with a specific polypeptide conformation. These local minima may result in the transient formation of partially folded species, accumulation of intermediates, or misfolded forms, depending on the trajectory taken by the protein along the energy surface.[12] The folding process involves a stochastic search of the many conformations accessible to a polypeptide chain.[13] As a polypeptide undergoes many conformational changes, there is a significant decrease in the system's conformational entropy as it approaches the native state.[14] Since there is a significant decrease in entropy as the protein becomes conformationally restricted, the enthalpic gain of forming hydrogen bonds and making favourable hydrophobic and hydrophilic contacts compensates for this entropic penalty.[10]

1.1.3 Protein Dynamics

The structure is an important aspect related to protein function however, it only represents a static image of a specific native conformation. Proteins are not static entities as they fluctuate naturally in solution. In recent years, it has become clear that in addition to structure, the dynamics of proteins play a significant role in their function.[15-16] Henzler-Wildman *et al* demonstrated that enzymatic activity requires a precise balance between flexibility and stability.[17] To function, enzymes must be stable enough to retain their native three-dimensional structure, but flexible enough to allow sufficient substrate binding. Protein dynamics are involved in a wide range of functions including catalytic turnover of enzymes, signaling/regulation, allostery, and some dynamics can enable some proteins to perform multiple functions.[18] In the case of catalytic turnover of enzymes, Eisenmesser *et al* have shown that during catalytic action of the enzyme cyclophilin A, conformational changes occurring in the active site are correlated with catalytic turnover of the enzyme.[19] This finding exemplifies the importance of enzyme dynamics and how they pertain to their catalytic function. Ultimately, protein function depends on their interaction with various other molecules such as cofactors, ligands, or other proteins. Depending on the specific molecules present and the physiological conditions, a protein can fluctuate between many different conformations. Protein dynamics are characterized by the timescales and the extent of the fluctuations.[15] These dynamic changes can be outstandingly apparent or subtle compared to their native conformation. The timescale and magnitude of protein dynamics are tied to a set of temperature, pH and solvent conditions. These conditions can be experimentally controlled to get a better understanding of the properties of protein dynamics to fully understand the role they have on their function.

Ultimately, the specific three-dimensional structure a protein adopts in its biological setting controls the protein's dynamic characteristics. This triad of structure, dynamics, and function is being actively studied to attain a comprehensive understanding of a protein's role in biological systems.[20-21]

1.1.4 Enzymes

Enzymes are proteins that act as catalysts for virtually all biological reactions. Enzymes are efficient catalysts that can accelerate reactions up to nineteen orders of magnitude.[22] They act by lowering the activation energy barrier of biochemical reactions, thereby increasing the reaction rate. The factors that enable enzymes to provide the large enhancement of reaction rates remain challenging to understand. Enzymes catalyze reactions on a wide range of timescales which are similar to the time-scales of various internal protein dynamics.[17] It is well known that protein dynamics is an essential aspect of enzyme function i.e., cofactor/substrate binding, catalytic conversion, and product release. Understanding their internal protein motions may give insights into enzyme catalysis.[23-24]

Enzymes are the key to carrying out biochemical processes, acting in a specific sequence of events which enable them to catalyze reactions that degrade nutrients, transform chemical energy, and synthesize biological macromolecules from simple precursors.[2] The activity of any given enzyme depends on several parameters such as substrate concentration, temperature, and pH. Assuming that an enzyme catalyzes a one-substrate reaction, the kinetics follow a simple mechanism shown in equation 1.1:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \qquad 1.1$$

This mechanism is characterized by a pre-equilibrium between the enzyme-substrate complex (ES) and the enzyme (E) and substrate (S) with a rate of k_1 and k_{-1}, respectively. This is followed by an irreversible decay (k_2) of the enzyme-substrate complex into the enzyme and products (P).[25] The Michaelis-Menten equation for this mechanism is expressed in equation 1.2:

$$v = \frac{V_{max} \cdot [S]}{K_M \cdot [S]} \qquad 1.2$$

where v is the rate of reaction, V_{max} is the maximum reaction rate which can only be enhanced by increasing the enzyme concentration, *[S]* is the substrate concentration, and K_M is the Michaelis constant that corresponds to the substrate concentration at ½ V_{max}. If [S] << K_M the rate of reaction follows first order kinetics. If [S] >> K_M, the reaction rate will reach saturation according to equation 1.3:

$$V_{max} = k_{cat} \cdot [E]_o \qquad 1.3$$

where k_{cat} and $[E]_o$ is the turnover number of the enzyme and initial enzyme concentration, respectively. [S], K_M, and $[E]_o$ define an enzyme's activity; these parameters are often measured and reported to characterize enzyme kinetics under different conditions.

Enzymes are intricate chemical-reaction machines and they often require other molecules to help carry out their function. These molecules are called *cofactors*. The most common cofactors are metal ions. Common examples of such cofactors are Zn^{2+}, $Fe^{2+/3+}$,

K^+, and Mg^{2+}. Cofactors can facilitate the following; binding and orienting the substrate in the active site, formation of electrostatic contacts with reaction intermediates or they can interact with a substrate to render them more electrophilic or nucleophilic.[26] A subgroup of cofactors called *prosthetic groups* are tightly integrated into the enzyme's structure by covalent or non-covalent interactions. These can be organic or inorganic molecules. Some common examples of prosthetic groups are flavin nucleotides, pyridoxal phosphate, biotin, or metal-containing porphyrins.

1.2 Methods for Studying Protein Structure and Dynamics

Numerous techniques are available to investigate the structure and dynamics of proteins. All the various experimental methods have advantages/disadvantages and vary in what type of information they can reveal about proteins and protein complexes.

1.2.1 X-ray Crystallography

X-ray crystallography is a technique that provides three-dimensional atomically resolved structures from a crystal. This works by exposing a crystalline sample to an X-ray beam. This process generates a unique diffraction pattern due to photon interactions with electrons within the crystal. Measuring the intensities of the diffraction pattern can reveal where the atoms are spatially located and this information is used to build a three-dimensional structure.[27]

X-ray crystallography remains the gold standard for atomically resolved structural data. *RCSB Protein Data Bank* (PDB) is a repository for three-dimensional biomolecular

structural data. Over 100,000 structures have been determined by crystallographic methods.[28] RCSB PDB statistics show that crystal structures can routinely be resolved under 2 Å. X-ray crystallography can be applied to protein complexes ranging all the way to mega Daltons.[29-30] Not only can this technique reveal high-resolution structures of large protein systems, it can also be applied to protein-ligand, protein-protein, and protein-DNA complexes.[31]

The main challenge associated with this technique is the crystallization process. Not all proteins form high-quality crystals. For X-ray crystallography to be successful for protein systems, the protein needs to be concentrated and high purity for a high quality, homogenous crystalline material.[32] For some protein systems, high concentrations can lead to aggregation.[33] As well, X-ray crystallography remains extremely challenging for disordered proteins.[27] Protein-protein contacts in crystal packing can differ from physiological protein-protein contacts in solution.[34] Overcoming the challenges of crystallization remains the biggest roadblock for this technique.

Nonetheless, surpassing all the challenges of X-ray crystallography immaculate structural data can be obtained. However, the data that is obtained only provides a static snapshot of proteins in their native structure. This type of information provides very little insight into their dynamics.[15] Due to this pitfall, researchers have developed other techniques that are able to provide high-resolution structural data as well as report on dynamic properties of protein systems.

1.2.2 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a technique that utilizes atomic nuclei that are intrinsically magnetic. For example, only selected isotopes such as ^{1}H, ^{13}C, ^{15}N, ^{17}O, and ^{19}F have a property called a *spin*. When spin-active nuclei is placed in a magnetic field and radio waves are applied, the nuclei will absorb a certain amount of energy. Due to this energy absorption at a given magnetic field strength, nuclei in different chemical environments will *resonate* at slightly different frequencies. These frequencies are reported as *chemical shifts* (δ) that are expressed in parts per million (ppm). As a result, chemical shift values are unique to a certain chemical environment of a spin active nuclei.[35] Numerous experiments can be designed to study changes to a particular chemical group under different conditions.

In 1952, the Nobel Prize in Physics was awarded to Felix Bloch and Edward Mills Purcell for their development of new methods for nuclear magnetic precision measurements of bulk materials.[36-37] Traditionally, NMR spectroscopy has been viewed as an analytical technique for the characterization of small organic and inorganic molecules. In 2002, Kurt Wüthrich won the Nobel Prize in Chemistry for developing the concept of interrogating biological macromolecules in solution using NMR spectroscopy. With the emergence of high-field magnets, this concept become possible to study macromolecular structures using NMR spectroscopy. Ever since, NMR spectroscopy has emerged as a major rival of X-ray crystallography for the determination of three-dimensional protein structures.

Like X-ray crystallography, NMR spectroscopy can generate high-resolution structural data on protein systems. However, NMR spectroscopy allows the determination

of atomically resolved structures in solution under near-native conditions.[38] Spin relaxation experiments are commonly used to investigate protein dynamics. For example ^{15}N experiments monitor the reorientation of ^{15}N – NH bonds, which are used to study backbone dynamics or side chain motions of residues containing nitrogen.[39] Methods have also been developed that measure ^{2}H relaxation properties at methyl (CH_2D) and methylene (CHD) side chain positions.[40] Two-dimensional homonuclear and heteronuclear experiments can be employed to help with spatial resolution of large protein systems.[41-42] Another popular NMR technique used for studying conformational dynamics is amide hydrogen/deuterium exchange (HDX).[43-44]

Since NMR spectroscopy has made many advances over the past half century, there are challenging experiments that can be implemented to study large protein complexes reaching 1 MDa.[45] The routine application of NMR methods to large protein systems (>40kDa) remains difficult. This is due to peak congestion, broadening, and overlap. Also, NMR requires proteins to be stable at high concentrations (0.1 to 1 mM) for hours or days.[46] Despite its size and concentration limitations, NMR spectroscopy remains a promising technique to study the structure and dynamics of protein complexes.

1.2.3 Optical Spectroscopy

Optical spectroscopy techniques may be the most popular methods to study protein complexes. This is due to their wide application, easy operation and their amenability to almost all protein systems. All spectroscopic methods fall under the umbrella of low resolution techniques. This is due to the fact that spectroscopic techniques provide some insights into global structural changes while detailed structural features remain hidden.[47]

1.2.3.1 Ultraviolet-Visible Spectroscopy

Ultraviolet-visible (UV-Vis) spectroscopy is an optical method used for many biological applications. Ultra-violet and visible light cover the 200-400 nm and 400-700 nm range, respectively, of the electromagnetic spectrum. When light of this energy interacts with biological molecules, it produces an electronic transition where π electrons or non-bonding electrons are promoted from their ground state to a high-energy state. The most basic experiments utilize the protein's side chain π bonds. The simplest example of this is used in determining a protein's concentration at wavelength 280 nm where the aromatic side chains absorb light. This is determined by using Beer Lambert's Law:

$$A = \varepsilon \cdot C \cdot l \qquad 1.4$$

where A is the absorbance, ε is the molar absorptivity, C is protein concentration, and l is the cuvette path length. If a protein contains an intrinsic chromophore, protein dynamics and structural changes can be investigated using UV-Vis methods. This approach is common for heme-containing proteins.[48-49] Heme is a porphyrin ring that contains an Fe^{2+} or Fe^{3+} center. The Soret peak at ~400 nm is distinctive to heme-proteins.[50] The spectral features in this region can give information regarding conformational changes that occur upon ligation or reduction/oxidation of the iron.[51]

1.2.3.2 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy measures the differences in the absorption of left- and right-circularly polarized light, a phenomenon that is generally encountered for chiral molecules. One of the most important applications of CD spectroscopy is studying the secondary structure of proteins. Different structural elements have characteristic CD

spectra in the 180-260 nm range. For example, spectra show unique minima for α-helices at 208 and 222 nm, β-sheets at 218 nm, and random coils at 195 nm.[52] CD spectroscopy can also be used to examine protein tertiary structure in the region of 280-320 nm, although this is less common.[53] Overall, CD spectroscopy is best for investigating whether a protein is folded, for monitoring the conformational stability of a protein under stress conditions (temperature, pH, denaturants), and for observing whether intramolecular binding interactions affect the secondary structure of a protein.[54]

1.2.3.3 Fluorescence Spectroscopy

Fluorescence spectroscopy studies protein behaviour in solution by monitoring photons emitted from a sample after it has absorbed light. This process is mediated by fluorophores with conjugated π bonds. Tryptophan (Trp) side chains are the most common intrinsic fluorophore used as a structural probe since they are sensitive to changes in the polarity of their environment.[55] By measuring the emission of Trp at λ_{max}=350 nm local environment properties of the protein can be investigated. For example, Trp residues are generally located in the core of a globular protein. However, when a protein becomes partially or completely unfolded Trp becomes solvent accessible giving insight into the overall structure of a protein. Many experiments can be employed to gain valuable information on folding/unfolding processes, subunit association, and substrate binding.[56-57]

1.3 Mass Spectrometry

Mass spectrometry (MS) measures the mass-to-charge ratio (m/z) of gas phase ions, allowing the implementation of a wide range of experiments. MS has become an important tool in structural biology. Investigating protein systems using MS requires a soft ionization (without fragmentation) technique to get biomolecules into the gas phase. Two ionization techniques that are used for studying proteins are matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI). In MALDI, the sample is prepared by mixing with a solution of an energy-absorbing, organic compound called matrix. When the matrix crystallizes upon drying, the sample within it co-crystallizes. The matrix is then exposed to a laser pulse, where through desorption and ionization analyte/matrix ions are produced.[58] Today, MALDI is primarily used for imaging applications. ESI has excelled as the preferred ionization technique to provide information on protein's solution-phase characteristics.[59-60] ESI was used in this work and in the following section it will be discussed in detail.

1.3.1 Electrospray Ionization

In the 1990s Fenn *et al* introduced the concept of ESI.[61] ESI is a process that permits the transfer of analytes from solution into the gas phase. Due to its 'gentle' nature, it has prevailed as the choice of ionization that is necessary for numerous MS applications particularly biological macromolecules.

ESI is initiated at atmospheric pressure by applying a high voltage to a metal capillary that acts as an outlet for the analyte solution. ESI can work in both positive and negative ion mode. In positive ion mode (that is exclusively used in this work), the high

positive voltage applied causes charge separation (for example, due to oxidation of water; $2H_2O \rightarrow 4H^+ + 4e^- + O_2$),[62] where electrons are removed from solution and the positively charged ions will emerge at the end of the capillary outlet. At this outlet, the solution will form a Taylor cone. Due to excess positive charge repulsion within the Taylor cone, the liquid becomes unstable and produces a plume of small charged droplets. These droplets undergo evaporation. The charge density on the shrinking droplet builds up until the surface tension γ is balanced by Coulombic repulsion. This situation corresponds to the Rayleigh limit, where the number (z_R) of elementary charges e is given by[63]

$$z_R = \frac{8\pi}{e}\sqrt{\varepsilon_0 \gamma r^3} \qquad 1.5$$

where r is the droplet radius, and ε_0 is the vacuum permittivity. Droplets at the Rayleigh limit undergo jet fission into smaller progeny droplets. This process repeats itself through several generations until desolvated gas-phase ions are released.[64] ESI produces multiply charged $[M + zH]^{z+}$ gas phase ions. Due to the high charges states formed during ionization, ESI has no upper size limit for studying proteins. Routinely ESI allows the transfer of kDa to MDa proteins into the gas phase.[65-66]

Three different mechanisms have been proposed for the final steps of the ESI process, i.e., the release of gas phase ions from nanometer-sized ESI droplets (Figure 1.2). A mechanism that likely applies for low molecular weight species and atomic ions is the ion evaporation model (IEM). This model envisions that direct emission of the analyte from the droplet will occur after the droplet radius is less than 10 nm. It is suggested that by

removing charge through this process, it suppresses the Coulomb fission of "late" ESI droplets.[64] Large globular proteins are released into the gas phase by the charged residue model (CRM). This model suggests that analyte in a droplet evaporate to dryness. The remaining charge in the solvent before the last evaporation step is believed to be transferred to the protein.[67] Kim *et al* recently demonstrated using MD simulations that peptides most likely become gas phase ions through the CRM process as well.[68] Lastly, unfolded proteins are believed to proceed through the chain ejection model (CEM). Since an unfolded protein will have solvent accessible hydrophobic residues, it will travel to the droplet surface where it will get ejected while taking charge with it.[69]

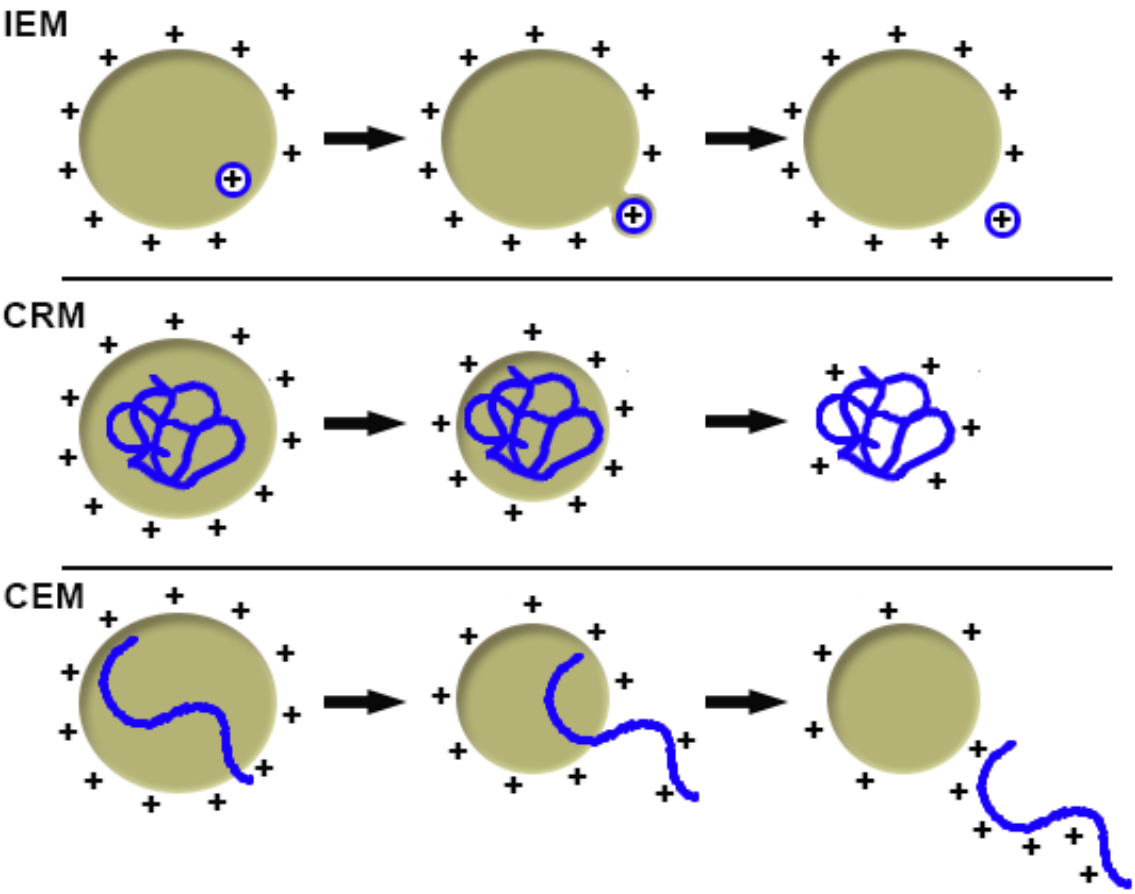

Figure 1.2. A schematic illustrating how gas phase ions are produced through the electrospray process through the IEM, CRM, and CEM mechanisms.

1.3.1.1 Protein Analyses

The investigation of proteins by MS can be carried out in many ways. Due to the gentle nature of ESI, some experiments investigate intact proteins from solution into the gas phase. Most popularly though, many MS experiments study proteins on the peptide level. These experiments involve the use of proteases, where proteins are enzymatically digested and then analyzed by MS. Two common proteases used are trypsin and pepsin. Trypsin cuts selectively after Lys and Arg residues.[70] This is advantageous as relatively few and well defined peptides are generated. For the analyses used in this work, pepsin digestions were employed. Pepsin has a fairly non-selective preference for what sequence it will cleave.[71] Even though the digestion patterns are somewhat unpredictable, the reproducibility of cleavage patterns is quite high under identical experimental conditions.[72] The use of a proteases to create peptides prior to analysis with MS is termed a *bottom-up* approach.

Since bottom-up approaches are commonly used, MS is routinely coupled with liquid chromatography (LC). This "LC/MS" approach allows the desalting of peptides or other analytes, allowing for higher quality mass analysis. As well, prefractionation of analyte prior to mass analyses simplifies MS measurements.

1.3.2 Mass Analyzers

The mass analyzer is used to separate gaseous ions based on their m/z. Various types of mass analyzers exist; these include quadrupoles, ion traps, Orbitraps, Fourier transform ion cyclotron (FT-ICR), and time-of-flight (TOF) analyzers. Quadrupoles and

TOF mass analyzers have been used in this work and will be discussed in detail in the following sections.

1.3.2.1 Quadrupole Mass Analyzers

A quadrupole mass analyzer consists of four parallel cylindrical rods. Each pair of opposing rods is always at the same electric potential. When a radio frequency (RF) voltage is applied to the quadrupole, all ions with different m/z will pass through the quadrupole and reach the detector as shown in Figure 1.3a. In this 'RF-only' mode, the quadrupole acts as an ion guide.[73] When both an RF and direct current (DC) voltage is applied to the rods, the quadrupole only allows the transmission of a certain m/z. This is due to the ion's stable trajectory through the quadrupole, where it then reaches the detector. Ions with other m/z values travelling through the quadrupole will have an unstable trajectory and collide with the rods and become neutralized as shown in Figure 1.3b. The changing of these voltages can result in the transmission of various m/z values.[74]

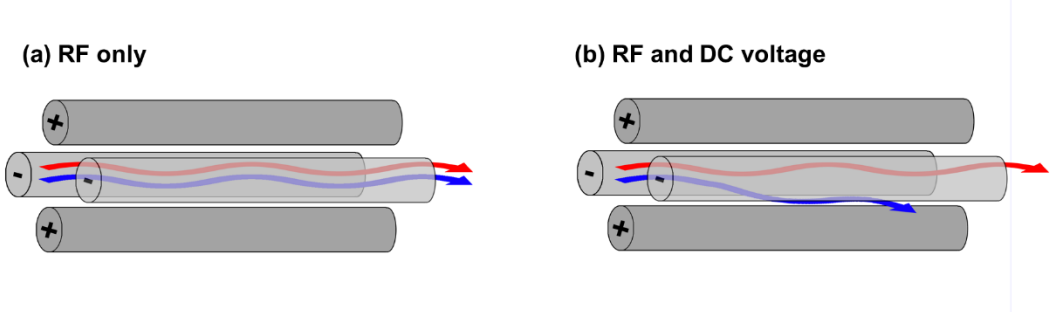

Figure 1.3. Schematic of a quadrupole mass analyzer. (a) In RF-only mode, all ions can be transmitted through the quadrupole to the detector. (b) When a certain RF/DC ratio is applied to the rods, only a specific m/z can pass through the quadrupole. All other m/z values will have an unstable trajectory and eventually hit the rods and become neutralized.

Quadrupole mass analyzers have relatively low resolution (~2000) but offer excellent sensitivity.[10] The m/z range of typical quadrupole MS is limited to 4000. Many modern mass spectrometers use quadrupoles in combination with other ion optics elements. Commonly, today they are used in Q-TOF (quadrupole time-of-flight) and Triple-Quadrupole instruments for performing tandem-MS. When quadrupoles are used in conjunction with another mass analyzer this increases the resolution of these instruments by an order of magnitude.

1.3.2.2 Time-of-Flight Mass Analyzers

A TOF mass analyzer measures an ion's m/z based on its flight time through a flight tube. Ions are accelerated into the flight tube by an electric pusher pulse. This voltage pulse gives all ions with the same m/z the same potential energy, which is converted into kinetic energy.

$$E_{pot} = E_{kin} \tag{1.6}$$

$$zeU = \frac{1}{2}mv^2 \tag{1.7}$$

Where U is the voltage pulse applied, v is the velocity of the ion, m is the mass of the ion, and z is the charge state of the ion. From equation 1.7, the time t it takes for an ion to traverse the flight tube with a length l and reach the detector is

$$t = l\sqrt{\frac{\mathrm{m}}{\mathrm{e}}\frac{1}{\mathrm{z2U}}} \tag{1.8}$$

Demonstrated in equation 1.8, the flight time of an ion depends on its m/z. Ions with different m/z values have different velocities. Ions with a smaller m/z will traverse the flight tube faster than ions with a larger m/z. In reality, ions with the same m/z may have different flight times as a result of the ions not receiving the exact same amount of energy (U) in equation 1.7. This means their arrival time to the detector will vary. This time difference leads to peak broadening. A reflectron is a device that is inserted into the flight path to compensate for this deleterious effect.

A reflectron is an ion mirror that has an electric field applied to its electrodes. The reflectron reverses the flight direction of ions to the detector. Faster ions with the same m/z will penetrate deeper into the reflectron and their flight time will become longer. The flight time of the two ions with the same m/z will be corrected and arrive at the detector at the same time. The introduction of reflectrons into TOF instruments decreases ion transmission however, it increases the resolution of these mass analyzers to routinely 50,000.

1.3.2.3 Tandem Mass Spectrometry

Tandem MS (MS/MS) is a technique that requires more than one mass analyzer. For example, the first mass analyzer (MS1) selects ions of a particular m/z (parent ion) that represents a single species from the ion source. These species are most commonly peptides. The ions are accelerated through a collision cell containing inert gas (N_2) to induce fragmentation. In this work, this process is called collision-induced dissociation (CID). The m/z values of the fragment ions are measured using the second mass analyzer (MS2). Triple-quadrupole and Q-TOF instruments are most commonly used for tandem MS. A schematic of the Q-TOF used in this work is shown in Figure 1.4.

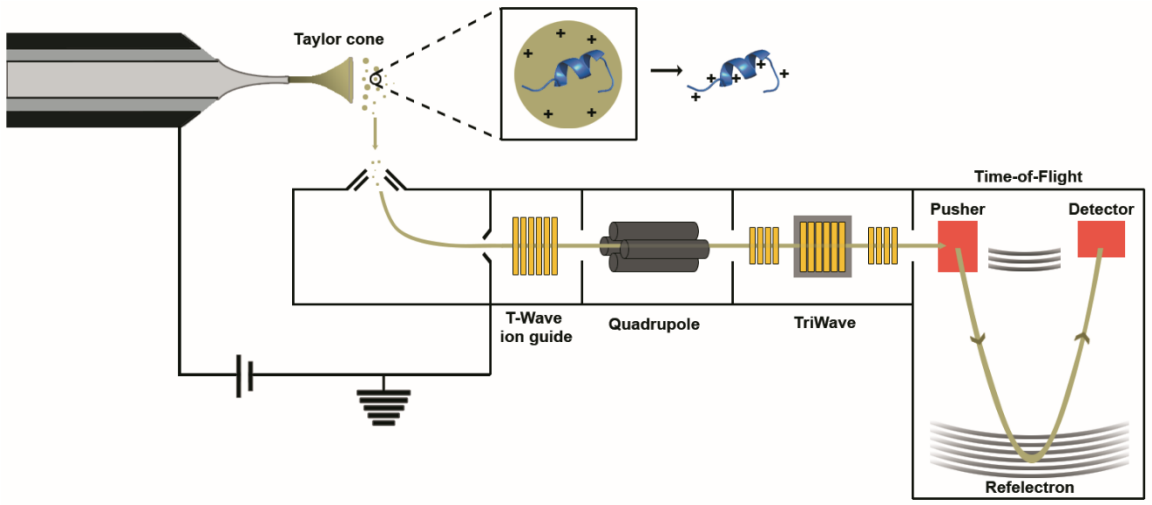

Figure 1.4. Electrospray ionization-mass spectrometer (Q-TOF) schematic used in this work. The main components of this Q-TOF is; the T-Wave ion guide, quadrupole, Tri-wave, and the time-of-flight detector.

In MS/MS experiments using this instrument, the quadrupole would be used for ion selection, the TriWave would be used for CID equipped with N_2 gas, and the generated ions would be detected using the TOF mass analyzer.

Fragmentation of proteins or peptides cause backbone cleavage. During CID, the peptide backbone will fragment once however, not all analytes will fragment in the same place. In this case, two types of ions are generated from the parent ion. In CID, the most common cleavage to occur is between the -OC-NH- bond resulting in *b and y ions*.[75] These ions are formed through cleavage of the peptide bonds, where b ions contain the N-terminus and y ions contain the C-terminus as displayed in Figure 1.5. The numbering indicates what peptide bond is being cleaved, counting from the N- or C-terminus.

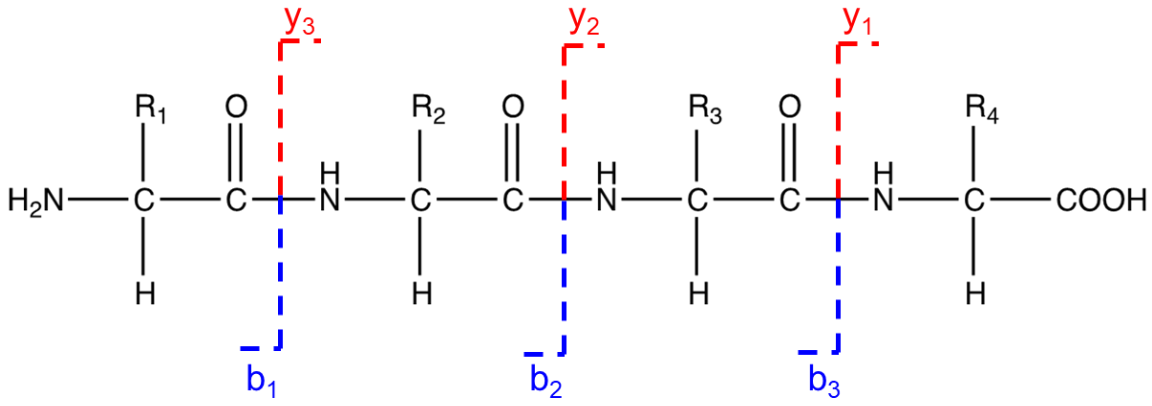

Figure 1.5. Nomenclature for b and y ions produced by CID. b ions (blue) contain the N-terminus and y ions (red) contain the C-terminus. The number indicates what peptide bond is being cleaved from the N- or C-terminus.

1.4 Structural Mass Spectrometry for Studying Protein Structure and Dynamics

The establishment of ESI in the early 1990s has open the doors to a vast array of MS methods to study macromolecular species. As the field evolves, structural MS techniques are routinely used to investigate protein structure and dynamics with ultimately no size limitations.

1.4.1 Native Mass Spectrometry

Native MS is a technique that is used to study intact protein complexes in the gas phase. For studying proteins in their native form, their structure must be retained from solution into the gas phase. Numerous examples in the literature show that proteins can retain 'native-like' structures in the vacuum which has established the field of gas-phase structural biology.[76-78] Native MS is used to identify proteins, sample heterogeneity, ligand binding, substrate turnover, structural topology, and dynamic properties of assembly.[77, 79]

To perform native MS experiments, the protein needs to be in ammonium acetate at pH ~7. The volatile nature of ammonium acetate (NH_4^+ (aq) + CH_3-COO^- (aq) → NH_3 (g) + CH_3 – COOH (g)) ensures that undesired adduct formation during ESI is minimized.[80] Once the protein complex is in the gas phase, various characteristics can be extracted from the mass spectra collected. For example, folded compact structures will give rise to narrow charge state distributions whereas unfolded conformers will generate a wide distribution of high charge states.[81]

Many research groups perform native MS studies coupled with ion mobility spectrometry (IMS).[82] IMS separates ions in the gas phase based on their drift time through an inert gas (N_2 or He) under the influence of a weak electric field E.[83] An ion with a more compact structure will have a smaller drift time due to less collisions with the background gas. In contrast, an unfolded structure will have a longer drift time.[84] By measuring the drift time t_d of ions with a charge z through a drift tube with length L, the collision cross section (CCS) Ω can be calculated using equation 1.9.

$$\Omega = z_{td}\frac{eE}{16NL}\left[\frac{18\pi}{\mu k_b T}\right]^{1/2}\frac{760\ Torr}{p}\frac{T}{273.2\ K} \qquad 1.9$$

where N is the background gas number density, μ the reduced mass of the ion-neutral pair, k_b is the Boltzmann's Constant, and T the gas temperature. This is another promising technique that is commonly used to ensure the solution-phase native structure is preserved in the gas-phase.[85-87]

1.4.2 Covalent Labeling

Covalent labeling is a technique that investigates conformational changes through modification reactions.[88] Solvent accessible or exposed side chain sites will react more quickly with probes or covalent labels compared to buried side chains on the protein. Information about a protein's structural changes is inferred from differential reactivity patterns of individual amino acids.[89] One commonly used covalent label is hydroxyl radicals (˙OH).[90] Quantitative analyses of labelled modifications to a protein are often conducted at the peptide level where the protein undergoes a tryptic digest before MS analyses. MS/MS experiments are employed to determine the labelled products.[91]

1.4.3 Covalent Cross-Linking

Like covalent labeling, chemical cross-linking also requires the addition of covalent modifications to the protein. This technique uses a bifunctional "cross-linker" of a specific length covalently coupled to two side chains on a protein that are separated by a certain distance. Various cross-link spacer lengths can be used to investigate different distance regimes.[92-93] Lysine side chains are the most commonly used cross-linking target however, other residues such as cysteine can also be used. The detection of cross-links using MS provides information on both structure and the interactions of proteins and protein assemblies.[94] Cross linking techniques have even been applied *in vivo*.[95-96]

Before the cross-linked protein is introduced into the MS, a trypsin digest is conducted. The analysis of cross-linked peptides can be difficult due to the low abundance of cross-linked peptides and a large portion of unmodified peptides present. Large databases have been created to help alleviate the analyses of cross-linking experiments.[97]

1.4.4 Hydrogen Deuterium Exchange

The development of hydrogen/deuterium exchange (HDX) began more than half a century ago.[98] HDX techniques monitor the exchange of labile amide hydrogens along the protein backbone with deuterons from D_2O solvent. Areas of the protein that are highly dynamic/unstructured or solvent accessible undergo fast exchange. In contrast, areas that are involved in hydrogen-bonded networks or are occluded from the solvent will undergo slower exchange. HDX can reveal various characteristics of a protein such as the presence or absence of protecting structure at peptide resolution, and aspects related to kinetic and thermodynamic stability.[99] HDX was initially used with NMR detection due to the different magnetic properties of ^1H and ^2H.[72] Since their magnetic properties are quite different (^1H spin= ½ whereas ^2H spin=1), ^2H would be silent in ^1H NMR and vice versa.[100] This approach is used to interrogate structural and dynamic properties of proteins and is still employed today.[101-102] However, in the early 1990s MS-measurements using HDX were introduced.[103] Since then, HDX has become an indispensable analytical tool that gives insights into the structure and dynamics of proteins.

Due to the 1 Da mass difference between hydrogen and deuterium, the incorporation of deuterium into a protein backbone is easily detected by MS. Figure 1.6 illustrates the workflow of a typical bottom-up experiment. The experiment starts with dilution of the protein in D_2O and the reaction proceeds for various time intervals ranging from seconds to hours (continuous-labeling). The exchange reaction is quenched by acidifying it to pH 2.4 and lowering the temperature to slow the exchange reaction down by orders of magnitude.[104] This is followed by online proteolysis by an acid-active enzyme.

Pepsin is almost exclusively used for digestion however; some other enzymes can be used as well.[105] Subsequently, the peptides are separated by LC and analyzed by MS.

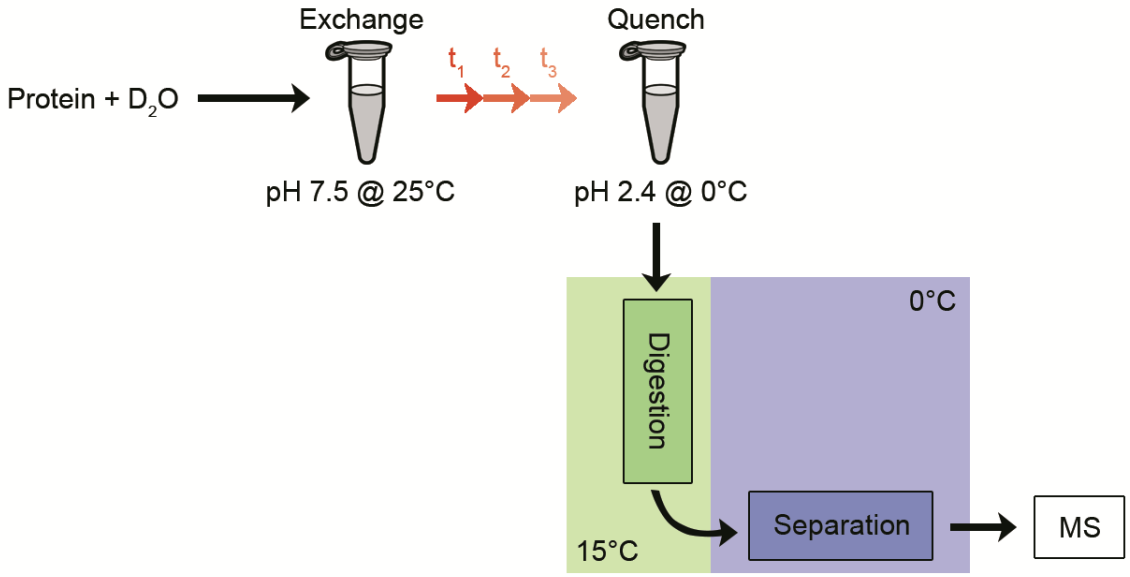

Figure 1.6. HDX workflow

The overall exchange mechanism for the backbone amides for a protein is usually described as

$$N - H_{closed} \underset{k_{cl}}{\overset{k_{op}}{\rightleftarrows}} N - H_{open} \xrightarrow{k_{ch}} N - D \qquad 1.10$$

where k_{op}, k_{cl}, and k_{ch} are the rate constants for an opening, closing, chemical transition, respectively. Each N-H group of the backbone has its own unique k_{op}, k_{cl}, and k_{ch} values. This exchange process is mediated by opening/closing conformational fluctuations where hydrogen bonds must be transiently broken.[102]

A freely exposed backbone amide hydrogen at pH 7 can undergo exchange with deuterons as quickly as 1 second when in D_2O.[106] However, protected amide hydrogens

(H-bonded or occluded from the solvent) rates spread over many orders of magnitude and may take weeks to months to undergo exchange.[99] The slow exchange rate of folded proteins are a result of intramolecular bonding or the restriction of solvent accessibility. The exchange rates of the slowest amide hydrogens can be reduced by 10^{-8} of their rates in unfolded forms of the same protein.[107]

1.4.4.1 EX1 and EX2 Exchange Kinetics

The overall rate for this mechanism in equation 1.10 is given by k_{HDX}. There are two limiting regimes that explain the HDX kinetics from equation 1.10. These are referred to as EX1 and EX2 kinetics. EX1 kinetics exhibits amide labeling that occurs during the first opening event where $k_{ch} >> k_{cl}$. The overall exchange rate for EX1 kinetics is simply

$$k_{HDX} = k_{op} \quad 1.11$$

In contrast, in the EX2 regime the probability of isotopic exchange occurring during a single opening event is small. It involves many opening/closing events to occur before deuterium exchange occurs.[108] Under EX2 conditions, $k_{cl} >> k_{ch}$ where the exchange mechanism can be characterized by:

$$k_{HDX} = K_{op}\, k_{ch} \quad 1.12$$

where $K_{op} = k_{op}/k_{cl}$ is the equilibrium constant for an opening event. Most proteins undergo EX2 exchange kinetics under physiological conditions, while EX1 is quite rare. However, EX1 kinetics can be induced by adding denaturants or by increasing the pH.[109] There are some examples where proteins can exhibit both EX1 and EX2 kinetics.[110] These two

regimes can be recognized by characteristic isotope patterns in mass spectra as shown in Figure 1.7.[111]

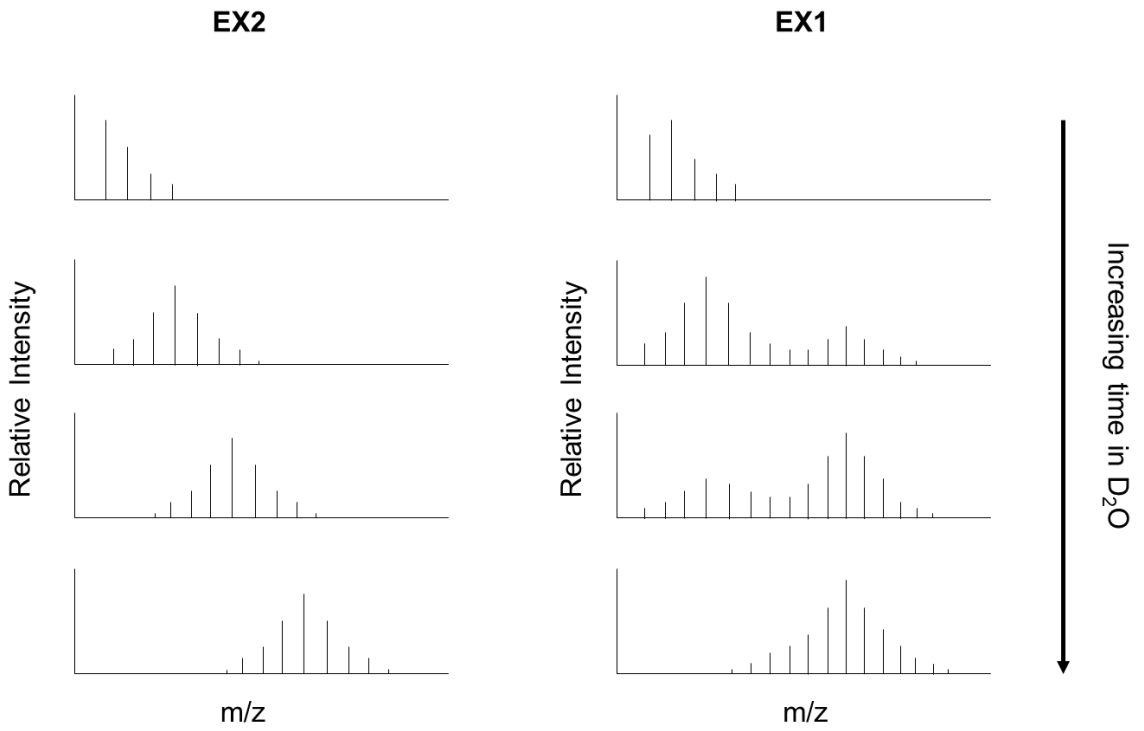

Figure 1.7. A cartoon schematic of characteristic isotope patterns of EX2 and EX1 mass spectra.

1.4.4.2 Dependence of k_{ch} on pH and Temperature

The pH dependence of k_{ch} is an important factor for the design of HDX experiments. An amide can exchange with solvent through acid-, or base-catalyzed reactions with a k_{ch} expressed by equation 1.13

$$k_{ch} = k_H[H^+] + k_{OH}[OH^-] \qquad 1.13$$

where k_H and k_{OH} are the intrinsic rate constants for the acid- and base-catalyzed reactions, respectively. Under physiological conditions, HDX proceeds with OH^- (or OD^-) catalysis.

The slowest exchange rates are observed around pH 2.4.[104] At pH values lower than 2.4 deuteration takes place with H^+ (or D^+) catalysis. The temperature also plays a critical role on the exchange kinetics. This relationship is dictated by the Arrhenius relationship

$$lnk_{ch} = lnA - \frac{E_a}{\mathrm{RT}} \qquad 1.14$$

where E_a is the activation energy for acid- or base-catalyzed exchanged ($E_a(k_H)$=14 kcal mol^{-1}, and $E_a(k_{OH})$= 17 kcal mol^{-1}); *R* is the gas constant, *T* is the temperature, and *A* is a constant.

Manipulating the pH and temperature during HDX experiments is essential for minimizing back-exchange. Back-exchange is an undesirable process where the deuterium on the backbone gets exchanged back with hydrogen from the solvent. When protein aliquots are removed after various time intervals, the pH is lowered to pH 2.4 to quench the exchange reaction. The quenching process slows the exchange rate down by four orders of magnitude. Decreasing the temperature to 0 ^oC results in an additional ~14-fold decrease. By simultaneously lowering the pH and temperature back exchange can thus be slowed by five orders of magnitude.[112] Chromatographic separation is performed at 0 ^oC in acidified mobile phase to minimize the back-exchange during LC runs. Side chains get labeled as well during incubation in D_2O however, these labels are removed due to quick 'back-exchange' during LC separation. The back-exchange of side chains is beneficial since HDX only measures the deuterium uptake of the amide hydrogens. All these different steps allow enough time for the determination of the deuterium uptake by LC-MS.

Valuable information is obtained by comparing the deuteration kinetics for different biochemical states of a protein. For example, this approach can be used for

interrogating protein-ligand or protein-protein interactions.[72] These interactions often manifest themselves as a reduction in deuterium uptake rates[113], although other scenarios are possible.[114] Soluble protein systems of all size remain routine for HDX experiments. However, HDX protocols for membrane proteins are still developing even though major advances have been achieved in the past few years.[115-117] HDX-MS methodologies have started to be exploited in the pharmaceutical industry in the development of therapeutics.[118-119]

1.5 Scope of Thesis

Proteins play a central role in all biological systems. Understanding how a protein functions has critical impact on many applications such as drug formulation, drug delivery, and cancer and disease therapies.[120-122] However, to fully understand a protein's function we must thoroughly understand their structure and dynamic characteristics. Many proteins follow the classic 'structure-function' paradigm meaning that a well-defined three-dimensional structure carries out a specific function. However, we know that proteins are not only static structures but also dynamic entities. Their dynamic properties play a key role in cofactor, ligand, and protein binding, having a direct impact on their functionally relevant three-dimensional structure.

The objective of this thesis is to use HDX-MS as an analytical technique to interrogate aspects of protein complexes that remain challenging with high-resolution techniques. These efforts will provide insights into structural and dynamic features that are related to the function of proteins that remain difficult to achieve with other techniques.

Chapter 2 investigates enzyme dynamics during catalysis. Structural fluctuations are essential for all aspects of enzyme function, i.e., substrate binding, catalytic conversion, and product release. It has been proposed that dynamics might be an 'intrinsic' property of enzymes, implying that structural fluctuations should remain unchanged in the resting state and during catalysis.[123] Other cases have shown that enzymes exhibit enhanced dynamics during catalysis.[124] HDX-MS is used as a sensitive tool for interrogating enzyme dynamics. HDX-MS comparative measurements are investigated during substrate turnover and in the resting state on an enzyme called pyruvate kinase. These comparative measurements are employed to give insights into the relationship between enzyme dynamics and catalysis.

Chapter 3 studies an intrinsically disordered protein (IDP) called nuclear factor erythroid 2- related factor 2 (Nrf2). IDPs are unique as they do not follow the structure-function paradigm. They have a high degree of dynamic flexibility and lack a well defined three-dimensional conformation. Research shows that IDPs carry out their specific function through protein-protein interactions where they undergo a disorder/order transition where they are thought to adopt their biologically relevant conformation.[125] High-resolution structural data on full length Nrf2 is not available. Many research groups have tried to investigate full length Nrf2 but have been unsuccessful due to its dynamic nature and its predisposition to aggregation during isolation and purification. HDX-MS is used in this work to understand the disordered nature of full length Nrf2, and what effect protein-protein interactions has on its structure and dynamics in the context of the full-length protein.

1.6 References

1. Redfern, O. C.; Dessailly, B.; Orengo, C. A., Exploring the structure and function paradigm. *Curr. Opin. Struct. Biol.* **2008,** *18* (3), 394-402.

2. Boyle, J., Lehninger principles of biochemistry (4th ed.): Nelson, D., and Cox, M. John Wiley & Sons Inc: USA, 2005; Vol. 33, pp 74-75.

3. Dobson, C. M.; Šali, A.; Karplus, M., Protein Folding: A Perspective from Theory and Experiment. *Angew. Chem. Int. Ed.* **1998,** *37* (7), 868-893.

4. Anfinsen, C. B., Principles that Govern the Folding of Protein Chains. *Science* **1973,** *181* (4096), 223-230.

5. Sali, A.; Shakhnovich, E.; Karplus, M., How does a protein fold? *Nature* **1994,** *369* (6477), 248-251.

6. Schuler, B.; Eaton, W. A., Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* **2008,** *18* (1), 16-26.

7. Karplus, M.; Šali, A., Theoretical studies of protein folding and unfolding. *Curr. Opin. Struct. Biol.* **1995,** *5* (1), 58-73.

8. Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L., Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling. *Anal. Chem.* **2013,** *85* (18), 8618-8625.

9. Davis, C. M.; Xiao, S.; Raleigh, D. P.; Dyer, R. B., Raising the Speed Limit for β-Hairpin Formation. *J. Am. Chem. Soc.* **2012,** *134* (35), 14476-14482.

10. Kaltashov, I. A.; Eyles, S. J., General Overview of Basic Concepts in Molecular Biophysics. In *Mass Spectrometry in Biophysics*, John Wiley & Sons, Inc.: 2005; pp 1-44.

11. Onuchic, J. N.; Wolynes, P. G.; Luthey-Schulten, Z.; Socci, N. D., Toward an outline of the topography of a realistic protein-folding funnel. *Pro. Natl. Acad. Sci. U. S. A.* **1995,** *92* (8), 3626-3630.

12. Tsai, C.-J.; Kumar, S.; Ma, B.; Nussinov, R., Folding funnels, binding funnels, and protein function. *Protein Sci.* **1999,** *8* (6), 1181-1190.

13. Dobson, C. M., Protein folding and misfolding. *Nature* **2003,** *426* (6968), 884-890.

14. Boehr, D. D.; Nussinov, R.; Wright, P. E., The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol* **2009,** *5* (11), 789-796.

15. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-972.

16. Karplus, M.; Kuriyan, J., Molecular dynamics and protein function. *Pro. Natl. Acad. Sci. U. S. A.* **2005,** *102* (19), 6679-6685.

17. Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D., A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **2007,** *450* (7171), 913-916.

18. Kleckner, I. R.; Foster, M. P., An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta* **2011,** *1814* (8), 942-968.

19. Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D., Enzyme Dynamics During Catalysis. *Science* **2002,** *295* (5559), 1520-1523.

20. Liang, B.; Tamm, L. K., NMR as a tool to investigate the structure, dynamics and function of membrane proteins. *Nat. Struct. Mol. Biol.* **2016,** *23* (6), 468-474.

21. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W., Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010,** *330* (6002), 341-346.

22. Radzicka, A.; Wolfenden, R., A proficient enzyme. *Science* **1995,** *267* (5194), 90-93.

23. Agarwal, P. K., Enzymes: An integrated view of structure, dynamics and function. *Microbial Cell Factories* **2006,** *5*, 2-2.

24. Benkovic, S. J.; Hammes-Schiffer, S., A Perspective on Enzyme Catalysis. *Science* **2003,** *301* (5637), 1196-1202.

25. Grunwald, P. P. D., *Biocatalysis: biochemical fundamentals and applications*. Imperial College Press: Hackensack, NJ;London;Singapore;, 2009.

26. Price, N. C.; Stevens, L., *Fundamentals of enzymology: the cell and molecular biology of catalytic proteins*. Oxford University Press: Oxford;New York;, 1999; Vol. 3rd.

27. Thibault, P.; Elser, V., X-ray diffraction microscopy. *Annu. Rev. Condens. Matter Phys.* **2010,** *1* (1), 237-255.

28. Su, X.-D.; Zhang, H.; Terwilliger, T. C.; Liljas, A.; Xiao, J.; Dong, Y., Protein Crystallography from the Perspective of Technology Developments. *Crystallogr. Rev.* **2015,** *21* (1-2), 122-153.

29. Hirano, Y.; Takeda, K.; Miki, K., Charge-density analysis of an iron–sulfur protein at an ultra-high resolution of 0.48 Å. *Nature* **2016,** *534* (7606), 281-284.

30. Shi, Y., A Glimpse of Structural Biology through X-Ray Crystallography. *Cell* **2014,** *159* (5), 995-1014.

31. Roemer, S. C.; Donham, D. C.; Sherman, L.; Pon, V. H.; Edwards, D. P.; Churchill, M. E. A., Structure of the Progesterone Receptor-Deoxyribonucleic Acid Complex: Novel Interactions Required for Binding to Half-Site Response Elements. *Mol. Endocrinol.* **2006,** *20* (12), 3042-3052.

32. Smyth, M. S.; Martin, J. H. J., x Ray crystallography. *J. Clin. Pathol.* **2000,** *53* (1), 8-14.

33. Hofmann, M.; Winzer, M.; Weber, C.; Gieseler, H., Prediction of Protein Aggregation in High Concentration Protein Solutions Utilizing Protein-Protein Interactions Determined by Low Volume Static Light Scattering. *J. Pharm. Sci.* **2016,** *105* (6), 1819-1828.

34. Carugo, O.; Argos, P., Protein-protein crystal-packing contacts. *Protein Sci.* **1997,** *6* (10), 2261-2263.

35. Encyclopedia of Analytical Chemistry In *Applications, Theory and Instrumentation* Meyers, R. A., Ed. 2011; p 2188.

36. Bloch, F.; Hansen, W. W.; Packard, M., The Nuclear Induction Experiment. *Phys. Rev.* **1946,** *70* (7-8), 474-485.

37. Purcell, E. M.; Torrey, H. C.; Pound, R. V., Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* **1946,** *69* (1-2), 37-38.

38. Montelione, G. T.; Zheng, D.; Huang, Y. J.; Gunsalus, K. C.; Szyperski, T., Protein NMR spectroscopy in structural genomics. *Nat. Struct. Mol. Biol.* **2000,** *7*, 982-985.

39. Kay, L. E., Protein dynamics from NMR. *Nat. Struct. Mol. Biol.* **1998**.

40. Millet, O.; Muhandiram, D. R.; Skrynnikov, N. R.; Kay, L. E., Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 1. Measurement of Five Relaxation Rates per Deuteron in 13C-Labeled and Fractionally 2H-Enriched Proteins in Solution. *J. Am. Chem. Soc.* **2002,** *124* (22), 6439-6448.

41. Sprangers, R.; Kay, L. E., Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* **2007,** *445* (7128), 618-622.

42. Kanelis, V.; Forman-Kay, J. D.; Kay, L. E., Multidimensional NMR Methods for Protein Structure Determination. *IUBMB Life* **2001,** *52* (6), 291-302.

43. Palmer, A. G., NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **2004,** *104* (8), 3623-3640.

44. Kuwata, K.; Matumoto, T.; Cheng, H.; Nagayama, K.; James, T. L.; Roder, H., NMR-detected hydrogen exchange and molecular dynamics simulations provide structural insight into fibril formation of prion protein fragment 106–126. *Pro. Natl. Acad. Sci. U. S. A.* **2003,** *100* (25), 14790-14795.

45. Rosenzweig, R.; Kay, L. E., Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR. *Annu. Rev. Biochem* **2014,** *83*, 291-315.

46. Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C., Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry—A review. *Anal. Chim. Acta* **2012,** *721*, 7-21.

47. Robinson, C. V.; Sali, A.; Baumeister, W., The molecular sociology of the cell. *Nature* **2007,** *450* (7172), 973-982.

48. Ghosh, K.; Thompson, A. M.; Oh, E.; Shi, X.; Goldbeck, R. A.; Zhiwu, Z.; Vulpe, C.; Holman, T. R., Spectroscopic and Biochemical Characterization of Heme Binding to Yeast Dap1p and Mouse PGRMC1p. *Biochemistry* **2005,** *44* (50), 16729-16736.

49. Nienhaus, K.; Nienhaus, G. U., Probing Heme Protein-Ligand Interactions by UV/Visible Absorption Spectroscopy. In *Protein-Ligand Interactions: Methods and Applications*, Ulrich Nienhaus, G., Ed. Humana Press: Totowa, NJ, 2005; pp 215-241.

50. Lykkegaard, M. K.; Ehlerding, A.; Hvelplund, P.; Kadhane, U.; Kirketerp, M.-B. S.; Nielsen, S. B.; Panja, S.; Wyer, J. A.; Zettergren, H., A Soret Marker Band for Four-Coordinate Ferric Heme Proteins from Absorption Spectra of Isolated Fe(III)-Heme+ and Fe(III)-Heme+(His) Ions in Vacuo. *J. Am. Chem. Soc.* **2008,** *130* (36), 11856-11857.

51. Kaminsky, L. S.; Miller, V. J.; Davison, A. J., Thermodynamic studies of the opening of the heme crevice of ferricytochrome c. *Biochemistry* **1973,** *12* (12), 2215-2221.

52. Brahms, S.; Brahms, J., Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J. Mol. Biol.* **1980,** *138* (2), 149-178.

53. Kelly, S. M.; Jess, T. J.; Price, N. C., How to study proteins by circular dichroism. *Biochim. Biophys. Acta* **2005,** *1751* (2), 119-139.

54. Kelly, S. M.; Price, N. C., The application of circular dichroism to studies of protein folding and unfolding. *Biochim. Biophys. Acta* **1997,** *1338* (2), 161-185.

55. Ghisaidoobe, A. B. T.; Chung, S. J., Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *Int. J. Mol. Sci.* **2014,** *15* (12), 22518-22538.

56. Roder, H.; Maki, K.; Cheng, H., Early Events in Protein Folding Explored by Rapid Mixing Methods. *Chem. Rev.* **2006,** *106* (5), 1836-1861.

57. Vivian, J. T.; Callis, P. R., Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* **2001,** *80* (5), 2093-2109.

58. Singhal, N.; Kumar, M.; Kanaujia, P. K.; Virdi, J. S., MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* **2015,** *6*, 791.

59. Cech, N. B.; Enke, C. G., Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom. Rev.* **2001,** *20* (6), 362-387.

60. Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R., Mass spectrometry-based methods to study protein architecture and dynamics. *Protein Sci.* **2013,** *22* (5), 530-544.

61. Yamashita, M.; Fenn, J. B., Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* **1984,** *88* (20), 4451-4459.

62. Van Berkel, G. J.; Kertesz, V., Using the electrochemistry of the electrospray ion source. *Anal. Chem.* **2007**.

63. Rayleigh, L., XX. On the equilibrium of liquid conducting masses charged with electricity. *Philosophical Magazine* **1882,** *14* (87), 184-186.

64. Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **2009,** *28* (6), 898-917.

65. Uetrecht, C.; Versluis, C.; Watts, N. R.; Roos, W. H.; Wuite, G. J. L.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. R., High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids. *Pro. Natl. Acad. Sci. U. S. A.* **2008,** *105* (27), 9216-9220.

66. Kaddis, C. S.; Lomeli, S. H.; Yin, S.; Berhane, B.; Apostol, M. I.; Kickhoefer, V. A.; Rome, L. H.; Loo, J. A., Sizing Large Proteins and Protein Complexes by Electrospray Ionization Mass Spectrometry and Ion Mobility. *J. Am. Soc. Mass. Spectrom.* **2007,** *18* (7), 1206-1216.

67. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013,** *85* (1), 2-9.

68. Kim, D.; Wagner, N.; Wooding, K.; Clemmer, D. E.; Russell, D. H., Ions from Solution to the Gas Phase: A Molecular Dynamics Simulation of the Structural Evolution of Substance P during Desolvation of Charged Nanodroplets Generated by Electrospray Ionization. *J. Am. Chem. Soc.* **2017,** *139* (8), 2981-2988.

69. Ahadi, E.; Konermann, L., Modeling the Behavior of Coarse-Grained Polymer Chains in Charged Water Droplets: Implications for the Mechanism of Electrospray Ionization. *J. Phys. Chem. B* **2012,** *116* (1), 104-112.

70. Olsen, J. V.; Ong, S.-E.; Mann, M., Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Mol. Cell. Proteomics.* **2004,** *3* (6), 608-614.

71. Hamuro, Y.; Coales, S. J.; Molnar, K. S.; Tuske, S. J.; Morrow, J. A., Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. *Rapid Commun. Mass Spectrom.* **2008,** *22* (7), 1041-1046.

72. Engen, J. R.; Wales, T. E., Analytical Aspects of Hydrogen Exchange Mass Spectrometry. *Annu. Rev. Anal. Chem.* **2015,** *8*, 127-148.

73. Henchman, M.; Steel, C., Understanding the Quadrupole Mass Filter through Computer Simulation. *J. Chem. Educ.* **1998,** *75* (8), 1049.

74. Leary, J. J.; Schmidt, R. L., Quadrupole Mass Spectrometers: An Intuitive Look at the Math. *J. Chem. Educ.* **1996,** *73* (12), 1142.

75. Hughes, C.; Ma, B.; Lajoie, G. A., De Novo Sequencing Methods in Proteomics. In *Proteome Bioinformatics*, Hubbard, S. J.; Jones, A. R., Eds. Humana Press: Totowa, NJ, 2010; pp 105-121.

76. Robinson, C. V., From molecular chaperones to membrane motors: through the lens of a mass spectrometrist. *Biochem. Soc. Trans.* **2017,** *45* (1), 251-260.

77. Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R., High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* **2012,** *9* (11), 1084-1086.

78. Leney, A. C.; Heck, A. J. R., Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass. Spectrom.* **2017,** *28* (1), 5-13.

79. Hyung, S.-J.; Ruotolo, B. T., Integrating Mass Spectrometry of Intact Protein Complexes into Structural Proteomics. *Proteomics* **2012,** *12* (10), 1547-1564.

80. Konermann, L.; Vahidi, S.; Sowole, M. A., Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. *Anal. Chem.* **2014,** *86* (1), 213-232.

81. Testa, L.; Brocca, S.; Santambrogio, C.; D'Urzo, A.; Habchi, J.; Longhi, S.; Uversky, V. N.; Grandori, R., Extracting structural information from charge-state distributions of intrinsically disordered proteins by non-denaturing electrospray-ionization mass spectrometry. *Intrinsically. Disord. Proteins* **2013,** *1* (1), e25068.

82. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V., Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* **2008,** *3* (7), 1139-1152.

83. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H., An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass spectrom.* **2007,** *261* (1), 1-12.

84. Barran, P.; Ruotolo, B. T., Ion Mobility Mass Spectrometry. *Analyst* **2015,** *140* (20), 6772-6774.

85. Chen, S.-H.; Russell, D. H., How Closely Related Are Conformations of Protein Ions Sampled by IM-MS to Native Solution Structures? *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (9), 1433-1443.

86. Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G., Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angew. Chem. Int. Ed.* **2016,** *55* (45), 14173-14176.

87. Wyttenbach, T.; Bowers, M. T., Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment. *J. Phys. Chem. B* **2011,** *115* (42), 12266-12275.

88. Xu, G.; Chance, M. R., Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007,** *107* (8), 3514-3543.

89. Mendoza, V. L.; Vachet, R. W., Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom. Rev.* **2009,** *28* (5), 785-815.

90. Konermann, L.; Stocks, B. B.; Pan, Y.; Tong, X., Mass spectrometry combined with oxidative labeling for exploring protein structure and folding. *Mass Spectrom. Rev.* **2010,** *29* (4), 651-667.

91. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W., Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999,** *121* (4), 806-812.

92. Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R., Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics. *Mol. Cell. Proteomics.* **2010,** *9* (8), 1634-1649.

93. Leitner, A.; Faini, M.; Stengel, F.; Aebersold, R., Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci* **2016,** *41* (1), 20-32.

94. Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R., Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **2015,** *12* (12), 1179-1184.

95. Holding, A. N., XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* **2015,** *89*, 54-63.

96. Bruce, J. E., In vivo protein complex topologies: Sights through a cross-linking lens. *Proteomics* **2012,** *12* (10), 1565-1575.

97. Liu, F.; van Breukelen, B.; Heck, A. J. R., Facilitating Protein Disulfide Mapping by a Combination of Pepsin Digestion, Electron Transfer Higher Energy Dissociation (EThcD), and a Dedicated Search Algorithm SlinkS. *Mol. Cell. Proteomics.* **2014,** *13* (10), 2776-2786.

98. Hvidt, A.; Nielsen, S. O., Hydrogen Exchange in Proteins. *Adv. Protein Chem.* **1966,** *21*, 287-386.

99. Krishna, M. M.; Hoang, L.; Lin, Y.; Englander, S. W., Hydrogen exchange methods to study protein folding. *Methods* **2004,** *34* (1), 51-64.

100. Englander, S. W.; Mayne, L., Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct.* **1992,** *21* (1), 243-265.

101. Roder, H.; Elöve, G. A.; Englander, S. W., Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* **1988,** *335* (6192), 700-704.

102. Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W., Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **2012,** *21* (7), 996-1005.

103. Katta, V.; Chait, B. T.; Carr, S., Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **1991,** *5* (4), 214-217.

104. Smith, D. L.; Deng, Y.; Zhang, Z., Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J. Mass Spectrom.* **1997,** *32* (2), 135-146.

105. Rey, M.; Yang, M.; Burns, K. M.; Yu, Y.; Lees-Miller, S. P.; Schriemer, D. C., Nepenthesin from Monkey Cups for Hydrogen/Deuterium Exchange Mass Spectrometry. *Mol. Cell. Proteomics.* **2013,** *12* (2), 464-472.

106. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* **2006,** *25* (1), 158-170.

107. Englander, S. W.; Kallenbach, N. R., Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **1983,** *16* (04), 521-655.

108. Konermann, L.; Pan, J.; Liu, Y.-H., Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **2011,** *40* (3), 1224-1234.

109. Deng, Y.; Smith, D. L., Identification of Unfolding Domains in Large Proteins by Their Unfolding Rates. *Biochemistry* **1998,** *37* (18), 6256-6262.

110. Xiao, Y.; Konermann, L., Protein structural dynamics at the gas/water interface examined by hydrogen exchange mass spectrometry. *Protein Sci.* **2015,** *24* (8), 1247-1256.

111. Miranker, A.; Robinson, C.; Radford, S.; Aplin, R.; Dobson, C., Detection of transient protein folding populations by mass spectrometry. *Science* **1993,** *262* (5135), 896-900.

112. Jensen, P. F.; Rand, K. D., *Hydrogen Exchange: A Sensitive Analytical Window into Protein Conformation and Dynamics*. 1 ed.; 2016.

113. Engen, J. R., Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Anal. Chem.* **2009,** *81* (19), 7870-7875.

114. Konermann, L.; Rodriguez, A. D.; Sowole, M. A., Type 1 and Type 2 scenarios in hydrogen exchange mass spectrometry studies on protein-ligand complexes. *Analyst* **2014,** *139* (23), 6078-6087.

115. Harrison, R. A.; Engen, J. R., Conformational insight into multi-protein signaling assemblies by hydrogen–deuterium exchange mass spectrometry. *Curr. Opin. Struct. Biol.* **2016,** *41*, 187-193.

116. Kim, D. K.; Yun, Y.; Kim, H. R.; Seo, M.-D.; Chung, K. Y., Different conformational dynamics of various active states of β-arrestin1 analyzed by hydrogen/deuterium exchange mass spectrometry. *J. Struct. Biol.* **2015,** *190* (2), 250-259.

117. Vadas, O.; Burke, John E., Probing the dynamic regulation of peripheral membrane proteins using hydrogen deuterium exchange–MS (HDX–MS). *Biochem. Soc. Trans.* **2015,** *43* (5), 773-786.

118. Anderson, K. W.; Mast, N.; Hudgens, J. W.; Lin, J. B.; Turko, I. V.; Pikuleva, I. A., Mapping of the Allosteric Site in Cholesterol Hydroxylase CYP46A1 for Efavirenz, a Drug That Stimulates Enzyme Activity. *J. Biol. Chem.* **2016,** *291* (22), 11876-11886.

119. Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F., Antibody Structural Integrity of Site-Specific Antibody-Drug Conjugates Investigated by Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2015,** *87* (11), 5669-5676.

120. Campbell, E.; Kaltenbach, M.; Correy, G. J.; Carr, P. D.; Porebski, B. T.; Livingstone, E. K.; Afriat-Jurnou, L.; Buckle, A. M.; Weik, M.; Hollfelder, F.; Tokuriki, N.; Jackson, C. J., The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol* **2016,** *12* (11), 944-950.

121. Liu, Z.; Huang, X.; Hu, L.; Pham, L.; Poole, K.; Tang, Y.; Mahon, B. P.; Tang, W.; Li, K.; Goldfarb, N. E., Implication of Natural Polymorphism in Hinge Region of HIV-1 Protease on Protein Conformations, Local Structures and Backbone Dynamics. *Biophys. J.* **2016,** *110* (3), 207a.

122. Wu, H.; Fuxreiter, M., The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* **2016,** *165* (5), 1055-1066.

123. Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D., Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007,** *450* (7171), 838-844.

124. Liuni, P.; Jeganathan, A.; Wilson, D. J., Conformer Selection and Intensified Dynamics During Catalytic Turnover in Chymotrypsin. *Angew. Chem. Int. Ed.* **2012,** *51* (38), 9666-9669.

125. Keppel, T. R.; Howard, B. A.; Weis, D. D., Mapping Unstructured Regions and Synergistic Folding in Intrinsically Disordered Proteins with Amide H/D Exchange Mass Spectrometry. *Biochemistry* **2011,** *50* (40), 8722-8732.

Chapter 2 - Changes in Enzyme Structural Dynamics Studied by Hydrogen Exchange-Mass Spectrometry: Ligand Binding Effects or Catalytically Relevant Motions?

2.1 Introduction

The catalytic power of enzymes is remarkable, with rate enhancements up to nineteen orders of magnitude to the corresponding uncatalyzed reactions.[1] This acceleration results from the lowered activation barriers, often in conjunction with the breakdown of reactions into multi-step sequences.[2-3] A central aspect of enzyme catalysis is the stabilization of transition states in the active site by H-bonds, the careful positioning of proton donors/acceptors, transiently formed covalent linkages, and the exclusion of water.[1, 4] Metal cofactors can provide electrostatic contacts, Lewis acid/base interactions, and electron shuttling.[5-6]

Like other proteins, enzymes undergo conformational dynamics. The corresponding motions cover a wide range, from subtle picosecond side chain fluctuations to major sub-second domain movements and unfolding/refolding transitions.[7-9] Enzyme dynamics are considered to be a key prerequisite for catalysis.[3, 10-18] This view is reinforced by X-ray data that show large structural changes upon binding of substrates or inhibitors, compared to the catalytically inactive resting state.[19-23] These crystallographically detectable changes typically lead from an unbound "open" conformation to a "closed" state where the substrate is occluded from the bulk solvent while catalysis takes place. Subsequent product release requires and opening event, before the next turnover event

commences.[24] Steady-state turnover must therefore be accompanied by incessant opening/closing transitions. Beyond binding and release, conformational fluctuations may also mediate catalytically important contacts that promote bond formation/dissociation.[3, 10-18] The exact implications of conformational dynamics for binding, release, and for the reshuffling of chemical bonds nonetheless remains controversial.[25-27]

A seemingly straightforward approach for deciphering the relationship between dynamics and catalysis is to compare the properties of working enzymes with those of the resting state. Experiments on dihydrofolate reductase suggest that substrates channel conformational motions along a specific reaction path.[28] The internal dynamics of chymotrypsin were reported to be enhanced during catalysis.[13] In contrast, the dynamics of thermolysin were indistinguishable during catalysis and in the resting state,[29] supporting the view that catalytically relevant motions are "intrinsic" and do not strongly depend on the presence of substrate (as previously suggested for the nano- to millisecond dynamics of adenylate kinase).[30] Thus a consistent picture has not emerged yet.

The interpretation of active/resting state comparative data is complicated by the fact that interactions with substrates, intermediates, or products may affect enzyme dynamics in ways that are not directly related to catalysis. Protein-ligand interactions generally affect the protein energy landscape, thereby altering conformational preferences.[31] Ligand-bound proteins are usually less dynamic than the corresponding apo-states. This stabilization arises from the cooperative stabilization of intramolecular linkages by intermolecular contacts.[32-45] However, local or global destabilization can take place as well.[46-49] It is therefore not obvious if altered enzyme dynamics during turnover are the result of catalytically relevant motions, of if they arise from rather trivial ligand

binding effects. For deciphering these different scenarios, it is necessary to supplement active/resting state comparisons with experiments that involve binding *without* turnover (e.g. by providing only one of two required substrates, or by exposing the enzyme to substrate mimics). It appears that the need for such control experiments has been under-appreciated in the literature.

Pyruvate kinase catalyzes the glycolytic conversion of phosphoenolpyruvate (PEP) and Mg-ADP to pyruvate and Mg-ATP. Four mammalian variants of this enzyme have been identified.[50-51] The L and R isoforms are found in liver and red blood cells, respectively. M2 is expressed in embryonic tissues and during tumorigenesis,[51-54] and it is allosterically regulated by fructose-1,6-bisphosphate (FBP).[55-56] The M1 isoform is found in adult brain tissue and skeletal muscle,[51, 53] it does not bind FBP and is constitutively active under typical conditions.[56]

Rabbit skeletal muscle pyruvate kinase (*r*M1-PK) represents the most extensively studied form of M1.[22, 57-62] The 232 kDa protein has a homotetrameric quaternary structure. Each subunit comprises four domains, termed N, A, B, and C. The active sites (one per subunit) are located in clefts between domains B and A (Figure 2.1).[22] Each active site contains one K^+ and one Mg^{2+} ion, for a total of two Mg^{2+} once Mg-ADP or Mg-ATP are bound.[63] The reaction (Figure 2.1b) involves binding of PEP and Mg-ADP with subsequent phosphate transfer from PEP to Mg-ADP. Mg-ATP then departs, leaving behind pyruvate in the enolate^{2-} form which has its negative charges coordinated by Mg^{2+}. Subsequent ketonization of the enolate produces pyruvate and supplies much of the driving force for the reaction.[22, 55] X-ray structures provide insights into some of the events associated with catalysis.[22] Crystals were grown in the presence of Mg-ATP and oxalate. The latter mimics

the enolate form of pyruvate.[64] The data showed three subunit conformations, a fully closed Mg-ATP/oxalate-bound state, a semi-open Mg-ATP/oxalate-bound structure, and an oxalate-bound open conformation (Figure 2.1c-e). The open/closed transition involves a 41° rotation of domain B relative to A. These structures[22] suggest that each turnover event is accompanied by large-scale opening/closing events of the active site, making *r*M1-PK an ideal model system for probing the relationship between dynamics and catalysis.

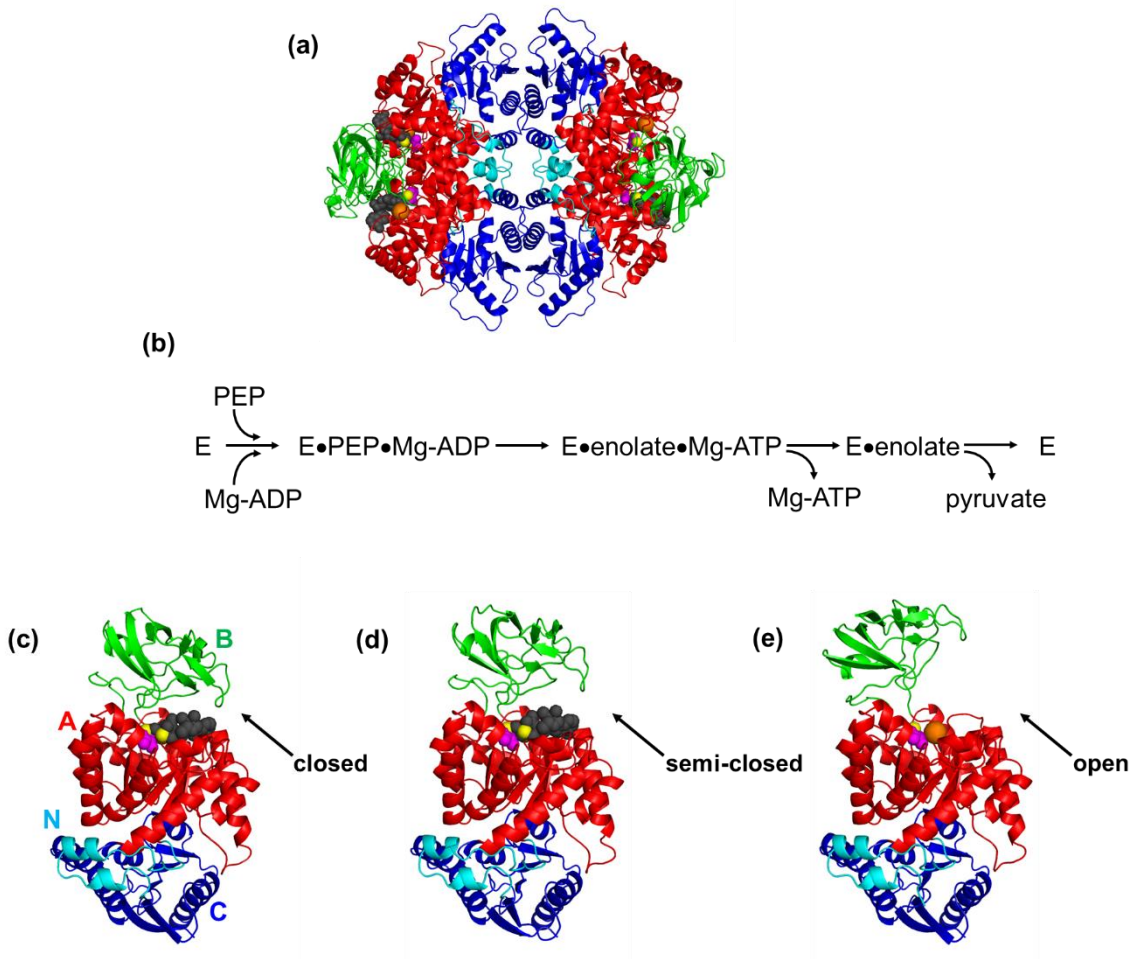

Figure 2.1. Crystal structure of rM1-PK (pdb code 1A49).[22] (a) Complete tetramer. Each subunit consists of four domains; N (cyan, residues 12-42), A (red, 43-115 and 219-387), B (green, 116-218); C (blue, 388-530). The active site is located in the cleft between domains A and B. (b) Catalytic reaction sequence, where the *r*M1-PK enzyme is denoted as "E". Dots indicate noncovalent protein-ligand contacts. Panels (c) – (e) illustrate active site closure for one of the subunits via rotation of domain B after Mg-ATP (gray) and oxalate (magenta) bind. K^+ and Mg^{2+} are shown in orange and yellow, respectively. (c) Closed state in the presence of Mg-ATP and oxalate. (d) Semi-closed state in the presence of Mg-ATP and oxalate. (e) Open state in the presence of oxalate, without nucleotide.

NMR spin relaxation experiments are a well-established tool for probing enzyme dynamics,[28, 30, 65] but the application of this technique to systems as large as *r*M1-PK is challenging. Also, NMR requires proteins to be stable at high concentrations (0.1 to 1 mM) for hours or days. Hydrogen/deuterium exchange (HDX) mass spectrometry (MS) represents an alternative approach[13, 16, 29] that overcomes these limitations.[35, 45, 66-68] HDX-MS reports on conformational dynamics by probing the stability of backbone amide hydrogen bonds. Segments that are tightly folded undergo slow deuteration, while flexible and/or disordered regions exhibit much faster HDX (see Chapter 1).[69] HDX-MS provides an integrative view of fluctuations that take place on time scales of microseconds to seconds.[70] This technique has previously been applied to M1 and other pyruvate kinase variants for probing allosteric regulation,[57, 71-72] but not for examining the relationship between dynamics and catalysis.

In the current work HDX-MS is used for characterizing the conformational dynamics of *r*M1-PK during catalysis, in the resting state, and in the presence/absence of various ligands that mimic different stages of the turnover process. The key question addressed in these experiments is to what extent working/resting state comparative measurements can provide information on catalytically relevant motions.

2.2 Methods

2.2.1 Materials

*r*M1-PK, Tris-Cl, ATP, ADP, potassium oxalate, rabbit muscle L-lactate dehydrogenase (LDH) and deuterium oxide were purchase by Sigma Aldrich (St. Louis,

MO). KCl and $MgCl_2$ were supplied by Fisher Scientific (Georgetown, ON). PEP and NADH (disodium salt) were purchased from Roche (Mississauga, ON). Stock solutions of metal-depleted *r*M1-PK were produced by protein incubation in 5 mM EDTA/50 mM Tris-HCl with subsequent dialysis against 50 mM Tris-HCl with multiple buffer exchanges. After dialysis, the *r*M1-PK concentration in the stock solution was 5 μM (as tetramer), as verified by UV-Vis measurements at 280 nm. All solutions had pH 7.5 and all experiments were carried out at room temperature (23 ± 1 °C), unless noted otherwise.

2.2.2 Enzyme Characterization

Several tests were conducted to verify the functional and structural integrity of *r*M1-PK. Activity measurements[73] were performed by monitoring the LDH-catalyzed reduction of pyruvate to lactate by NADH. The conversion of NADH to NAD^+ was monitored by UV-Vis spectroscopy at 340 nm resulting in a k_{cat} of 274 ± 30 s^{-1} per *r*M1-PK active site, which is close to the literature value of 250 s^{-1} per active site (Figure 2.2).[74]

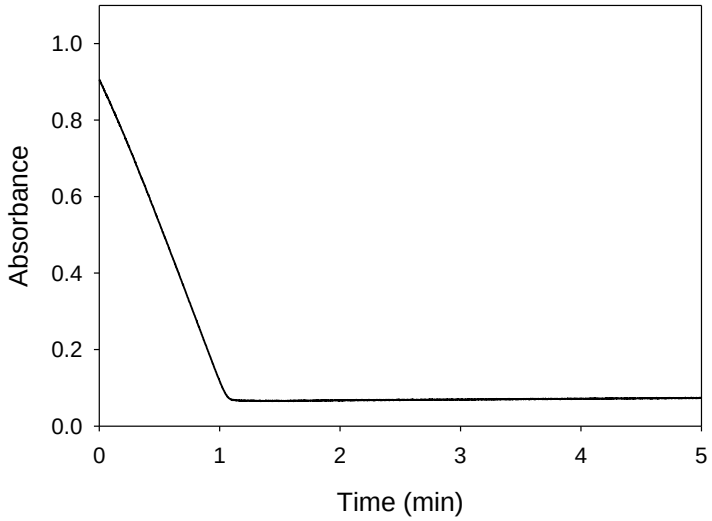

Figure 2.2. Pyruvate kinase activity was determined by the lactate dehydrogenase coupled assay.[1] Reaction assays contained 50 mM Tris buffer, 100 mM KCl, 8 mM $MgCl_2$, 1 mM PEP, 2 mM ADP, 0.25 mM NADH, 2.5 μM lactate dehydrogenase, and 0.002 μM *r*M1-PK (as tetramer). The reaction was initiated by adding 990 μL of the reaction mixture to 10 μL of *r*M1-PK. The assay was run in triplicates, resulting in k_{cat}= 274 ± 30 s^{-1} per active site.

Native nanoESI-MS confirmed that the enzyme existed in its canonical tetrameric quaternary structure shown in Figure 2.3.[75]

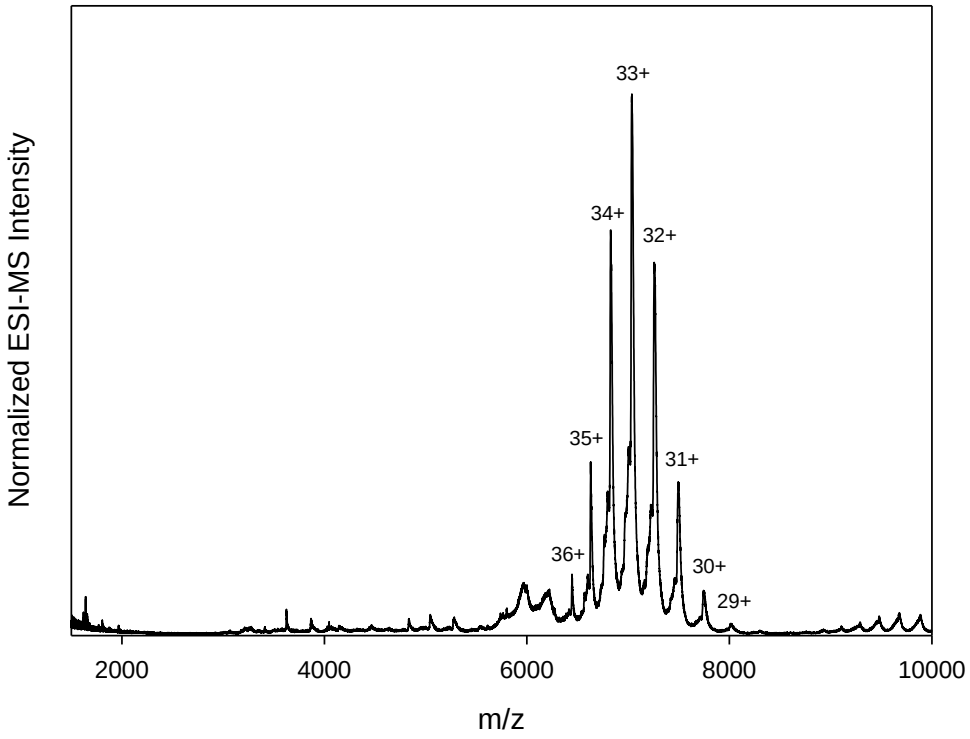

Figure 2.3. Native nanoESI mass spectrum of rM1-PK (protein concentration 2.5 μM as tetramer, metal-depleted sample) in aqueous solution containing 50 mM ammonium acetate, showing various charge states of the protein tetramer. ESI capillary voltage 1.8 kV, sample cone 175 V, extraction cone 3 V, source temperature 80 °C, desolvation temperature 200 °C. Similar data have been reported previously.[2]

Mass spectrum recoreded under denaturing conditions showed monomers with a measured mass of 57956 ± 3 Da, consistent with the *r*M1-PK sequence (Uniprot identifier PKM1/P11974-1) after loss of the N-terminal methionine and subsequent N-terminal acetylation (57959 Da, Figure 2.4).[76] These findings imply that the HDX kinetics discussed below are dominated by native functional *r*M1-PK; any "contaminating" contributions caused by degraded protein are minimal.

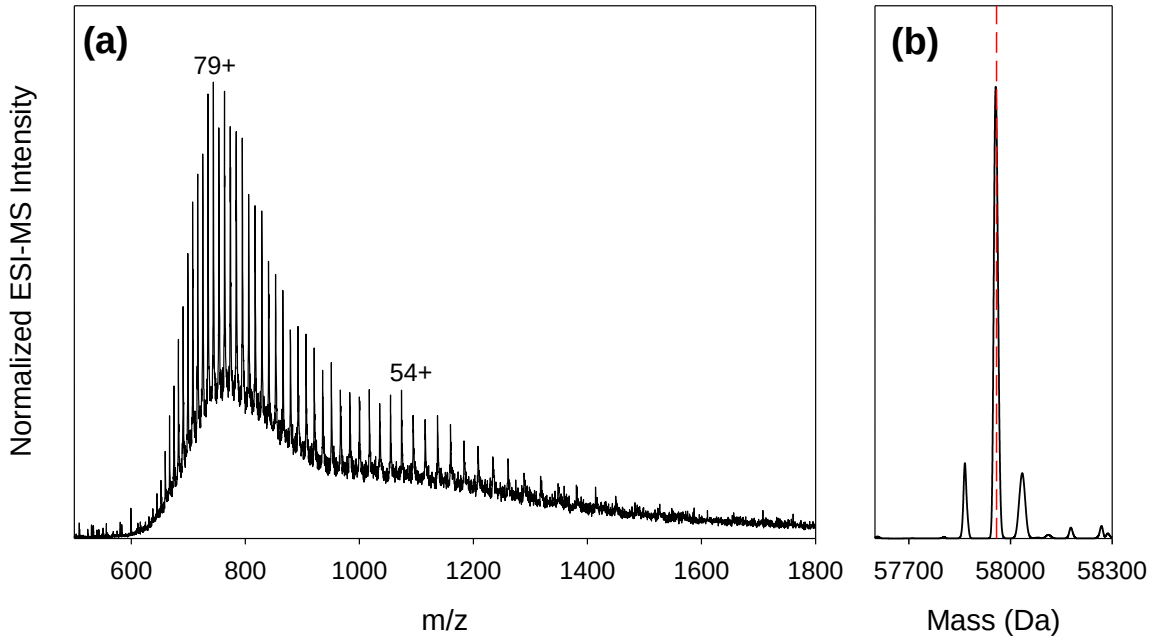

Figure 2.4. ESI mass spectrum of denatured 5 μM monomeric *r*M1-PK acquired after passing the protein through a C4 column using a water/acetonitrile gradient in the presence of 0.1% formic acid. (a) Complete mass spectrum. (b) Deconvoluted mass distribution, resulting in a measured protein mass of 57956 ± 3 Da. The red dashed line indicates the calculated mass of the protein (57959 Da) after loss of Met1 and N-terminal acetylation.[3]

2.2.3 Solution Conditions

Unless noted otherwise, all experiments were performed in 50 mM Tris buffer, 100 mM KCl, and 8 mM $MgCl_2$ at pH 7.5 with *r*M1-PK concentration of 0.125 μM (as tetramer). This salt environment has previously been shown to be adequate for ensuring *r*M1-PK activity.[77] Different types of samples were tested. (i) In the absence of any other solutes, the aforementioned solution conditions provide the *r*M1-PK resting state. (ii) Working state experiments were conducted in the presence of 10 mM ADP and 7.5 mM PEP. Both of these concentrations are well above the K_M values of the two substrates, K_M(PEP) = 0.07 mM, K_M(ADP) = 0.3 mM.[78] The time period during which the enzyme maintains turnover can be estimated as (limiting substrate concentration)/(k_{cat} × active site

concentration) = 7500 μM / (274 s^{-1} × 0.5 μM) ≈ 1 min. Keeping the enzyme in the working state for a longer time interval would require higher substrate concentrations. Unfortunately, this was not possible due to ADP-mediated suppression of peptide signals after peptic digestion. Some signal suppression was already apparent at an ADP concentration of 10 mM, but the effect was still tolerable. An alternative strategy to extend active state conditions would be to lower enzyme concentrations, but this would result in undesirable loss of S/N ratio during MS analysis. (iii) Substrate binding experiments were performed in the presence of either 2 mM ADP or PEP. (iv) "Product" binding experiments were performed with 5 mM ATP and/or 1 mM potassium oxalate. These ligand concentrations mirror those used in crystallization experiments of the corresponding bound states.[22] (v) Metal-depleted *r*M1-PK (produced by dialysis in the presence of EDTA, see above) was studied in 50 mM Tris buffer; for K^+ and Mg^{2+} binding experiments the solutions were supplemented with 100 mM KCl or 8 mM $MgCl_2$, respectively.

2.2.4 HDX-MS

Deuteration was carried out in 90% D_2O at pD 7.5 (pD = pH meter reading + 0.4).[79] Aliquots were removed at various time points (0.1, 1, 10, 30, and 100 minutes) after initiation of deuteration. The aliquots were quenched by lowering the pH to 2.4 using 10% (v/v) formic acid (FA) and flash freezing in liquid N_2. The samples were thawed and injected into a nanoACQUITY UPLC equipped with HDX technology (Waters, Milford, MA). Digestion was carried out online using a POROS pepsin column (2.1 mm × 30 mm) from Life Technologies/Applied Biosystems (Carlsbad, CA) held at 15 ºC. Subsequently, the peptides were trapped on a C_{18} BEH130 VanGuard column (5 mm × 1 mm, 1.7 μm) for

two minutes. Peptides were separated on a C_8 BEH130 column (50 mm × 2.1 mm, 1.7 μm) at 90 μL min^{-1} using a water/acetonitrile gradient with 0.1 % FA. The LC outflow was directed to a Waters Synapt Q-TOF G2 ESI mass spectrometer. The instrument was operated using a capillary voltage of +2.8 kV and a cone voltage of 20 V. The desolvation and source temperatures were set to 250 ºC and 80 ºC, respectively. Mass spectra were acquired in resolution mode. Traveling-wave ion mobility separation was employed to separate overlapping peptides.[68] Peptides were identified using MS^E acquisition in the absence of D_2O. MS^E data were analyzed using ProteinLynx Global Server 2.5.3. All software-based peptide assignments were verified manually. Online digestion yielded 108 peptides with a sequence coverage of 74.4% for the working state. All other states yielded 125 peptides, corresponding to a sequence coverage of 80.4%. As noted earlier, this difference results from ADP-induced peptide signal suppression in the working state.

2.2.5 Data Analysis

DynamX 3.0 (Waters) was used for HDX data analysis. Deuterium levels were corrected for artificial in-exchange and back-exchange using unlabeled (m_0) and full deuterated (m_{100}) control samples, respectively. The latter were produced by incubation of acid-denatured *r*M1-PK at pD 2.4 in 90% D_2O for 15 h at 37 ºC. Deuterium uptake for each peptide at labeling time *t* was calculated as:

$$\% D(t) = \frac{m_t - m_0}{m_{100} - m_0} \times 100\%$$

Residue numbering refers to the Uniprot sequence PKM1/P11974-1, which has the N-terminal methionine as residue 1 (MSKSH…). This is in contrast to some other studies[22]

where the subsequent Ser2 was considered to be the first residue (SKSH…). Error bars represent standard deviations derived from three independent sets of experiments for each solution condition.

2.3 Results and Discussion

2.3.1 Conformational Dynamics During Catalysis

HDX-MS measurements on *r*M1-PK were carried out under various conditions. In all cases the K^+/Mg^{2+}- saturated resting state served as reference. Data for the working enzyme was obtained by monitoring deuterium incorporation into *r*M1-PK while the PEP + Mg-ADP → pyruvate + Mg-ATP conversion proceeded in solution. As noted in the Methods section, these turnover conditions could be maintained for one minute. HDX-MS yielded information on the deuteration behaviour of peptides throughout the protein, but particular attention was focused on three segments located in the active site (Figure 2.5a). Segment 110-123 binds K^+ and is in contact with the phosphate groups of Mg-ADP or Mg-ATP. Also, this region forms part of the hinge that allows for the movement of domain B relative to A during turnover. Segment 292-298 accommodates the non-nucleotide Mg^{2+}, and it interacts with PEP and pyruvate (or oxalate in the crystal structure).[22, 55] Segment 365-374 interacts with the adenine moiety of Mg-ADP or Mg-ATP.

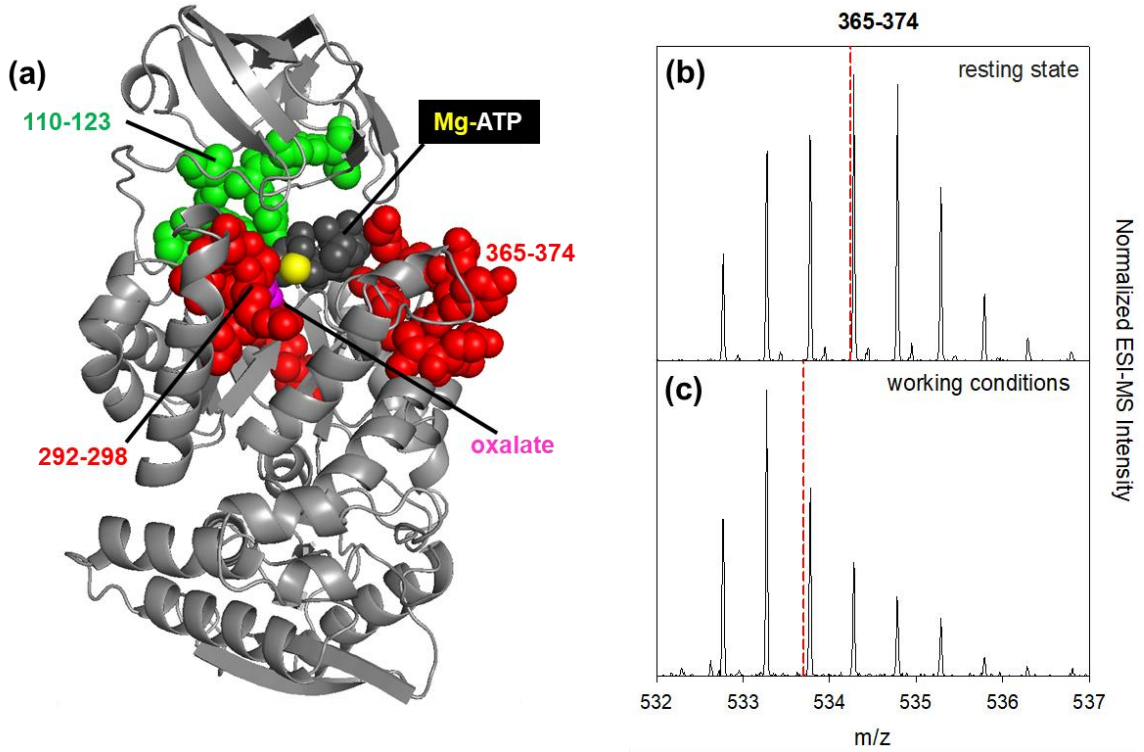

Figure 2.5. (a) Locations of peptide 110-123, 292-298, and 365-374 in the *r*M1-PK active site.[22] Unprocessed mass spectra of segment 365-374 after 1 min of HDX are shown for the resting state (b), and under working conditions with substrate turnover (c). Vertical red dashed lines in panels b and c represent isotope distribution centroids.

Mass spectra for segment 365-374 in the resting state and under working conditions for t = 1 min are shown in Figure 2.5b and c. Even without detailed analysis, it is apparent that working conditions caused significantly reduced deuteration compared to the resting state. After normalization (see Methods), it was found that the deuteration difference for segment 365-374 at t = 1 min was 36%. Even for measurements taken after 6 s this segment already showed a deuteration difference of 18% (Figure 2.6c). Reduced HDX levels under working conditions were also observed for the active site peptide 110-123 and 292-297, although the differences were not as dramatic as for 365-374 (Figure 2.6a, b).

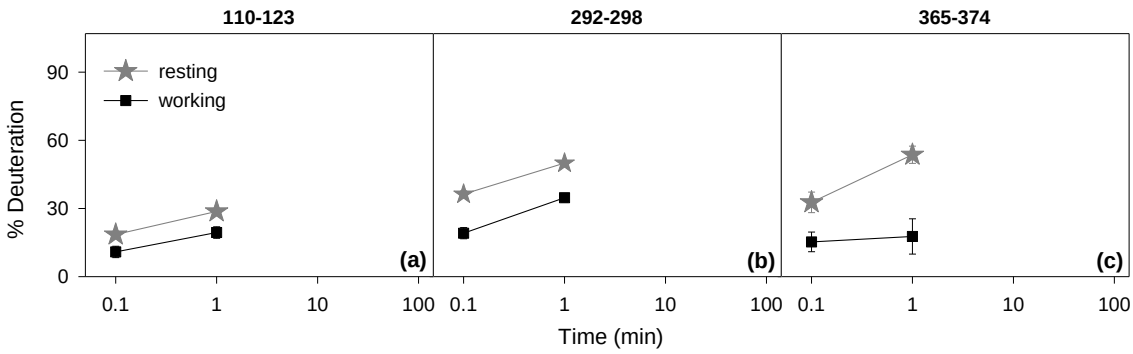

Figure 2.6. HDX kinetic plots for selected *r*M1-PK peptides found in the active site (110-123, 292-298, and 365-374, as noted along the top) under working conditions. (a-c) Comparison of resting state (gray) and working (black) state.

For visualizing the HDX response across the entire protein the measured deuteration differences were mapped to the crystal structure (Figure 2.7). All colour maps in the following sections use a scale where blue indicates regions that are less dynamic (i.e. with less deuteration) than the resting state and segments shown in red are more dynamic, as indicated in the legend of Figure 2.7. Figure 2.7 reinforces the finding that peptides in the vicinity of the active site exhibited strongly reduced deuteration under working conditions. Slightly elevated HDX levels (~10%) were seen in some domain B segments after 1 min, while the N and C domains were not affected.

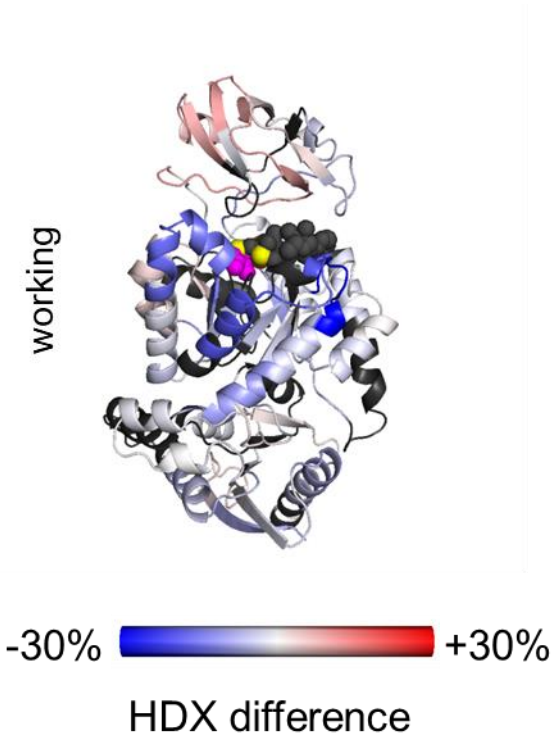

Figure 2.7. HDX difference plot for the working state at t = 1 min relative to the resting state, i.e., [%D(current state) - %D(resting state)], mapped to the crystal structure of one (closed) *r*M1-PK subunit. The ligands shown represent the available crystal structures[22] and do not always correspond to the ligands present in the working state experiment.

The reduced deuteration in the *r*M1-PK active site during catalysis is unexpected. Studies on other enzymes under turnover conditions found that structural dynamics were enhanced,[13] altered,[28] or unchanged.[29-30] Turnover-induced rigidification of the H-bonding network, as observed here for *r*M1-PK, has not been previously reported. From the measured k_{cat} it can be estimated that each subunit undergoes 274 s^{-1} × 60 s ≈ 16,000 turnover events during the experimental time window, which must be accompanied by at least as many active site opening/closing transitions (Figure 2.1c, e). One might have expected that these large-scale movements contribute to enhanced dynamics in the active site or in the domain B/A hinge region during turnover (Figure 2.1c-e). However, this expectation is not confirmed by the experimental data. Several earlier studies indicated that

reduced conformational dynamics generally suppress catalytic efficiency.[16-17, 80] The rigidification seen here for catalytically active *r*M1-PK thus seems counterintuitive, considering the widely-accepted paradigm that conformational fluctuations are a prerequisite for enzyme function.[3, 10-18] Due to this observed behaviour, a thorough examination of the response of the protein to the presence or absence of a range of mechanistically relevant ligands, is described in the following sections.

2.3.2 Substrate Binding

Catalytic turnover by *r*M1-PK requires the presence of both PEP and Mg-ADP (Figure 2.1b). By adding just one of these compounds it is possible to examine changes in enzyme dynamics in response to substrate *binding*, as opposed to turnover. HDX-MS profiles measured upon addition of PEP (green) showed markedly reduced deuteration in a number of peptides, particularly 292-298 and 365-374 in the active site (Figure 2.8b, c).

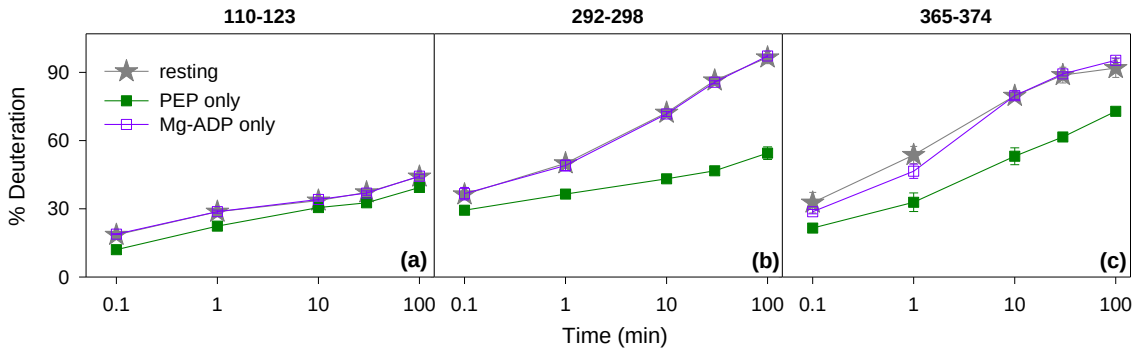

Figure 2.8. HDX kinetic plots for selected *r*M1-PK peptides found in the active site (110-123, 292-298, and 365, as noted along the top) under substrate binding conditions. (a-c) Comparison of the resting state (gray), PEP only (green) and Mg-ADP only (purple) states.

The magnitude and spatial distribution of the PEP-induced changes resembled the effects seen under turnover conditions (Figure 2.9a). The addition of Mg-ADP (without PEP) only

caused relatively minor changes in the deuteration behaviour relative to the resting state displayed in Figure 2.9b.

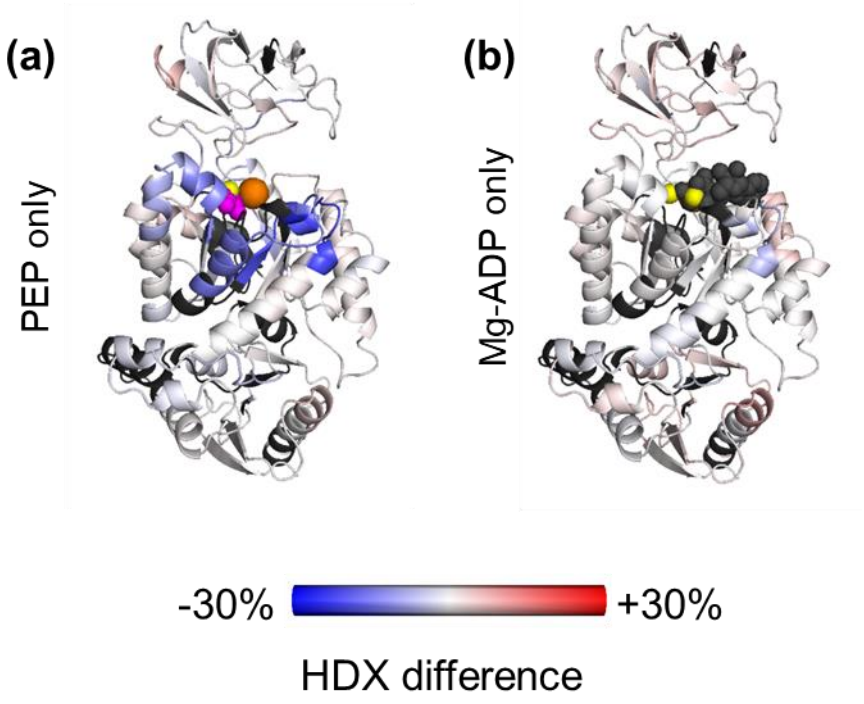

Figure 2.9. HDX difference plots for substrate binding at t = 1 min relative to the resting state, mapped to the crystal structures of one (semi-closed) rM1-PK subunit. (a) PEP binding, and (b) Mg-ADP binding. The ligands shown represent the available crystal structures[22] and do not always exactly correspond to the ligands present in the substrate binding experiments.

2.3.3 "Product" Binding

Next, investigations into characterizing the response of *r*M1-PK to reaction products in the solution was performed. The canonical products are pyruvate and Mg-ATP. However, the binding affinity of *r*M1-PK to pyruvate is low such that unreasonably high concentrations (>> 10 mM) would be required to achieve significant levels of bound complex.[22] The tendency of pyruvate to dimerize under such conditions would cause additional complications.[81] Following earlier crystallographic studies on *r*M1-PK,[22] the protein was exposed to oxalate, which has a much higher binding affinity than pyruvate.

Oxalate mimics the enolate that is transiently formed in the active site as an obligatory intermediate during PEP → pyruvate conversion (Figure 2.1b)

Shown in Figure 2.10, exposure to Mg-ATP (purple) caused only minor changes in the deuteration behaviour of *r*M1-PK relative to the resting state. In contrast, the presence of oxalate (green) significantly lowered deuteration levels in the active site. Even more pronounced effects were seen when both species were present simultaneously, demonstrating that oxalate and Mg-ATP stabilize the active site H-bond network in a synergistic fashion (Figure 2.10 a and c, blue).

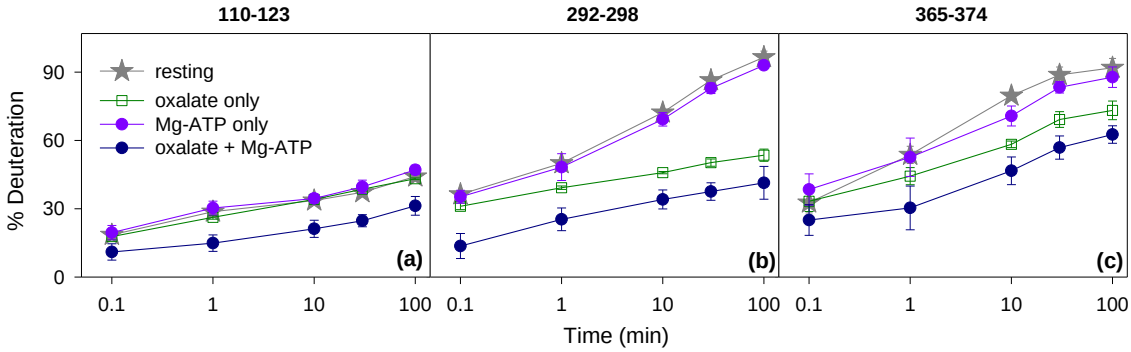

Figure 2.10. HDX kinetic plots for selected *r*M1-PK peptides found in the active site (110-123, 292-298, and 365, as noted along the top) under "product" binding conditions. (a-c) Comparison of the resting state (gray), oxalate only (green), Mg-ATP only (purple), and oxalate + Mg-ATP (blue).

The HDX difference maps generated for the oxalate and oxalate/Mg-ATP-bound protein (Figure 2.11a, c) closely resemble that of the working enzyme (Figure 2.7). In all three cases, the blue-coloured regions in domain A reflect a rigidification of the H-bond network in the vicinity of the active site.

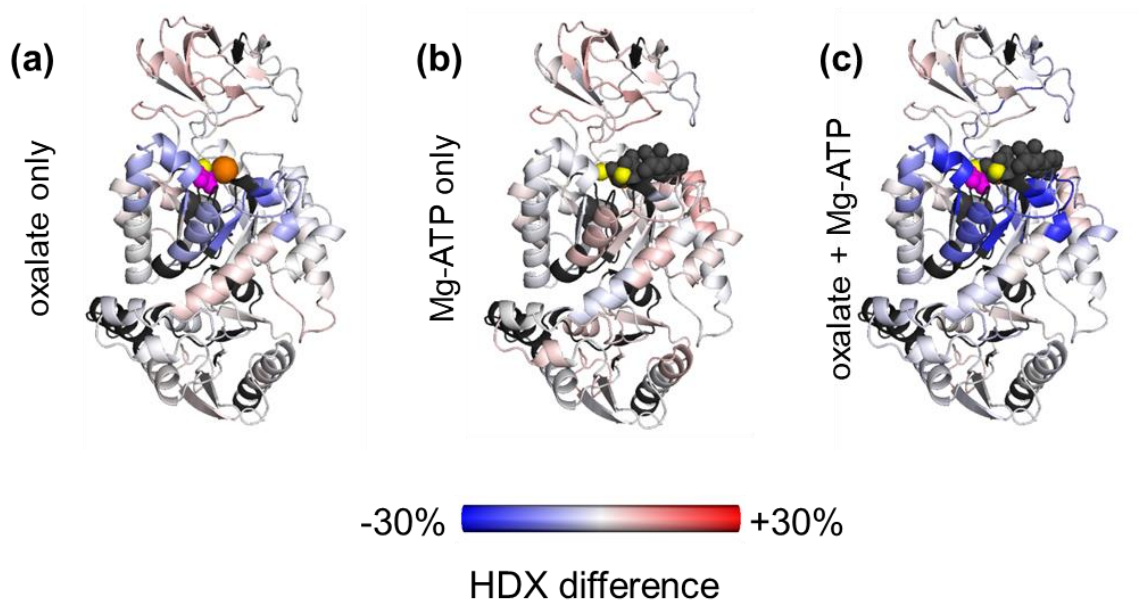

Figure 2.11. HDX difference plots for product binding at t = 1 min relative to the resting state, mapped to the crystal structures of one (semi-closed) *r*M1-PK subunit. (a) Oxalate binding; (b) Mg-ATP binding; and (c) Oxalate + Mg-ATP. The ligands shown represent the available crystal structures[22] and do not always exactly correspond to the ligands present in the product binding experiments.

2.3.4 Effects of Metal Cofactors

To further characterize the response of *r*M1-PK to intermolecular contacts the consequence of K^+ and Mg^{2+} depletion, being aware of the fact that such conditions are not necessarily relevant in a physiological context, was examined. K^+/Mg^{2+} depletion induced dramatically increase deuteration in the active site where the metal binding sites are located (Figure 2.12a-c). Other protein regions were affected to a much less extent (Figure 2.13a).

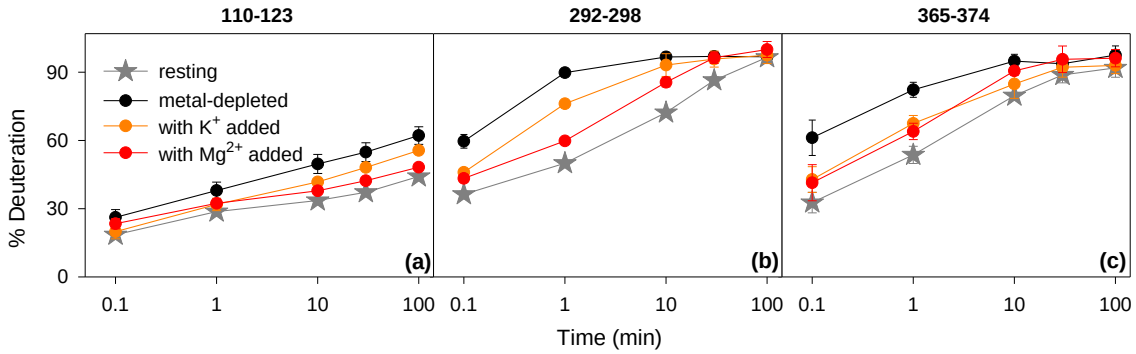

Figure 2.12. HDX kinetic plots for selected *r*M1-PK peptides found in the active site (110-123, 292-298, and 365, as noted along the top) under metal binding conditions. (a-c) Comparison of the resting state (gray), metal-depleted rM1-PK (black), K^+ only (orange), and Mg^{2+} only (red).

Upon supplementing the metal-depleted protein with K^+ or Mg^{2+} it was found that the latter was more effective in terms of restoring stability of the H-bonding network (Figure 2.13b, c). This behaviour is consistent with the fact that divalent ions usually interact more strongly with proteins than monovalent species.[5] In *r*M1-PK the non-nucleotide Mg^{2+} is

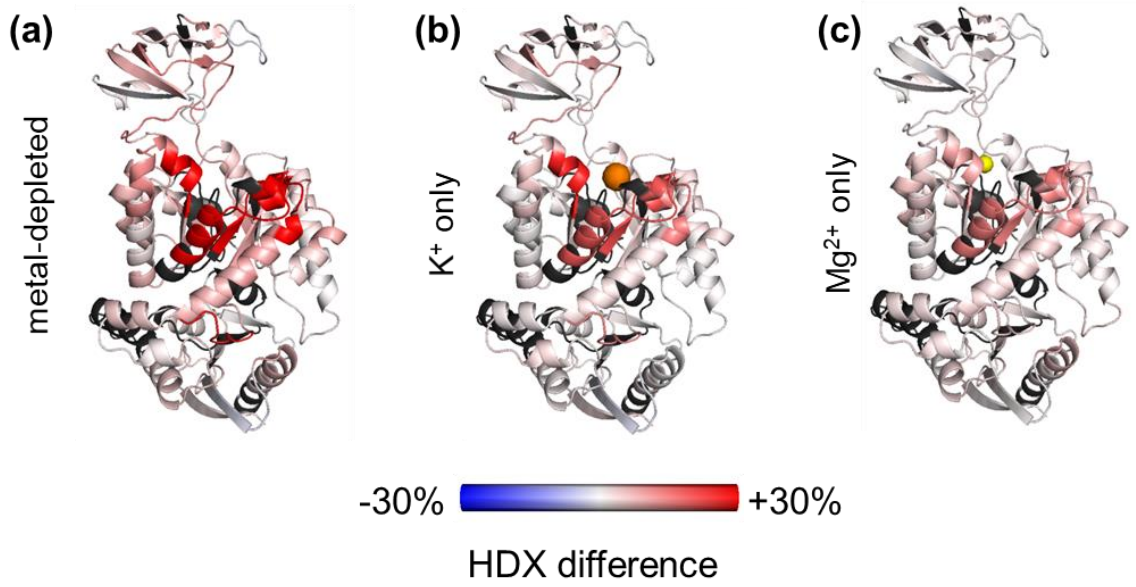

Figure 2.13. HDX difference plots for metal binding at t = 1 min relative to the resting state, mapped to the crystal structures of one (open) rM1-PK subunit. (a) Metal-depleted rM1-PK; (b) K^+ binding; and (c) Mg^{2+} binding. The ligands shown represent the available crystal structures[22] and do not always exactly correspond to the ligands present in the metal binding experiments.

tightly chelated by Glu273 and Asp296, while K^+ binds Asp113 and interacts with two side chains and one main chain oxygen.[22]

2.3.5 Ligand-Induced Active Site Stabilization vs. Catalytically Relevant Motions

The HDX-MS binding experiments described above reveal the H-bond fluctuations of *r*M1-PK are strongly affected by the presence of ligands. Binding to PEP, oxalate, or oxalate/Mg-ATP cause significant stabilization of the active site. This ligand-induced stabilization suggests that the active site follows an induced-fit scenario, where pliable elements fold around the ligand. Such behaviour has previously been documented for many other protein-ligand interactions.[82-83] In all these cases, intermolecular contacts trigger conformational changes that stabilize the protein scaffold surrounding the ligand recognition site.[32-45] The high sensitivity of the active site for ligand-induced stabilization is further emphasized by the finding that Mg^{2+} removal dramatically increase the extent of deuteration, implying that metal binding is subject to induced behaviour as well.

Close inspection of Figure 2.9 and Figure 2.11 colour maps reveal red hues in domain B, indicating that his region which is somewhat remote from the active site, shows a slight destabilization in the presence of catalytically relevant ligands. This behaviour is consistent with earlier reports that ligand binding can cause stabilization as well as destabilization in different protein regions.[46-49]

Strikingly, the HDX patterns observed upon binding of PEP, oxalate, or oxalate/Mg-ATP under conditions that preclude substrate turnover (Figure 2.9a and Figure 2.11a, c) are very similar to the HDX pattern observed under working conditions (Figure 2.7). From this finding it is concluded that changes in the *r*M1-PK H-bonding network seen

under working conditions primarily reflect the protein response to *binding* interactions, rather than the occurrence of catalytically relevant enzyme motions. In other words, H-bond stabilization during catalysis arises from the fact that the protein cycles through various ligand-bound states under working conditions (E·PEP·Mg-ADP, E·enolate·Mg-ATP, E·enolate, see Figure 2.1b) These observations imply that interactions with PEP and enolate are primarily responsible for the rigidification during catalysis, while the presence of Mg-ADP and Mg-ATP has less pronounced effects on H-bond stability.

Very minor differences are apparent between the working state HDX map (Figure 2.7) and the data obtained for PEP, oxalate, and oxalate/Mg-ATP-bound states (Figure 2.9a and Figure 2.11a, c). However, attempts to interpret these small effects as a signature of unique catalytic motions would be unjustified, keeping in mind that no ligand-bound state will ever represent a perfect comparator for the working enzyme.

It is not disputed that conformational dynamics play a role for the *r*M1-PK catalytic mechanism, as implied by the closed/open crystal structure of Figure 1 c and e.[22] In addition to these large domain movements, small-scale fluctuations may facilitate the chemical conversion steps taking place during catalysis.[17] However, the HDX-MS data fail to uncover the presence of any conformational events that are uniquely linked to catalysis. Rather, the deuteration kinetics are dominated by ligand-induced stabilization that is encountered regardless whether turnover takes place or not. Any efforts to specifically identify catalytically relevant motions within the wide spectrum of *r*M1-PK dynamics[8-9] are hindered by the masking contributions of rather mundane ligand binding effects. Clearly, binding-induced stabilization is associated with *r*M1-PK catalysis, but this

stabilization does not represent a *unique* feature of the catalytically active protein because the same stabilization effects can be induced in the absence of turnover.

2.4 Conclusions

The question addressed in this work is whether working/resting state comparative HDX-MS measurements can provide direct information on the nature of catalytically relevant enzyme motions. This concept is analogous to trying to decipher the inner workings of an industrial assembly line by examining a factory on a Sunday when the machinery is on standby, followed by a comparative inspection during the week when raw material is converted to product. Within this simple analogy, it will undoubtedly be possible to uncover how motions of individual machine parts contribute to the overall turnover process.

In the case of enzymes, the situation is more complex. Incessant thermal motions take place on a wide range of length and time scales, as proteins cycle through their Boltzmann-allowed conformational space.[8-9] Some of these fluctuations may be catalytically relevant, while others represent unproductive "noise". Catalytically relevant motions may be facilitated by the presence of substrate in the active site,[13, 28] while others are intrinsic.[29-30] During enzymatic turnover there will be a certain fraction of time during which the active site is bound to substrates, intermediates or products; these bound conditions will alternate with periods where the active site is vacant and primed for the binding of new substrates that would initiate the next turnover cycle. When aiming to design comparative measurements for identifying motions that are uniquely linked to catalysis, it may be impossible to select a suitable state against which the working enzyme

can be appraised. The resting state is not appropriate because it lacks pertinent ligand-protein interactions. On the other extreme, binding of catalytically relevant species to the active site without turnover (as in Figure 2.9 and Figure 2.11) results in ligand-protein interactions that are much more permanent than those encountered during turnover. Neither condition provides a proper comparison base, which is why HDX difference maps will generally *not* reveal dynamic features that are uniquely linked to catalysis. Other enzymes may be less prone to binding-induced alterations than *r*M1-PK,[29-30] but this does not imply that differences seen in working/resting state comparisons for those other systems will necessarily pinpoint catalytically relevant motions.

Overall, the current work demonstrates that efforts to elucidate linkages between enzyme dynamics and catalysis via comparative measurements should be treated with caution, although there is no shortage of such experiments in the literature. This conclusion is not limited to HDX/MS but also applies to other bioanalytical tools. In the case of *r*M1-PK, dramatic changes were observed in the extent of conformational dynamics during catalysis. These changes are not unique to catalysis, but they arise from binding interactions that can also be implemented in the absence of turnover (compared in Figure 2.14). Comparable binding-related stabilization[32-45] (or destabilization)[46-49] effects have been reported for non-enzymatic proteins. Based on current available data it appears that the magnitude of such binding-associated dynamic changes dwarfs subtle conformational events that might be uniquely linked to enzymatic turnover.

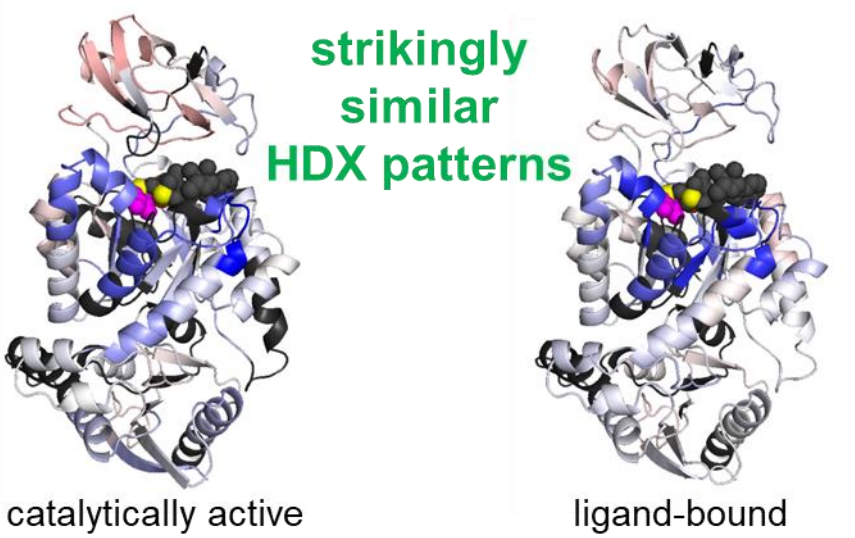

Figure 2.14. Dynamic changes observed during catalysis are quite similar to dynamic changes observed upon ligand binding.

2.5 References

1. Wolfenden, R.; Snider, M. J., The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res.* **2001,** *34* (12), 938-945.

2. Cleland, W. W.; Northrop, D. B., Energetics of substrate binding, catalysis, and product release. *Methods Enzymol.* **1999,** *308*, 3-27.

3. Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V.; Alber, T.; Hagan, M. F.; Kern, D., The energy landscape of adenylate kinase during catalysis. *Nat Struct Mol Biol* **2015,** *22* (2), 124-131.

4. Vocadlo, D. J.; Davies, G. J.; Laine, R.; Withers, S. G., Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **2001,** *412* (6849), 835-838.

5. Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayyum, H.; Pecoraro, V. L., Protein Design: Toward Functional Metalloenzymes. *Chem. Rev.* **2014,** *114* (7), 3495-3578.

6. Valdez, C. E.; Smith, Q. A.; Nechay, M. R.; Alexandrova, A. N., Mysteries of Metals in Metalloenzymes. *Acc. Chem. Res.* **2014,** *47* (10), 3110-3117.

7. Dielmann-Gessner, J.; Grossman, M.; Conti Nibali, V.; Born, B.; Solomonov, I.; Fields, G. B.; Havenith, M.; Sagi, I., Enzymatic turnover of macromolecules generates long-lasting protein–water-coupled motions beyond reaction steady state. *Pro. Natl. Acad. Sci. U. S. A.* **2014,** *111* (50), 17857-17862.

8. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D., A unified model of protein dynamics. *Pro. Natl. Acad. Sci. U. S. A.* **2009,** *106* (13), 5129-5134.

9. Xu, Y.; Mayne, L.; Englander, S. W., Evidence for an unfolding and refolding pathway in cytochrome c. *Nat. Struct. Mol. Biol.* **1998,** *5* (9), 774-778.

10. Hammes-Schiffer, S., Catalytic Efficiency of Enzymes: A Theoretical Analysis. *Biochemistry* **2013,** *52* (12), 2012-2020.

11. Boehr, D. D.; Dyson, H. J.; Wright, P. E., An NMR Perspective on Enzyme Dynamics. *Chem. Rev.* **2006,** *106* (8), 3055-3079.

12. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-972.

13. Liuni, P.; Jeganathan, A.; Wilson, D. J., Conformer Selection and Intensified Dynamics During Catalytic Turnover in Chymotrypsin. *Angew. Chem. Int. Ed.* **2012,** *51* (38), 9666-9669.

14. Campbell, E.; Kaltenbach, M.; Correy, G. J.; Carr, P. D.; Porebski, B. T.; Livingstone, E. K.; Afriat-Jurnou, L.; Buckle, A. M.; Weik, M.; Hollfelder, F.; Tokuriki, N.; Jackson, C. J., The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol* **2016,** *12* (11), 944-950.

15. Kohen, A.; Klinman, J. P., Protein Flexibility Correlates with Degree of Hydrogen Tunneling in Thermophilic and Mesophilic Alcohol Dehydrogenases. *J. Am. Chem. Soc.* **2000,** *122* (43), 10738-10739.

16. Liang, Z.-X.; Lee, T.; Resing, K. A.; Ahn, N. G.; Klinman, J. P., Thermal-activated protein mobility and its correlation with catalysis in thermophilic alcohol dehydrogenase. *Pro. Natl. Acad. Sci. U. S. A.* **2004,** *101* (26), 9556-9561.

17. Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E., A Dynamic Knockout Reveals That Conformational Fluctuations Influence the Chemical Step of Enzyme Catalysis. *Science* **2011,** *332* (6026), 234-238.

18. Kovermann, M.; Ådén, J.; Grundström, C.; Elisabeth Sauer-Eriksson, A.; Sauer, U. H.; Wolf-Watz, M., Structural basis for catalytically restrictive dynamics of a high-energy enzyme state. *Nat. Commun.* **2015,** *6*, 7644.

19. Holland, D. R.; Tronrud, D. E.; Pley, H. W.; Flaherty, K. M.; Stark, W.; Jansonius, J. N.; McKay, D. B.; Matthews, B. W., Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis. *Biochemistry* **1992,** *31* (46), 11310-11316.

20. Müller, C. W.; Schlauderer, G. J.; Reinstein, J.; Schulz, G. E., Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure* **1996,** *4* (2), 147-156.

21. Bystroff, C.; Kraut, J., Crystal structure of unliganded Escherichia coli dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding. *Biochemistry* **1991,** *30* (8), 2227-2239.

22. Larsen, T. M.; Benning, M. M.; Rayment, I.; Reed, G. H., Structure of the Bis(Mg2+)−ATP−Oxalate Complex of the Rabbit Muscle Pyruvate Kinase at 2.1 Å Resolution: ATP Binding over a Barrel. *Biochemistry* **1998,** *37* (18), 6247-6255.

23. Joseph, D.; Petsko, G.; Karplus, M., Anatomy of a conformational change: hinged "lid" motion of the triosephosphate isomerase loop. *Science* **1990,** *249* (4975), 1425-1428.

24. Gutteridge, A.; Thornton, J., Conformational change in substrate binding, catalysis and product release: an open and shut case? *FEBS Lett.* **2004,** *567* (1), 67-73.

25. Kamerlin, S. C.; Warshel, A., Reply to Karplus: Conformational dynamics have no role in the chemical step. *Pro. Natl. Acad. Sci. U. S. A.* **2010,** *107* (17), E72-E72.

26. Karplus, M., Role of conformation transitions in adenylate kinase. *Pro. Natl. Acad. Sci. U. S. A.* **2010,** *107* (17), E71.

27. Pisliakov, A. V.; Cao, J.; Kamerlin, S. C. L.; Warshel, A., Enzyme millisecond conformational dynamics do not catalyze the chemical step. *Pro. Natl. Acad. Sci. U. S. A.* **2009,** *106* (41), 17359-17364.

28. Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E., The Dynamic Energy Landscape of Dihydrofolate Reductase Catalysis. *Science* **2006,** *313* (5793), 1638-1642.

29. Liu, Y.-H.; Konermann, L., Conformational Dynamics of Free and Catalytically Active Thermolysin Are Indistinguishable by Hydrogen/Deuterium Exchange Mass Spectrometry. *Biochemistry* **2008,** *47* (24), 6342-6351.

30. Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D., Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007,** *450* (7171), 838-844.

31. Verkhivker, G. M.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Freer, S. T.; Rose, P. W., Simulating disorder–order transitions in molecular recognition of unstructured proteins: Where folding meets binding. *Pro. Natl. Acad. Sci. U. S. A.* **2003,** *100* (9), 5148-5153.

32. Creighton, T. E., *Proteins: structures and molecular properties*. Macmillan: 1993.

33. Eliezer, D.; Yao, J.; Dyson, H. J.; Wright, P. E., Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Mol. Biol.* **1998,** *5* (2), 148-155.

34. Keppel, T. R.; Howard, B. A.; Weis, D. D., Mapping Unstructured Regions and Synergistic Folding in Intrinsically Disordered Proteins with Amide H/D Exchange Mass Spectrometry. *Biochemistry* **2011,** *50* (40), 8722-8732.

35. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* **2006,** *25* (1), 158-170.

36. Zhu, M. M.; Rempel, D. L.; Du, Z.; Gross, M. L., Quantification of Protein−Ligand Interactions by Mass Spectrometry, Titration, and H/D Exchange: PLIMSTEX. *J. Am. Chem. Soc.* **2003,** *125* (18), 5252-5253.

37. Trelle, M. B.; Hirschberg, D.; Jansson, A.; Ploug, M.; Roepstorff, P.; Andreasen, P. A.; Jørgensen, T. J. D., Hydrogen/Deuterium Exchange Mass Spectrometry Reveals Specific Changes in the Local Flexibility of Plasminogen Activator Inhibitor 1 upon Binding to the Somatomedin B Domain of Vitronectin. *Biochemistry* **2012,** *51* (41), 8256-8266.

38. Mayne, L.; Paterson, Y.; Cerasoli, D.; Englander, S. W., Effect of antibody binding on protein motions studied by hydrogen-exchange labeling and two-dimensional NMR. *Biochemistry* **1992,** *31* (44), 10678-10685.

39. Wang, C.; Pawley, N. H.; Nicholson, L. K., The role of backbone motions in ligand binding to the c-Src SH3 domain11Edited by P. E. Wright. *J. Mol. Biol.* **2001,** *313* (4), 873-887.

40. Williams, D. C.; Rule, G. S.; Poljak, R. J.; Benjamin, D. C., Reduction in the amide hydrogen exchange rates of an anti-lysozyme Fv fragment due to formation of the Fv-Lysozyme complex11Edited by P. E. Wright. *J. Mol. Biol.* **1997,** *270* (5), 751-762.

41. Massiah, M. A.; Saraswat, V.; Azurmendi, H. F.; Mildvan, A. S., Solution Structure and NH Exchange Studies of the MutT Pyrophosphohydrolase Complexed with Mg2+ and 8-Oxo-dGMP, a Tightly Bound Product. *Biochemistry* **2003,** *42* (34), 10140-10154.

42. Emerson, S. D.; Palermo, R.; Liu, C.-M.; Tilley, J. W.; Chen, L.; Danho, W.; Madison, V. S.; Greeley, D. N.; Ju, G.; Fry, D. C., NMR characterization of interleukin-2 in complexes with the IL-2Rα receptor component, and with low molecular weight compounds that inhibit the IL-2/IL-Rα interaction. *Protein Science : A Publication of the Protein Society* **2003,** *12* (4), 811-822.

43. Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R., Mass spectrometry-based methods to study protein architecture and dynamics. *Protein Sci.* **2013,** *22* (5), 530-544.

44. Powell, K. D.; Ghaemmaghami, S.; Wang, M. Z.; Ma, L.; Oas, T. G.; Fitzgerald, M. C., A General Mass Spectrometry-Based Assay for the Quantitation of Protein−Ligand Binding Interactions in Solution. *J. Am. Chem. Soc.* **2002,** *124* (35), 10256-10257.

45. Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C., Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry—A review. *Anal. Chim. Acta* **2012,** *721*, 7-21.

46. Bobst, C. E.; Zhang, M.; Kaltashov, I. A., Existence of a noncanonical state of iron-bound transferrin at endosomal pH revealed by hydrogen exchange and mass spectrometry. *J. Mol. Biol.* **2009,** *388* (5), 954-967.

47. Sowole, M. A.; Alexopoulos, J. A.; Cheng, Y.-Q.; Ortega, J.; Konermann, L., Activation of ClpP Protease by ADEP Antibiotics: Insights from Hydrogen Exchange Mass Spectrometry. *J. Mol. Biol.* **2013,** *425* (22), 4508-4519.

48. Burke, J. E.; Babakhani, A.; Gorfe, A. A.; Kokotos, G.; Li, S.; Woods Jr, V. L.; McCammon, J. A.; Dennis, E. A., Location of inhibitors bound to group IVA phospholipase A2 determined by molecular dynamics and deuterium exchange mass spectrometry. *J. Am. Chem. Soc.* **2009,** *131* (23), 8083-8091.

49. Shukla, A. K.; Westfield, G. H.; Xiao, K.; Reis, R. I.; Huang, L.-Y.; Tripathi-Shukla, P.; Qian, J.; Li, S.; Blanc, A.; Oleskie, A. N., Visualization of arrestin recruitment by a G protein-coupled receptor. *Nature* **2014,** *512* (7513), 218.

50. Imamura, K.; Tanaka, T., Multimolecular forms of pyruvate kinase from rat and other mammalian tissues. *The Journal of Biochemistry* **1972,** *71* (6), 1043-1051.

51. Liu, V. M.; Vander Heiden, M. G., The Role of Pyruvate Kinase M2 in Cancer Metabolism. *Brain Pathol.* **2015,** *25* (6), 781-783.

52. Christofk, H. R.; Vander Heiden, M. G.; Harris, M. H.; Ramanathan, A.; Gerszten, R. E.; Wei, R.; Fleming, M. D.; Schreiber, S. L.; Cantley, L. C., The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* **2008,** *452* (7184), 230-233.

53. Anastasiou, D.; Yu, Y.; Israelsen, W. J.; Jiang, J.-k.; Boxer, M. B.; Hong, B. S.; Tempel, W.; Dimov, S.; Shen, M.; Jha, A.; Yang, H.; Mattaini, K. R.; Metallo, C. M.; Fiske, B. P.; Courtney, K. D.; Malstrom, S.; Khan, T. M.; Kung, C.; Skoumbourdis, A. P.; Veith, H.; Southall, N.; Walsh, M. J.; Brimacombe, K. R.; Leister, W.; Lunt, S. Y.; Johnson, Z. R.; Yen, K. E.; Kunii, K.; Davidson, S. M.; Christofk, H. R.; Austin, C. P.; Inglese, J.; Harris, M. H.; Asara, J. M.; Stephanopoulos, G.; Salituro, F. G.; Jin, S.; Dang, L.; Auld, D. S.; Park, H.-W.; Cantley, L. C.; Thomas, C. J.; Vander Heiden, M. G., Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol* **2012,** *8* (10), 839-847.

54. Iqbal, M. A.; Gupta, V.; Gopinath, P.; Mazurek, S.; Bamezai, R. N. K., Pyruvate kinase M2 and cancer: an updated assessment. *FEBS Lett.* **2014,** *588* (16), 2685-2692.

55. Dombrauckas, J. D.; Santarsiero, B. D.; Mesecar, A. D., Structural Basis for Tumor Pyruvate Kinase M2 Allosteric Regulation and Catalysis. *Biochemistry* **2005,** *44* (27), 9417-9429.

56. Jurica, M. S.; Mesecar, A.; Heath, P. J.; Shi, W.; Nowak, T.; Stoddard, B. L., The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure* **1998,** *6* (2), 195-210.

57. Prasannan, C. B.; Villar, M. T.; Artigues, A.; Fenton, A. W., Identification of regions of rabbit muscle pyruvate kinase important for allosteric regulation by phenylalanine, detected by H/D exchange mass spectrometry. *Biochemistry* **2013,** *52* (11), 1998-2006.

58. Ou, Y.; Tao, W.; Zhang, Y.; Wu, G.; Yu, S., The conformational change of rabbit muscle pyruvate kinase induced by activating cations and its substrates. *Int. J. Biol. Macromol.* **2010,** *47* (2), 228-232.

59. Yu, S.; Lee, L. L. Y.; Ching Lee, J., Effects of metabolites on the structural dynamics of rabbit muscle pyruvate kinase. *Biophys. Chem.* **2003,** *103* (1), 1-11.

60. Muñoz, M. E.; Ponce, E., Pyruvate kinase: current status of regulatory and functional properties. *Comp. Biochem. Physiol. B Biochem, Mol. Biol.* **2003,** *135* (2), 197-218.

61. Cheng, X.; Friesen, R. H.; Lee, J. C., Effects of conserved residues on the regulation of rabbit muscle pyruvate kinase. *J. Biol. Chem.* **1996,** *271* (11), 6313-6321.

62. Consler, T. G.; Jennewein, M. J.; Cai, G. Z.; Lee, J. C., Energetics of allosteric regulation in muscle pyruvate kinase. *Biochemistry* **1992,** *31* (34), 7870-7878.

63. Oria-Hernández, J.; Cabrera, N.; Pérez-Montfort, R.; Ramírez-Silva, L., Pyruvate Kinase Revisited: The Activating Effect of K+. *J. Biol. Chem.* **2005,** *280* (45), 37924-37929.

64. Lodato, D. T.; Reed, G. H., Structure of the oxalate-ATP complex with pyruvate kinase: ATP as a bridging ligand for the two divalent cations. *Biochemistry* **1987,** *26* (8), 2243-2250.

65. Long, D.; Marshall, C. B.; Bouvignies, G.; Mazhab-Jafari, M. T.; Smith, M. J.; Ikura, M.; Kay, L. E., A comparative CEST NMR study of slow conformational dynamics of small GTPases complexed with GTP and GTP analogues. *Angew. Chem. Int. Ed.* **2013,** *125* (41), 10971-10974.

66. Rand, K. D.; Zehl, M.; Jørgensen, T. J., Measuring the hydrogen/deuterium exchange of proteins at high spatial resolution by mass spectrometry: overcoming gas-phase hydrogen/deuterium scrambling. *Acc. Chem. Res.* **2014,** *47* (10), 3018-3027.

67. Marciano, D. P.; Dharmarajan, V.; Griffin, P. R., HDX-MS guided drug discovery: small molecules and biopharmaceuticals. *Curr. Opin. Struct. Biol.* **2014,** *28*, 105-111.

68. Vahidi, S.; Bi, Y.; Dunn, S. D.; Konermann, L., Load-dependent destabilization of the γ-rotor shaft in FOF1 ATP synthase revealed by hydrogen/deuterium-exchange mass spectrometry. *Pro. Natl. Acad. Sci. U. S. A.* **2016,** *113* (9), 2412-2417.

69. Englander, S. W.; Mayne, L.; Krishna, M. M., Protein folding and misfolding: mechanism and principles. *Q. Rev. Biophys.* **2007,** *40* (4), 1-41.

70. Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W., Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **2012,** *21* (7), 996-1005.

71. Donovan, K. A.; Zhu, S. L.; Liuni, P.; Peng, F.; Kessans, S. A.; Wilson, D. J.; Dobson, R. C. J., Conformational Dynamics and Allostery in Pyruvate Kinase. *J. Biol. Chem.* **2016,** *291* (17), 9244-9256.

72. Prasannan, C. B.; Artigues, A.; Fenton, A. W., Monitoring allostery in D2O: a necessary control in studies using hydrogen/deuterium exchange to characterize allosteric regulation. *Anal. Bioanal. Chem.* **2011,** *401* (3), 1083-1086.

73. Bücher, T.; Pfleiderer, G., Pyruvate kinase from muscle. *Methods Enzymol.* **1955,** *1*, 435-440.

74. Kayne, F. J., In *The Enzymes*, Boyer, P. D., Ed. Academic Press: New York, 1973; Vol. 3, pp 353-382.

75. Belov, M. E.; Damoc, E.; Denisov, E.; Compton, P. D.; Horning, S.; Makarov, A. A.; Kelleher, N. L., From Protein Complexes to Subunit Backbone Fragments: A Multi-stage Approach to Native Mass Spectrometry. *Anal. Chem.* **2013,** *85* (23), 11163-11173.

76. Cottam, G. L.; Hollenberg, P. F.; Coon, M. J., Subunit structure of rabbit muscle pyruvate kinase. *J. Biol. Chem.* **1969,** *244* (6), 1481-1486.

77. Kayne, F., Thallium (I) activation of pyruvate kinase. *Arch. Biochem. Biophys.* **1971,** *143* (1), 232-239.

78. Bergmeyer, H. U.; Gawehn, K.; Grassel, M., In *Methods of Enzymatic Analysis*, 2 ed.; Bergmeyer, H. U., Ed. Academic Press: New York, 1974; Vol. 1, pp 509-511.

79. Glasoe, P. K.; Long, F., Use of glass electrodes to measure acidities in deuterium oxide1, 2. *J. Phys. Chem.* **1960,** *64* (1), 188-190.

80. Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D., Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struct. Mol. Biol.* **2004,** *11* (10), 945-949.

81. Gallo, A. A.; Sable, H. Z., Rate enhancement of pyruvate aldolization by divalent cations: A model for class II aldolases. *Biochim. Biophys. Acta* **1973,** *302* (2), 443-456.

82. Boehr, D. D.; Nussinov, R.; Wright, P. E., The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol* **2009,** *5* (11), 789-796.

83. Koshland, D., Application of a theory of enzyme specificity to protein synthesis. *Pro. Natl. Acad. Sci. U. S. A.* **1958,** *44* (2), 98-104.

Chapter 3 – Protein-Protein Interactions in the Antioxidant Pathway: Probing the Disordered Nature of Full-length Nrf2 Using Hydrogen/Deuterium Exchange-Mass Spectrometry

3.1 Introduction

Intrinsically disordered proteins (IDPs) have a high degree of dynamic flexibility and lack a well-defined three-dimensional conformation. IDPs can have isolated disordered domains but in many cases this disorder extends to the entire protein in which case it is called 'natively' unfolded.[1-2] Characteristic of IDPs is their low level of sequence complexity and that their sparsity of hydrophobic and bulky residues result in the absence of a tertiary structure.[3] Due to their distinctive sequences, it is relatively straightforward to develop computer-based algorithms to identify these types of proteins in biological databases.[4-5] These algorithms have shown that IDPs encompass roughly one third of all known proteins in eukaryotic organisms.[6]

Many proteins conform to the classical structure-function paradigm, where a distinct sequence gives rise to a well-defined three-dimensional structure that is closely associated with a biological function (e.g., in the case of enzymes, see Chapter 2). IDPs do not follow this classic structure-function paradigm, since they lack a well-defined three-dimensional structure. This makes it difficult to understand their structural and dynamic characteristics as it is often the disordered regions that carry out their specific function by mediating protein-protein interactions.[7-9] The disordered regions involved in these interactions usually undergo a disorder/order transition, .i.e., they adopt their biologically

relevant conformation upon target binding.[10-11] IDPs interact with their targets with relatively high specificity and low affinity.[12] This process allows them to associate rapidly with their binding partners and a initiate a process, but they can also dissociate easily when this process is completed.[13]

IDPs play a crucial role in transcriptional regulation, translation and cellular signalling resulting in the activation/deactivation of certain genes.[14] One important example of this is cellular protection under endogenous or xenobiotic stressors such as reactive oxygen species (ROS), heavy metals, or other electrophilic stressors.[15] When large quantities of these cytotoxic species are present in cells, a mechanism is implemented to eliminate them to avoid cell damage which could lead to impaired cellular function. Specific proteins are involved in activating certain genes that encode cytoprotective enzymes. One protein that plays a key role in the activation of cytoprotective genes is the nuclear factor erythroid 2- related factor 2 (Nrf2).

Nrf2 is a 68 kDa protein that plays an important role in resistance to oxidative and electrophilic stress. It is a member of the cap ‘n’ collar subfamily of basic leucine zipper (bZIP) transcription factors. Research has shown Nrf2 plays a crucial role in cancer, neurodegenerative diseases, and many autoimmune and inflammatory disorders.[16-19] Distinctive to bZIP transcription factor proteins, Nrf2 consists of seven conserved domains displayed in Figure 3.1. Neh1 contains the bZIP DNA binding and heterodimerization domain through which Nrf2 interacts with small musculoaponeurotic fibrosarcoma (Maf) proteins in the nucleus. Neh2 contains a low affinity binding site (sequence DLG) and high affinity (ETGE) binding site that interact with Kelch-like ECH-associated protein 1 (Keap1). Each binding site on Nrf2 interacts with a separate Keap1 protein resulting in a

1:2 binding stoichiometry. Neh3-5 are involved in binding other transcription factors in the nucleus. Neh6 contains two redox-independent motifs that are targeted for degradation of Nrf2.[20] Lastly, Neh7 is involved in the interaction with the DNA-binding domain of retinoic X receptor α that has been believed to supress Nrf2.[21-23]

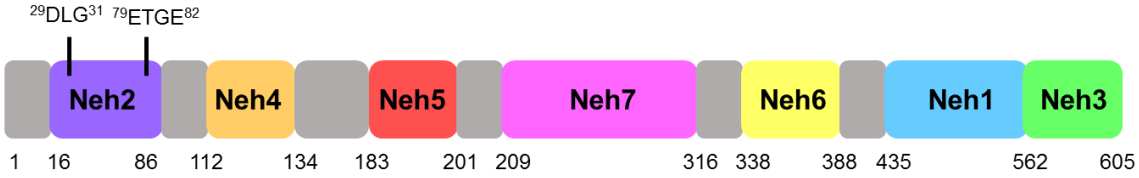

Figure 3.1. The position and residue numbering of Nrf2's seven conserved domains. Neh2 contains a low (DLG, residues 29-31) and a high (ETGE, residues 79-82) affinity binding motif that interact with two separate Keap1 proteins.

Keap1 is a 70 kDa protein that belongs to the BTB-Kelch family. Keap1 consists of four domains; the N-terminal Broad complex, Tramtrack, and Bric-a-Brac (BTB) domain, central intervening region (IVR), Kelch domain, and the C-terminal region shown in Figure 3.2. The N-terminal BTB domain is involved in homodimerization of Keap1. The central intervening domain is rich in cysteine, making this region redox sensitive. Lastly, the C-terminal Kelch domain is involved in interacting with Nrf2 through the ETGE and DLG motifs.[24-25]

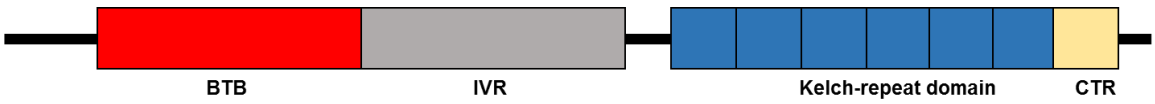

Figure 3.2. Keap1 domains. BTB domain (red), IVR domain (gray), Kelch domain (blue), and the C-terminal region (yellow). Two separate Kelch domains are involved in the interaction with Nrf2's low and high affinity binding motifs.

Under normal conditions (Figure 3.3a), Nrf2's DLG and ETGE motifs interact with Keap1 in a 1:2 ratio of Nrf2:Keap1 in the cytoplasm. When Nrf2-Keap1 interact with each other, this results in the polyubiquitination of Nrf2. Due to polyubiquitination of the Nrf2-Keap1 complex, Nrf2 is rapidly degraded by the 26S proteasome, with a short half-life of approximately 20 minutes in the cytoplasm.[26]

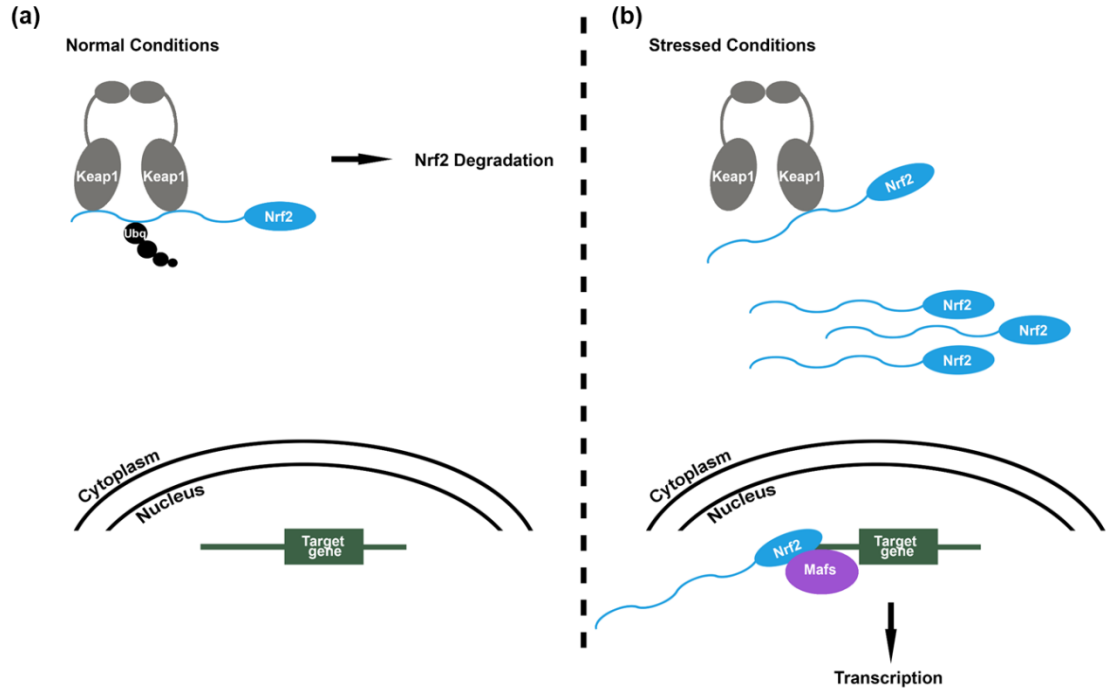

Figure 3.3. Nrf2-Keap1 redox control pathway. (a) Under normal conditions, Nrf2 and Keap1 interact in the cytoplasm, which results in the polyubiquitination of Nrf2. This leads to the degradation of Nrf2 by the 26S proteasome. (b) Under stressed conditions, Keap1 is redox sensitive protein where specific thiols become modified under oxidative or electrophilic stress, this alters the Nrf2-Keap1 complex. This causes the DLG motif to dissociate which results in the accumulation of new synthesized Nrf2 in the cytoplasm where it can then enter the nucleus and bind the respective target gene.

Keap1 is a cysteine rich protein that acts as a redox sensor under oxidative/electrophilic stress primarily in its IVR. Keap1 contains numerous reactive cysteines with a pKa value around 4-5 due to the microenvironment of the surrounding amino acids.[27] Therefore, at physiological pH, these reactive cysteines exist as thiolate

anions (RS^-).[28] When a cell is under oxidative/electrophilic stress, certain cysteines are more reactive towards specific stressors. Modification of cysteines (for example, formation of disulfide bonds) on Keap1 alter the structure of the Nrf2-Keap1 complex. This causes the lower affinity DLG binding site to dissociate from Keap1, subsequently inhibiting polyubiquitination of Nrf2 displayed in Figure 3.3b. This allows newly synthesized Nrf2 to enter the nucleus.[21] Once in the nucleus, Nrf2 binds small Maf proteins and various other transcription factors where it acts upon the antioxidant response element/electrophile response element in the regulatory region of its target genes.[29] This causes the activation for various target genes that can result in the following; glutathione synthesis, elimination of ROS, xenobiotic metabolism, and drug transport.[30]

The structure of Nrf2 is poorly characterized. A solution NMR structure shows that isolated Neh1 (residues 405-483) adopts a helical conformation (Figure 3.4). NMR data suggest that the N-terminal domain (Neh2) is intrinsically disordered.[31] No other structural information is available for the full-length protein.

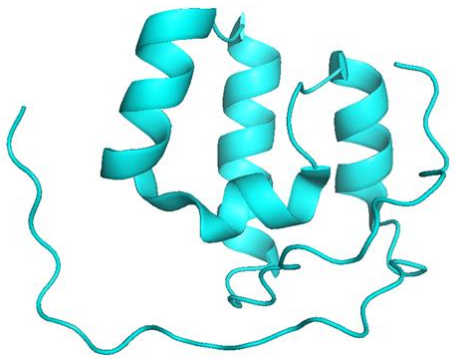

Figure 3.4. Neh1solution NMR structure. Residues 405-483 of Nrf2 display a helical architecture (pdb: 2LZ1).

Like Nrf2, there are no full-length high-resolution structures of Keap1. Single particle electron microscopy revealed that Keap1 resembled a 'cherry-bob' structure.[32] Displayed in Figure 3.5 are partial crystal structures that have been determined for the BTB and Kelch domain.[33-34] Figure 3.5a shows the BTB domain consists of a three-stranded β-sheet flanked by six α-helices, the first of which forms a key part of the dimerization interface with another Keap1 protein. The Kelch domain (Figure 3.5b) is a six-bladed β-propeller where each blade of the propeller is made up of four β-strands.[35] Each motif of Nrf2 binds in a pocket at the bottom of the entrance of the β-propeller.

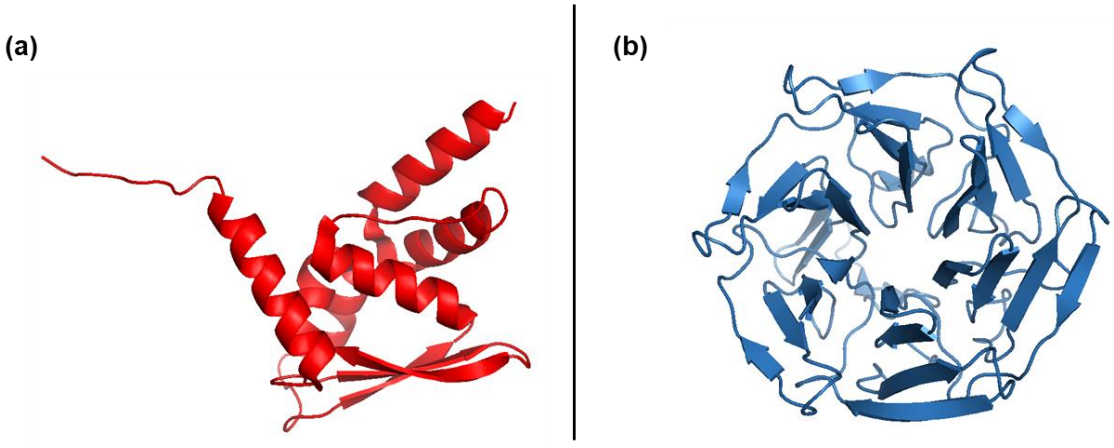

Figure 3.5. Partial crystal structures of Keap1. (a) Crystal structure of the BTB domain from residues 50-179 (pdb: 4CXI). (b) Crystal structure of the Kelch domain from residues 325-609 (pdb: 2FLU).

Due to the disordered nature of Nrf2 and the complexity of studying protein-protein interactions, investigating Nrf2 has proven to be challenging. Characterizing structural and dynamic features of Nrf2 and its interactions with the Kelch domain is essential for understanding how Nrf2 carries out its function. X-ray crystallography remains difficult for full-length Nrf2 due to its disordered nature. NMR based methods have also been

employed to study the structural and dynamic aspects,[19, 36] but those experiments are difficult due to the size and disordered nature of full-length Nrf2.

Backbone amide hydrogen/deuterium exchange (HDX) coupled with ESI-MS has become a popular method for studying protein behaviour in solution.[37-38] Regions that are involved in hydrogen-bonding networks or are occluded from the solvent will undergo slow exchange. In contrast, regions that are disordered or solvent accessible will undergo fast exchange upon protein exposure in D_2O-based solvent. The exchange process is due to dynamic fluctuations that disrupt amide hydrogen bonding. By measuring deuterium incorporation, it is possible to determine the relative flexibility of backbone segments of a protein (refer to Chapter 1 where this has been discussed in more detail).

HDX-MS has become a popular tool to interrogate IDPs that remain challenging to characterize by other techniques.[2, 39-41] Therefore, HDX-MS is a promising avenue to investigate the disordered full-length Nrf2. Since HDX-MS has almost no size limitation, investigating the Nrf2-Keap1 complex can easily be achieved. Using this approach will provide an understanding of the dynamic nature of full-length Nrf2 (see Figure 3.1) and pinpoint regions in Nrf2 and the Kelch domain (see Figure 3.5b) that change upon interacting with one another.

3.2 Experimental

3.2.1 Protein Expression and Purification

The protein used in this work was expressed and purified by Nadun Karunatilleke from Dr. James (Wing-Yiu) Choy's Laboratory in the UWO Biochemistry Department. Nrf2 and the truncated Kelch domain were cloned into the expression vector pDEST17. Nrf2 was isolated from E. *coli* (Rosetta 2(De3) pLysS). The Kelch domain was isolated from E. *coli* (BL21). Both proteins were purified using Ni-Sepharose affinity chromatography. The protein concentration of Nrf2 and Kelch domain were calculated by the Lowry and Bradford protein assay, respectively. These concentration values were confirmed by amino acid analysis.

Final protein stock solutions contained 50 mM phosphate buffer, 100 mM NaCl, and 1 mM dithiothreitol (DTT). Na_2PO_4, NaH_2PO_4, NaCl, and D_2O were purchased from Sigma Aldrich (St. Louis, MO).

3.2.2 Hydrogen/Deuterium Exchange-Mass Spectrometry

All HDX experiments were performed in 50 mM sodium phosphate buffer and 100 mM NaCl at pH 7 in 90% D_2O with a final protein concentration of 2 μM for Nrf2 and Kelch. In the binding experiments Nrf2 protein concentration was kept constant at 2 μM and the Kelch protein concentration was varied from 2, 4, and 6 μM in the 1:1, 1:2, and 1:3 (Nrf2:Kelch) binding experiments, respectively. Aliquots were removed at time points 0.1, 0.2, 0.4, 5, 30, and 60 minutes. The various aliquots were quenched by lowering the pH to 2.5 using 20% (v/v) formic acid (FA) and flash freezing in liquid N_2. The samples were

thawed and injected into a nanoACQUITY UPLC equipped with HDX technology (Waters, Milford, MA). Digestion was carried out online using a POROS pepsin column (2.1 mm x 30 mm) from Life Technologies/Applied Biosystems (Carlsbad, CA) held at 15 ^{o}C. Subsequently, the peptides were trapped on a C_{18} BEH130 VanGuard column (5 mm x 1 mm, 1.7 μm) for three minutes at 80 μL/min. The peptides were separated on a C_8 (50 mm x 2.1 mm, 1.7 μm) at a flow rate of 100 μL/min using a water/acetonitrile gradient acidified with 0.1 % FA.

The LC outflow was directed to a 2nd generation Synapt Q-TOF G2 instrument (Waters, Milford, MA). The instrument was operated in positive ion mode at 2.8 kV and a cone voltage of 20.0 V. The desolvation and source temperatures were set to 250 ^{o}C and 80 ^{o}C, respectively. Mass spectra were acquired for 25 minutes and obtained within 50-2000 m/z range in resolution mode. Ion mobility separation was employed to aid in separating overlapping isobaric peaks.

3.2.3 HDX Data Analysis

Peptide identification was performed using three separate label free data independent acquisitions (MS^E). MS^E data were analyzed using Protein Lynx Global Server 2.5.3. DynamX 3.0 was used for HDX data analysis and all computer assignments were verified manually. Deuterium uptake levels were corrected for artificial in-exchange and back-exchange using a protein sample which exhibits the minimum exchange that could possibly occur (m_0) and a fully deuterated protein sample (m_{100}), respectively. Deuterium uptake for each peptide at exposure time *t*, was calculated according to:

$$\% D(t) = \frac{m_t - m_0}{m_{100} - m_0} \times 100\%$$

Online digestion yielded 101 peptides for Nrf2 resulting in 86.3% coverage shown in Figure 3.6. Online digestion of Kelch yield 109 peptides resulting in 99.7% coverage displayed in Figure 3.7.

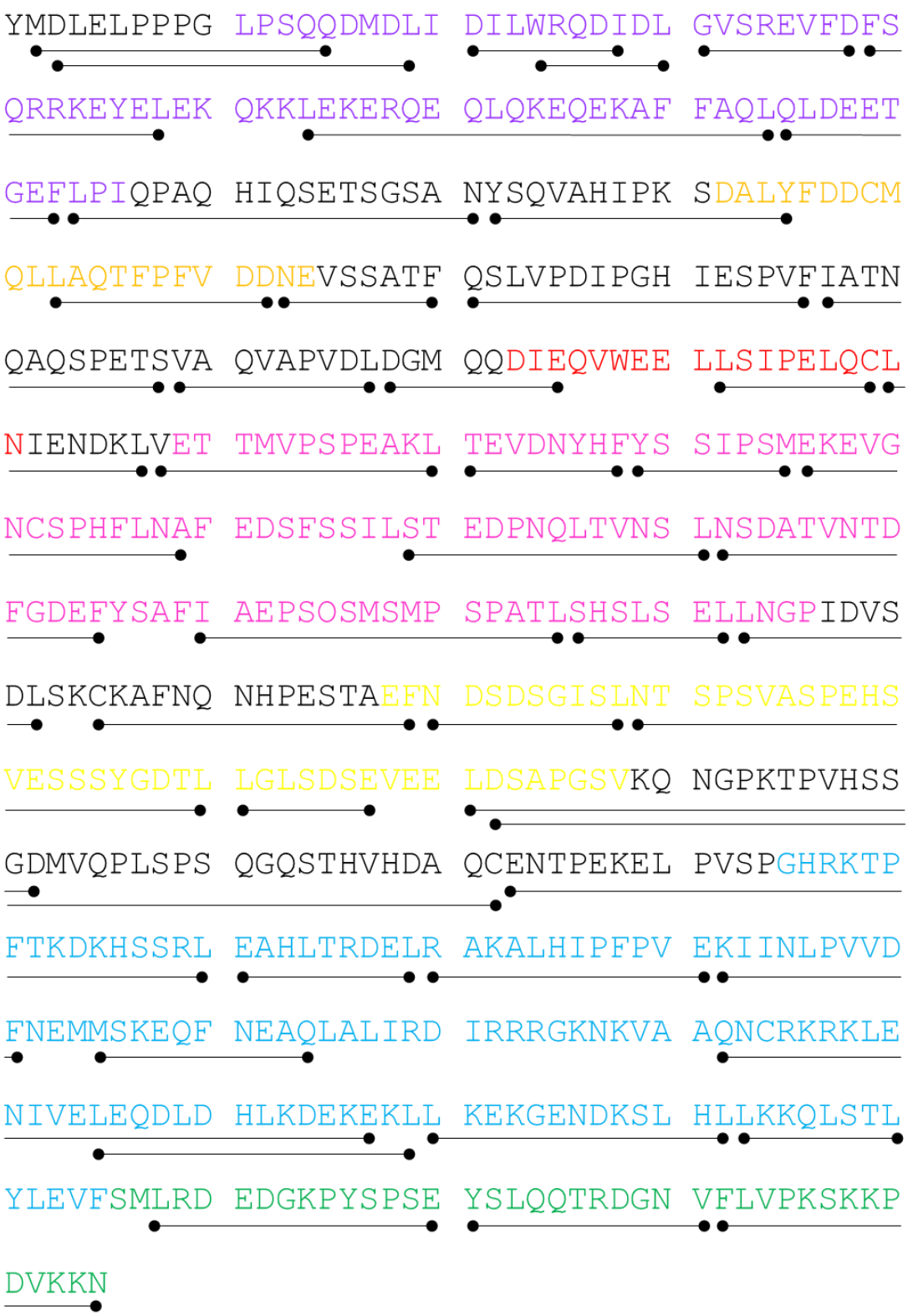

Figure 3.6. Nrf2 peptide coverage. Online digestion yielded 101 peptides resulting in 86.3% coverage. All seven domains were covered in the HDX-MS experiments. The sequence colouring corresponds to the domain colours in Figure 3.1.

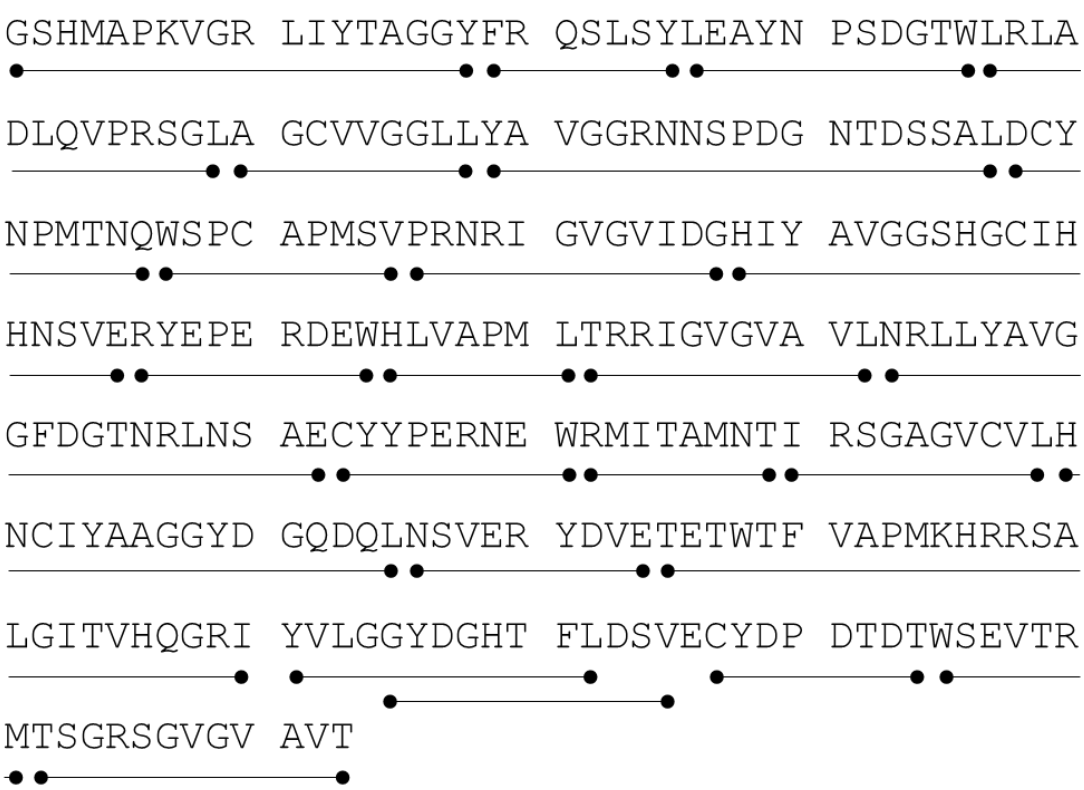

Figure 3.7. Kelch sequence coverage. Online digestion yielded 109 peptides resulting in 99.7% coverage.

3.3 Results and Discussion

It has remained difficult to study full-length Nrf2 and its interaction with the Kelch domain using X-ray crystallography and NMR methods due to its large size and dynamic nature. Here, HDX-MS is used to gain insights into structural and dynamic aspects of Nrf2 and its interaction with the Kelch domain.

Experiments show that the Neh2 domain of Nrf2 can bind to the full-length Keap1 at low nanomolar concentrations.[31, 42] Initial isothermal calorimetry from our collaborators showed the K_d values of the ETGE and DLG binding motifs in the full-length Nrf2 construct when bound to truncated Kelch are 3.01, and 274 nM, respectively owing to more

electrostatic interactions between Kelch and ETGE motif.[43] These values are consistent with what has been demonstrated in the literature.[31]

3.3.1 Effects of Kelch on Nrf2

Initial HDX-MS data showed that free full-length Nrf2 is highly dynamic throughout its entire sequence as seen from the data in Figure 3.8 (black). Fast HDX kinetics are observed throughout the protein, where all peptides exhibit 100% deuteration after 12 seconds. Surprisingly, the Neh1 domain exhibits no signs of a helical architecture that was displayed in its NMR structure.[44] This could be an artifact of the high

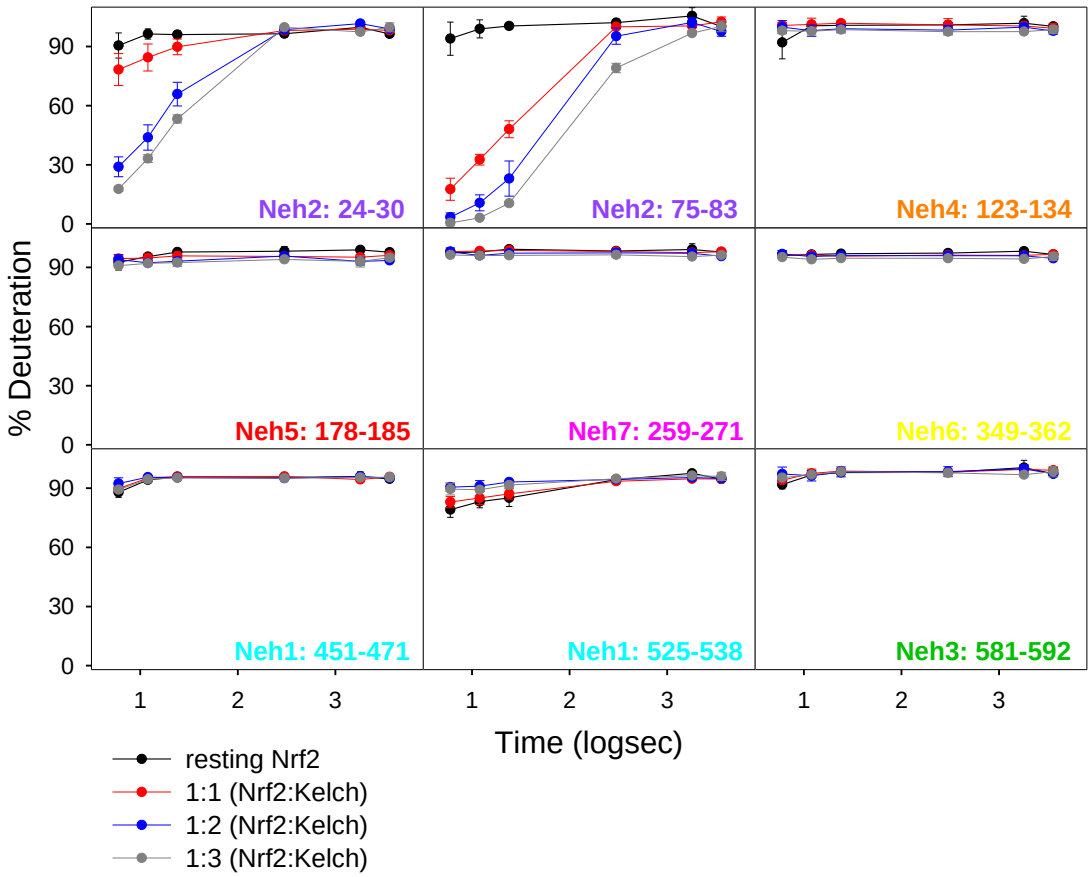

Figure 3.8. Nrf2 uptake plots of the seven different domains for free Nrf2 (black), 1:1 (red). 1:2 (blue), and 1:3 (gray) Nrf2:Kelch.

concentration of arginine (50 mM) that was used in the solution NMR experiment where the Neh1 structure was generated.[44] Arginine has been found to suppress protein aggregation by increasing protein stability.[45] This could be a possible reason for the absence of the helical structure in this work.

The effects of Kelch on the structure and dynamics of full-length Nrf2 was also investigated. Various conditions were tested. The Nrf2 concentration was held constant at 2 μM, while the Kelch concentration varied from 2, 4, and 6 μM which correspond to Nrf2-Kelch ratios of 1:1, 1:2, and 1:3, respectively. By testing these different ratios of protein concentration, it is possible to selectively populate the high affinity motif or both binding sites due to the different K_d values of the ETGE and DLG motifs. In Figure 3.8, major HDX changes are observed in the Neh2 domain when Kelch is added. Upon 1:1 addition of Kelch (red), a large reduction in HDX kinetics is displayed in the ETGE binding motif and a small reduction in deuterium uptake in the DLG motif, which is consistent with their K_d values. In a physiological setting, Nrf2 interacts with two Kelch domains. By interrogating the 1:2 (blue) state, a large reduction in deuterium uptake is observed in the DLG binding site and an even greater reduction in HDX in the ETGE motif. This effect was enhanced in both binding motifs in the 1:3 state (gray). In all three states, the addition of Kelch had no visible effect on the other domains of Nrf2.

Other experimental studies on Nrf2's binding motifs with the Kelch domain have shown to display consistent results with the observed HDX data presented in this work. The DLG motif (29-31) and ETGE motif (79-81) insert into the bottom of the Kelch domain in a similar manner. Data has shown that the ETGE peptide adopts a β-turn conformation that is buried into the binding surface of the Kelch domain.[25, 35] Various

electrostatic interactions are made by Glu79 and Glu82. Another X-ray structure investigated the DLG region complexed with Kelch and showed that this peptide formed an N-terminal helix from residues 19-25 and two short 3_{10} helices from residues 28-30 and 34- 37.[34] HDX is not able to differentiate between solvent accessibility and H-bonding. Therefore, the strong protection patterns observed in these regions when Nrf2 interacts with Kelch could be due to solvent occlusion as the binding motifs insert into the bottom pocket of Kelch, or due to secondary structure formation as shown from previous work.[25, 34-35]

Figure 3.9 shows the HDX % difference plots at t = 6 s of Nrf2 in the 1:1, 1:2, and 1:3 states. These maps use a scale where blue regions indicate less HDX than free Nrf2; red regions are indicative of more HDX (more dynamic) than free Nrf2. These plots clearly display the enhanced effect that is observed in the Neh2 domain by increasing the Kelch concentration. Surprisingly, all the other domains experience slightly more flexibility at the earliest time point upon binding the Kelch domain (red hues). These data suggest that when full-length Nrf2 binds to the Kelch domain the rest of the protein becomes somewhat more dynamic. Other IDPs that were studied using a full-length construct have also shown a more dynamic nature upon target binding distant from the binding site.[40] Due to the absence of studies on full-length Nrf2, this highlights a limitation when using truncated proteins in biophysical studies.

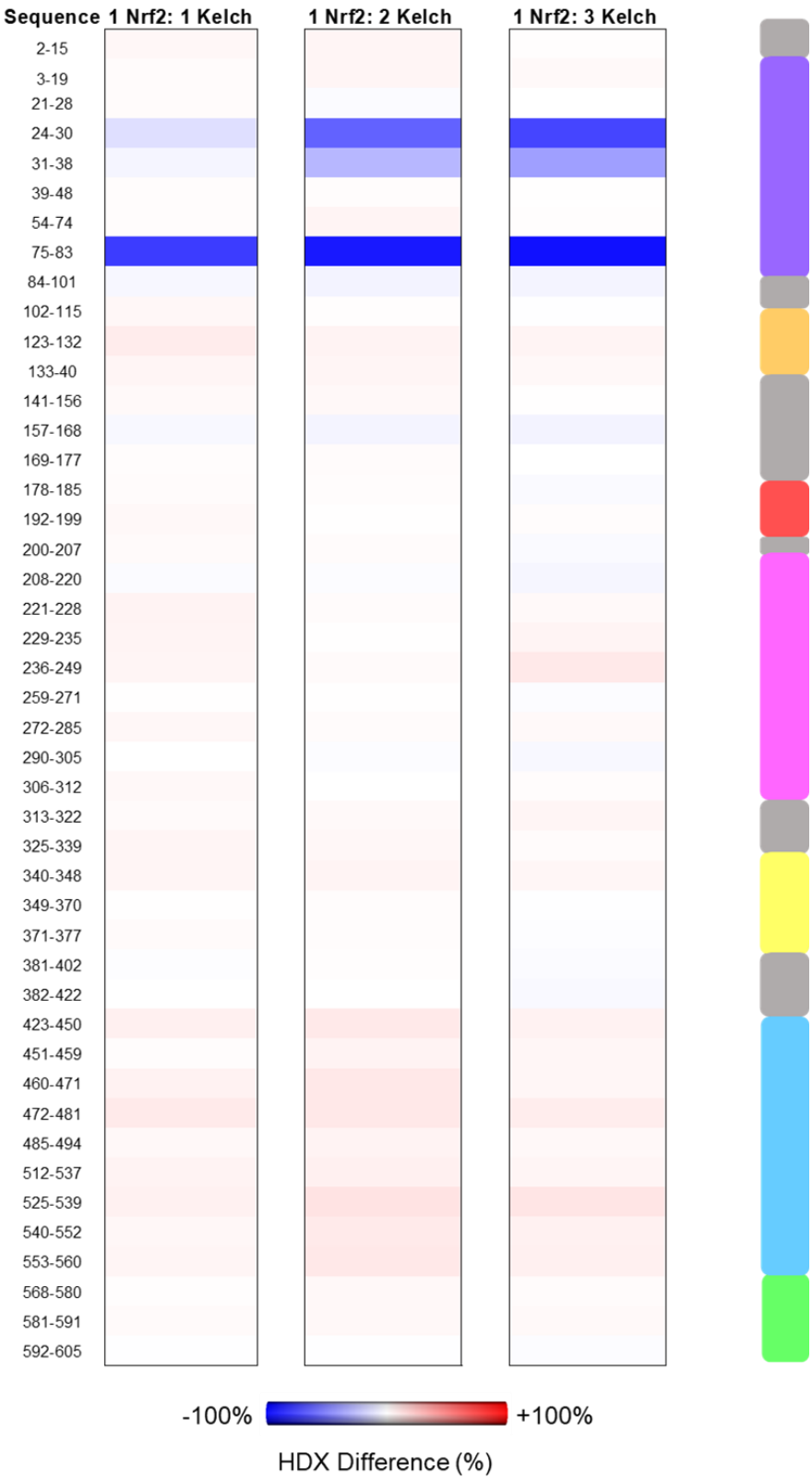

Figure 3.9. HDX % difference plots of Nrf2 when it interacts with Kelch in the 1:1, 1:2, and 1:3 states compared to free Nrf2 at t = 6 s. The side panel indicates the domain location as shown in Figure 3.1 (same colouring).

3.3.2 Effects of Nrf2 on the Kelch Domain

The effects of Nrf2 on Kelch can also be investigated. The deuteration kinetics of free Kelch (black) displayed in Figure 3.10 show slow uptake patterns throughout its entire sequence, consistent with its tightly folded β-propeller structure.[35] Previous X-ray data using truncated peptides of Nrf2 showed that the ETGE motif forms various hydrogen bonds when it interacts with Kelch. Padmanabhan *et al* demonstrate that Glu79 forms H-bonds with Kelch residues Arg415, Arg483, and Ser508, whereas Glu82 H-bonds with residues Ser363, Asn382, and Arg380.[25, 46] These regions may be of interest when investigating the effect Nrf2 has on Kelch since H-bonding could result in stabilization of the Kelch domain.

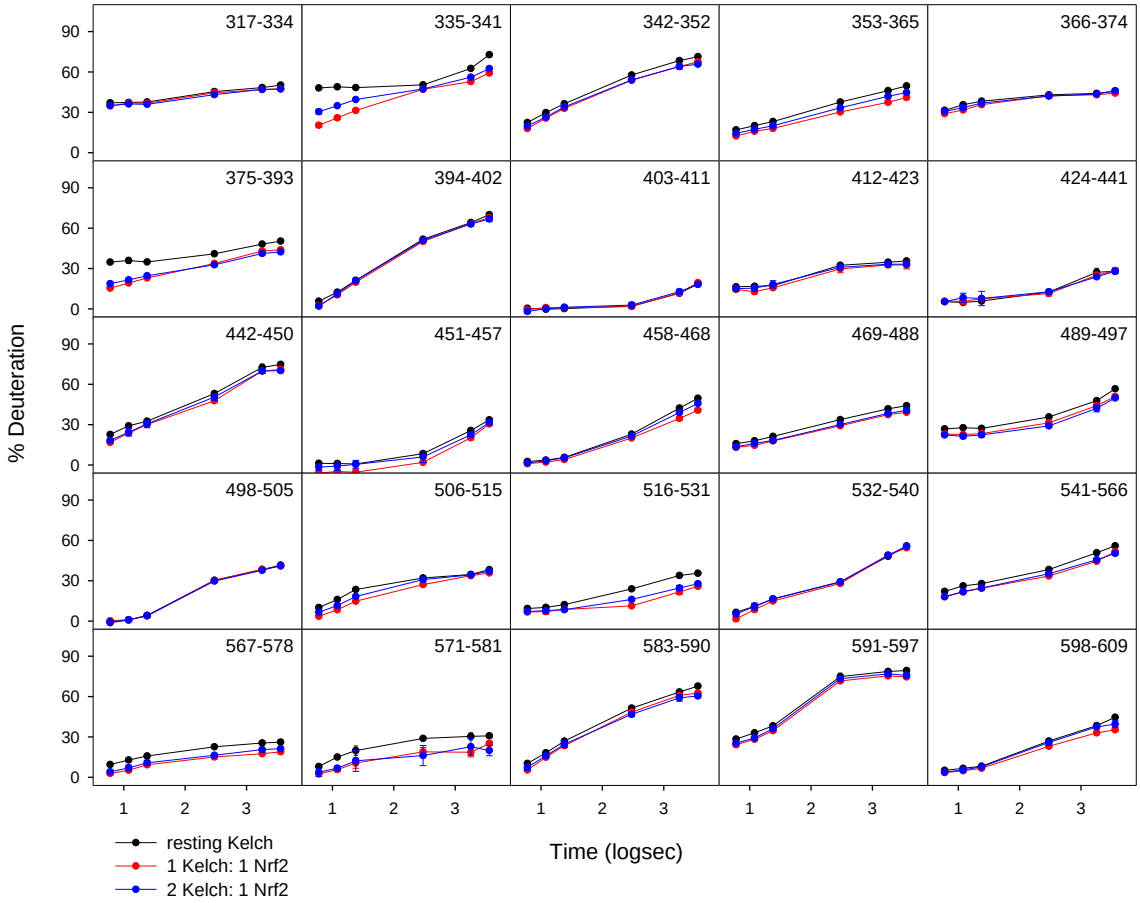

Figure 3.10. % Deuteration uptake plots for free Kelch (blue), 1:1 (red), and 2:1 (blue) Kelch:Nrf2.

In the 1:1 and 2:1 (Kelch:Nrf2) states, the overall HDX kinetics remained the same throughout the entire Kelch sequence. Peptides 335-341, 375-393, and 571-581 are the only three peptides that show a slight reduction in deuterium uptake when Nrf2 is added. Peptide 375-393 contains residue Asn382 that is found to form an H-bond with Nrf2. The HDX data is consistent with this as peptide 375-393 display a reduction in deuterium uptake indicative of protection of this site.

To understand the global effect Nrf2 has on Kelch, the HDX % difference of the 1:1 state was plotted at time points 0.1, 0.4, 30 minute compared to free Kelch. This is displayed in Figure 3.11 using a top and side view of Kelch. As stated above, the maps use a scale where blue regions indicate less HDX than free Kelch; red regions are indicative

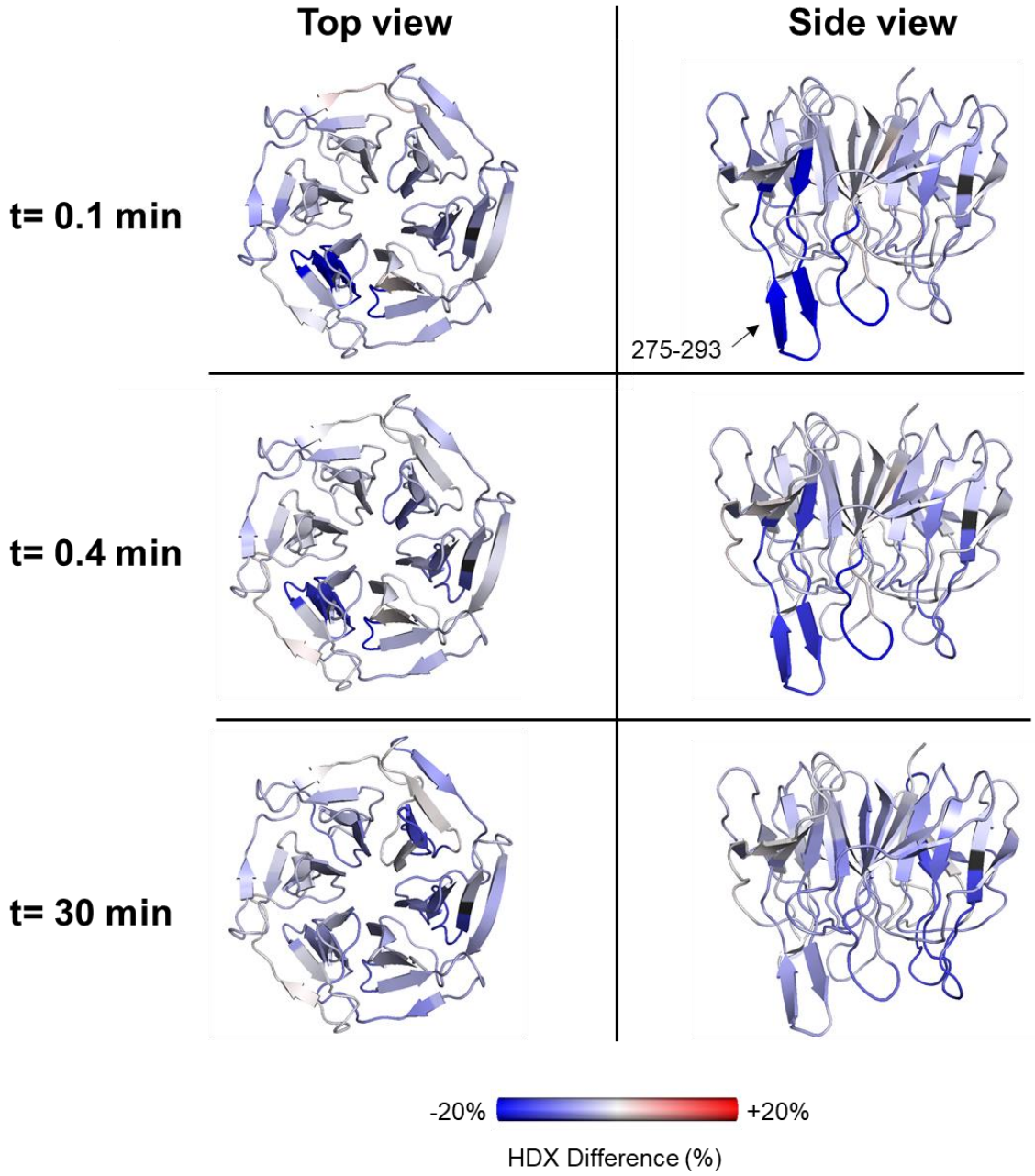

Figure 3.11. HDX % difference of Kelch when Nrf2 is added at 0.1, 0.4, and 30 mins compared to free Kelch.

of more HDX (more dynamic) than free Kelch. At 0.1 and 0.4 minutes a large reduction in deuterium uptake is observed in peptide 275-293 consistent with H-bond formation of this site when the Kelch domain interacts with Nrf2. This strong effect disappears at the 30-minute time point. Overall, the data show that there is a general stabilization of the entire Kelch domain when Nrf2 binds as a result of a reduction in deuterium uptake. Interestingly, this may be the first account of this global stabilization of the Kelch domain when it interacts with full-length Nrf2.

The data presented in this Chapter highlight the importance of using full-length proteins to investigate their structural and dynamic characteristics using biophysical techniques. Earlier investigations using full-length Nrf2 have been unsuccessful. Using the full-length Nrf2 construct in this work uncovers important structural and dynamic characteristics that provide insights into Nrf2's biological role. It was uncovered that free Nrf2 is highly disordered throughout its entire sequence, contrary to what has been reported before.[44] When Nrf2 interacts with the Kelch domain, a large reduction in deuterium uptake is observed in the DLG and ETGE binding motifs of Nrf2. The rest of Nrf2 became slightly more dynamic shown by an increase in HDX. It was also found that when Nrf2 interacts with Kelch, it resulted in a global stabilization of the entire Kelch domain. These findings highlight the limitations of interrogating truncated proteins when using biophysical techniques. Since full-length Nrf2 has not been successfully investigated previously, HDX-MS has been a successful technique in providing new insights into the function of this important protein. These findings could help in numerous applications such as drug therapies related cancer, neurodegenerative diseases, as well as autoimmune and inflammatory diseases where Nrf2 has been found to play an important role.

3.4 Conclusions

HDX-MS is used in this work as a technique that allows the interrogation of large disordered proteins and protein complexes. Full-length Nrf2 has not yet been successfully studied to date, due to its large size and highly dynamic nature. Here, full-length Nrf2 is investigated using HDX-MS. The data show that in the context of the full-length protein, Nrf2 is exclusively disordered opposing what has previously been reported.[44] When Nrf2 is in the presence of the Kelch domain, a large reduction in deuterium uptake is observed in the ETGE and DLG binding motifs and the rest of the protein become slightly more dynamic. The effect Nrf2 has on the structure of Kelch was also investigated and show that full-length Nrf2 induces a global stabilization on the Kelch domain which has not been reported before.

Full-length Nrf2 has continued to be difficult to study with classic high-resolution techniques such as X-ray crystallography and NMR. However, using HDX-MS in this work has provided insight into the dynamic disposition of Nrf2 and the effect Nrf2 and Kelch have on each other when they interact. The findings in this work emphasize the notable power of HDX-MS as an analytical tool. Since one third of all eukaryotic proteins are known to be intrinsically disordered, it is imperative to understand their structural and dynamic characteristics and how they are related to their function. HDX-MS is a promising technique that can be utilized to study other disordered proteins and protein-protein interactions like Nrf2.

3.5 References

1. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M., Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014,** *114* (13), 6589-6631.

2. Beveridge, R.; Chappuis, Q.; Macphee, C.; Barran, P., Mass spectrometry methods for intrinsically disordered proteins. *Analyst* **2013,** *138* (1), 32-42.

3. Dyson, H. J.; Wright, P. E., Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **2005,** *6* (3), 197-208.

4. Ishida, T.; Kinoshita, K., PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res.* **2007,** *35* (Web Server issue), W460-W464.

5. Yang, Z. R.; Thomson, R.; McNeil, P.; Esnouf, R. M., RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **2005,** *21* (16), 3369-3376.

6. Gsponer, J.; Futschik, M. E.; Teichmann, S. A.; Babu, M. M., Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* **2008,** *322* (5906), 1365-1368.

7. Oldfield, C. J.; Dunker, A. K., Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem* **2014,** *83*, 553-584.

8. Sigler, P. B., Acid blobs and negative noodles. *Nature* **1988,** *333* (6170), 210-212.

9. Gsponer, J.; Madan Babu, M., The rules of disorder or why disorder rules. *Prog. Biophys. Mol. Biol.* **2009,** *99* (2), 94-103.

10. Wright, P. E.; Dyson, H. J., Linking folding and binding. *Curr. Opin. Struct. Biol.* **2009,** *19* (1), 31-38.

11. Spolar, R.; Record, M., Coupling of local folding to site-specific binding of proteins to DNA. *Science* **1994,** *263* (5148), 777-784.

12. Sugase, K.; Dyson, H. J.; Wright, P. E., Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **2007,** *447* (7147), 1021-1025.

13. Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J., Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **2011,** *21* (3), 432-440.

14. Wright, P. E.; Dyson, H. J., Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **2015,** *16* (1), 18-29.

15. He, X.; Chen, M. G.; Lin, G. X.; Ma, Q., Arsenic induces NAD (P) H-quinone oxidoreductase I by disrupting the Nrf2· Keap1· Cul3 complex and recruiting Nrf2· Maf to the antioxidant response element enhancer. *J. Biol. Chem.* **2006,** *281* (33), 23620-23631.

16. Sykiotis, G. P.; Bohmann, D., Stress-Activated Cap'n'collar Transcription Factors in Aging and Human Disease. *Science Signaling* **2010,** *3* (112), re3-re3.

17. Suzuki, T.; Motohashi, H.; Yamamoto, M., Toward clinical application of the Keap1–Nrf2 pathway. *Trends Pharmacl Sci* **2013,** *34* (6), 340-346.

18. Chen, P.-C.; Vargas, M. R.; Pani, A. K.; Smeyne, R. J.; Johnson, D. A.; Kan, Y. W.; Johnson, J. A., Nrf2-mediated neuroprotection in the MPTP mouse model of Parkinson's disease: Critical role for the astrocyte. *Pro. Natl. Acad. Sci. U. S. A.* **2009,** *106* (8), 2933-2938.

19. Khan, H.; Killoran, Ryan C.; Brickenden, A.; Fan, J.; Yang, D.; Choy, W.-Y., Molecular effects of cancer-associated somatic mutations on the structural and target recognition properties of Keap1. *Biochem. J* **2015,** *467* (1), 141-151.

20. McMahon, M.; Thomas, N.; Itoh, K.; Yamamoto, M.; Hayes, J. D., Redox-regulated Turnover of Nrf2 Is Determined by at Least Two Separate Protein Domains, the Redox-sensitive Neh2 Degron and the Redox-insensitive Neh6 Degron. *J. Biol. Chem.* **2004,** *279* (30), 31556-31567.

21. Canning, P.; Sorrell, F. J.; Bullock, A. N., Structural basis of Keap1 interactions with Nrf2. *Free Radic Biol Med* **2015,** *88* (Pt B), 101-107.

22. Nioi, P.; Nguyen, T.; Sherratt, P. J.; Pickett, C. B., The Carboxy-Terminal Neh3 Domain of Nrf2 Is Required for Transcriptional Activation. *Mol Cell Biol* **2005,** *25* (24), 10895-10906.

23. Itoh, K.; Wakabayashi, N.; Katoh, Y.; Ishii, T.; Igarashi, K.; Engel, J. D.; Yamamoto, M., Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev* **1999,** *13* (1), 76-86.

24. Zheng-Yu, J.; Hong-Xi, C.; Mei-Yang, X.; Ting-Ting, Y.; Jian-Min, J.; Jing-Jie, H.; Xiao-Ke, G.; Xiao-Jin, Z.; Qi-Dong, Y.; Hao-Peng, S., *Domain structure of Nrf2 and Keap1*. 2013.

25. Canning, P.; Sorrell, F. J.; Bullock, A. N., Structural basis of Keap1 interactions with Nrf2. *Free Radic Biol Med* **2015,** *88, Part B*, 101-107.

26. Kobayashi, A.; Kang, M.-I.; Okawa, H.; Ohtsuji, M.; Zenke, Y.; Chiba, T.; Igarashi, K.; Yamamoto, M., Oxidative Stress Sensor Keap1 Functions as an Adaptor for Cul3-Based E3 Ligase To Regulate Proteasomal Degradation of Nrf2. *Mol Cell Biol* **2004,** *24* (16), 7130-7139.

27. Ma, Q., Role of Nrf2 in Oxidative Stress and Toxicity. *Annu. Rev. Pharmacool. Toxicol.* **2013,** *53*, 401-426.

28. Baird, L.; Dinkova-Kostova, A. T., The cytoprotective role of the Keap1–Nrf2 pathway. *Arch. Toxicol.* **2011,** *85* (4), 241-272.

29. Itoh, K.; Tong, K. I.; Yamamoto, M., Molecular mechanism activating Nrf2–Keap1 pathway in regulation of adaptive response to electrophiles. *Free Radic Biol Med* **2004,** *36* (10), 1208-1213.

30. Bryan, H. K.; Olayanju, A.; Goldring, C. E.; Park, B. K., The Nrf2 cell defence pathway: Keap1-dependent and -independent mechanisms of regulation. *Biochem. Pharmacol.* **2013,** *85* (6), 705-717.

31. Tong, K. I.; Katoh, Y.; Kusunoki, H.; Itoh, K.; Tanaka, T.; Yamamoto, M., Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. *Mol Cell Biol* **2006,** *26* (8), 2887-2900.

32. Ogura, T.; Tong, K. I.; Mio, K.; Maruyama, Y.; Kurokawa, H.; Sato, C.; Yamamoto, M., Keap1 is a forked-stem dimer structure with two large spheres enclosing the intervening, double glycine repeat, and C-terminal domains. *Pro. Natl. Acad. Sci. U. S. A.* **2010,** *107* (7), 2842-2847.

33. Cleasby, A.; Yon, J.; Day, P. J.; Richardson, C.; Tickle, I. J.; Williams, P. A.; Callahan, J. F.; Carr, R.; Concha, N.; Kerns, J. K.; Qi, H.; Sweitzer, T.; Ward, P.; Davies, T. G., Structure of the BTB Domain of Keap1 and Its Interaction with the Triterpenoid Antagonist CDDO. *PLoS ONE* **2014,** *9* (6), e98896.

34. Fukutomi, T.; Takagi, K.; Mizushima, T.; Ohuchi, N.; Yamamoto, M., Kinetic, Thermodynamic, and Structural Characterizations of the Association between Nrf2-DLGex Degron and Keap1. *Mol Cell Biol* **2014,** *34* (5), 832-846.

35. Lo, S.-C.; Li, X.; Henzl, M. T.; Beamer, L. J.; Hannink, M., Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *EMBO J* **2006,** *25* (15), 3605-3617.

36. Cino, E.; Fan, J.; Yang, D.; Choy, W.-Y., 1H, 15N and 13C backbone resonance assignments of the Kelch domain of mouse Keap1. *Biomol NMR Assign* **2013,** *7* (2), 149-153.

37. Iacob, R. E.; Engen, J. R., Hydrogen exchange mass spectrometry: Are we out of the quicksand? *J. Am. Soc. Mass. Spectrom.* **2012,** *23* (6), 1003-1010.

38. Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C., Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry—A review. *Anal. Chim. Acta* **2012,** *721*, 7-21.

39. Keppel, T. R.; Howard, B. A.; Weis, D. D., Mapping Unstructured Regions and Synergistic Folding in Intrinsically Disordered Proteins with Amide H/D Exchange Mass Spectrometry. *Biochemistry* **2011,** *50* (40), 8722-8732.

40. Goswami, D.; Callaway, C.; Pascal, Bruce D.; Kumar, R.; Edwards, Dean P.; Griffin, Patrick R., Influence of Domain Interactions on Conformational Mobility of the Progesterone Receptor Detected by Hydrogen/Deuterium Exchange Mass Spectrometry. *Structure* **2014,** *22* (7), 961-973.

41. Keppel, T. R.; Weis, D. D., Mapping Residual Structure in Intrinsically Disordered Proteins at Residue Resolution Using Millisecond Hydrogen/Deuterium Exchange and Residue Averaging. *J. Am. Soc. Mass. Spectrom.* **2015,** *26* (4), 547-554.

42. Eggler, A. L.; Liu, G.; Pezzuto, J. M.; van Breemen, R. B.; Mesecar, A. D., Modifying specific cysteines of the electrophile-sensing human Keap1 protein is insufficient to disrupt binding to the Nrf2 domain Neh2. *Pro. Natl. Acad. Sci. U. S. A.* **2005,** *102* (29), 10070-10075.

43. Tong, K. I.; Padmanabhan, B.; Kobayashi, A.; Shang, C.; Hirotsu, Y.; Yokoyama, S.; Yamamoto, M., Different Electrostatic Potentials Define ETGE and DLG Motifs as Hinge and Latch in Oxidative Stress Response. *Mol Cell Biol* **2007,** *27* (21), 7511-7521.

44. Eletsky, A.; Pulavarti, S.; Lee, D.; Kohan, E.; Janjua, H.; Xiao, R.; Acton, T.; Everett, J.; Montelione, G.; Szyperski, T., Solution NMR Structure of the DNA-Binding Domain of Human NF-E2 Releated Factor 2, Northeast Structural Genomics Consortium (NESG) Target HR3520O.

45. Tsumoto, K.; Umetsu, M.; Kumagai, I.; Ejima, D.; Philo, J. S.; Arakawa, T., Role of Arginine in Protein Refolding, Solubilization, and Purification. *Biotechnol. Progr.* **2004,** *20* (5), 1301-1308.

46. Padmanabhan, B.; Tong, K. I.; Ohta, T.; Nakamura, Y.; Scharlock, M.; Ohtsuji, M.; Kang, M.-I.; Kobayashi, A.; Yokoyama, S.; Yamamoto, M., Structural Basis for Defects of Keap1 Activity Provoked by Its Point Mutations in Lung Cancer. *Mol Cell* **2006,** *21* (5), 689-700.

Chapter 4 - Conclusions and Future Work

4.1 Conclusions

Proteins are macromolecules that carry out essential biological functions. Understanding a protein's three-dimensional native structure is imperative to understanding their function. In recent years, it has become clear that in addition to structure, dynamics play a significant role for protein function.[1-2] A variety of biophysical techniques have been introduced to study protein structure and dynamics. X-ray crystallography and NMR spectroscopy represent techniques that interrogate proteins on an atomic scale. X-ray crystallography remains the gold standard for atomically resolved structural data however, it only provides static snap shots.[1, 3] NMR spectroscopy can probe both structural and dynamic characteristics at near-native solution conditions, however, its application to large protein systems remains challenging.[4-5] Optical spectroscopy provides insights into global structural changes while detailed structural features remain hidden.[3, 6] Ideally, all protein studies would involve the use of high-resolution techniques, but this is not always feasible. HDX-MS has become an indispensable analytical tool that gives insight into both structure and dynamics in solution and that is widely applicable regardless of system size.

In Chapter 2, rabbit muscle pyruvate kinase was investigated using HDX-MS. This work concluded that efforts to elucidate linkages between enzyme dynamics and catalysis via comparative measurements should be treated with caution. Dramatic changes were observed in the extent of conformational dynamics during catalysis. However, this work showed that these changes are not unique to catalysis, but they arise from binding

interactions that can also be implemented in the absence of turnover. It appeared that the magnitude of such binding-associated dynamic changes dwarfs subtle conformational events that might be uniquely linked to enzymatic turnover.

Chapter 3 focused on interrogating the intrinsically disordered protein called Nrf2. Full-length Nrf2 and its interaction with the Kelch domain continue to pose experimental challenges due to their size and dynamic predisposition. HDX-MS showed that full-length Nrf2 is quite disordered throughout its entire sequence contrary to previous proposals in the literature generated on the basis of truncated protein constructs.[7] Interactions between Nrf2 and the Kelch domain showed a reduction in deuterium uptake for both Nrf2 binding sites. However, the rest of Nrf2 became slightly more dynamic upon binding Kelch. Such results have not been reported before, as full-length Nrf2 remains challenging to study. This work highlights the limitations of using truncated proteins in biophysical studies.

4.2 Future Work

4.2.1 Application of HDX to Other Enzymes

HDX-MS has proven to be a powerful technique for probing protein structure and dynamics. In the future, HDX-MS can be applied to other enzymes for obtaining a better understanding of the relationship between protein dynamics and catalysis. Recent investigations on enzymes revealed that the protein dynamics during catalysis are an intrinsic property,[8] where other studies have shown enhanced dynamics during catalysis.[9] These findings may be on a case-by-case basis so investigating other enzymes using HDX-MS is important for understanding the basis of protein dynamics during catalysis. Since

HDX-MS has no upper size limit, studying the dynamics of larger enzymes that remain challenging for other techniques is promising as well.

4.2.2 Application of HDX to Other Intrinsically Disordered Proteins

IDPs are a unique class of proteins as they lack a well-defined three-dimensional structure. They play an important role in cancer and many neurodegenerative diseases.[10-11] Since IDPs are difficult to study by other techniques, truncated segments are commonly investigated. This work has proven that this approach is not adequate to only study segments of IDPs. HDX-MS has demonstrated to be a key technique to study IDPs since it routinely reports on disorder regions. Studying full-length IDPs whose structure remains unknown, may give important insights into how they behave in solution and how they interact with their binding partner. IDPs encompass one third of all proteins in eukaryotic organisms, such investigations could provide important insights that can be used in drug therapies for many human diseases.

4.3 References

1. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-972.

2. Karplus, M.; Kuriyan, J., Molecular dynamics and protein function. *Pro. Natl. Acad. Sci. U. S. A.* **2005,** *102* (19), 6679-6685.

3. Robinson, C. V.; Sali, A.; Baumeister, W., The molecular sociology of the cell. *Nature* **2007,** *450* (7172), 973-982.

4. Montelione, G. T.; Zheng, D.; Huang, Y. J.; Gunsalus, K. C.; Szyperski, T., Protein NMR spectroscopy in structural genomics. *Nat. Struct. Mol. Biol.* **2000,** *7*, 982-985.

5. Kay, L. E., Protein dynamics from NMR. *Nat. Struct. Mol. Biol.* **1998**.

6. Brahms, S.; Brahms, J., Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J. Mol. Biol.* **1980,** *138* (2), 149-178.

7. Eletsky, A.; Pulavarti, S.; Lee, D.; Kohan, E.; Janjua, H.; Xiao, R.; Acton, T.; Everett, J.; Montelione, G.; Szyperski, T., Solution NMR Structure of the DNA-Binding Domain of Human NF-E2 Releated Factor 2, Northeast Structural Genomics Consortium (NESG) Target HR3520O.

8. Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D., Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007,** *450* (7171), 838-844.

9. Liuni, P.; Jeganathan, A.; Wilson, D. J., Conformer Selection and Intensified Dynamics During Catalytic Turnover in Chymotrypsin. *Angew. Chem. Int. Ed.* **2012,** *51* (38), 9666-9669.

10. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annu Rev Biophys* **2008,** *37* (1), 215-246.

11. Kumar, D.; Sharma, N.; Giri, R., Therapeutic Interventions of Cancers Using Intrinsically Disordered Proteins as Drug Targets: c-Myc as Model System. *Cancer Inform* **2017,** *16*, 1176935117699408.

Curriculum vitae

Courtney S. Fast

M.Sc. Candidate

University of Western Ontario

Education

Master of Science (M.Sc): Physical/Analytical Chemistry Sept 2015-Aug 2017
University of Western Ontario, London, ON

Bachelor of Science (B.Sc): Honors Specialization in Chemistry Sept 2011-Apr 2015
University of Western Ontario, London, ON

Employment

Graduate Research Assistant Sept 2015-Aug 2017
University of Western Ontario, London, ON

Teaching Assistant Sept 2011-Apr 2015
University of Western Ontario, London, ON

Peer-Reviewed Publications

1. Fast, C. S.; Vahidi, S.; Konermann, L. Using Hydrogen/Deuterium Exchange-Mass Spectrometry to Probe Changes in Enzyme Dynamics During Catalysis. (**Submitted**)

Conference Presentations

1. Fast, C.S.; Vahidi, S.; Konermann, L. "Using Hydrogen/Deuterium Exchange-Mass Spectrometry to Probe Changes in Enzyme Dynamics During Catalysis." 65th American Society for Mass Spectrometry (ASMS) Conference, Indianapolis, IN, **June 2017**. (Poster Presentation).
2. Fast, C.S.; Vahidi, S.; Dunn, S.D.; Konermann, L. "Examining the Importance of Conformational Dynamics for the Function of Pyruvate Kinase: Insight from HDX-MS." 64th American Society for Mass Spectrometry (ASMS) Conference, San Antonio, TX, **June 2016**. (Poster Presentation).
3. Fast, C.S; Vahidi, S.; Busnello, C.; Bi, Y.; Dunn, S.D.; Konermann, L. "Unraveling Protein Interactions within the F_OF_1 ATP Synthase Complex by HDX-MS." 63th Post-American Society for Mass Spectrometry (ASMS) Conference, Toronto, ON, **October 2015**. (Poster Presentation).
4. Fast, C.S; Vahidi, S.; Busnello, C.; Bi, Y.; Dunn, S.D.; Konermann, L. "Unraveling Protein Interactions within the F_OF_1 ATP Synthase Complex by HDX-MS." 63th American Society for Mass Spectrometry (ASMS) Conference, St. Louis, MO, **June 2015**. (Poster Presentation).