Conformational Analysis of Disordered Proteins Using H/D Exchange Mass Spectrometry

By

Theodore Robert Keppel

Submitted to the graduate degree program in Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Chairperson David Weis

_____
Heather Desaire

_____
Mario Rivera

_____
Carey Johnson

_____
Roberto De Guzman

Date Defended: 01/31/2013

The Dissertation Committee for Theodore R. Keppel certifies that this is the approved version for the following dissertation:

Conformational Analysis of Disordered Proteins Using H/D Exchange Mass Spectrometry

_____

Chairperson David Weis

Date Approved: 02/01/2013

Abstract

Intrinsically disordered proteins (IDPs) represent a growing area of scientific interest. The prevalence of IDPs within the proteomes of complex organisms and the importance of these proteins in cellular functions has become apparent. The application of hydrogen/deuterium exchange mass spectrometry (H/D-MS) to study IDPs is presented in this dissertation. Results that demonstrate improvements to the H/D-MS method are also presented. A thermoelectrically cooled refrigeration system was developed in order to house LC components necessary for bottom-up H/D-MS. This system was used to keep solvent temperatures low and stable, ensuring reproducible and minimized back exchange. Two model IDPs, the interacting domains of CREB binding protein (CBP, residues 2059-2117) and activator of thyroid and retinoid receptors (ACTR, residues 1018-1088), were analyzed using H/D-MS. CBP and ACTR represent two classes of IDPs: the molten globule and random coil, respectively. The mutual synergistic folding phenomenon observed when a complex is formed between these two proteins was also analyzed. A lower limit of several seconds of D_2O exposure time is imposed when manual pipetting is used to add label and quench buffers to protein samples. A quench flow apparatus was used to extend the lower time scale limit of exposure times to 42 milliseconds. When CBP and ACTR were labeled using the quench flow apparatus, analyses revealed subtle exchange protection in specific regions of both proteins. An H/D-MS study of the conformational dynamics of β-glucuronidase was also conducted. The destabilization of β-glucuronidase that results from the W492G mutation and the partial conformational rescue that results from indole addition were analyzed. A new method to interpret and normalize H/D-MS data was introduced to improve intra- and inter-laboratory comparability.

Acknowledgements

I would like to express my gratitude to the personnel at the University of Kansas for their aid and guidance throughout the course of my research. Specifically, and in no particular order, I would like to thank Prof. John Karanicolas for grant support and the fascinating β-glucuronidase project. I would like to thank Dr. Todd Williams for his help in conducting MS/MS analyses. I would like to thank Dr. Travis Witte for his help in the student analytical labs. I would like to thank Prof. Yong Zeng and Dr. Tanyu Wang for their help with capturing microscope images. I would like to thank Prof. Jennifer Laurence for her help with conducting gel electrophoresis and imaging. I'd like to thank my primary investigator, Prof. David Weis, for being my mentor these last few years and helping me to learn and master the concepts presented herein. I'd like to also thank the members of my research committee, including Prof. Heather Desaire, Prof. Mario Rivera, Prof. Carey Johnson, and Prof. Roberto De Guzman.

I'd also like to thank all of the teachers I've had throughout my life, starting with those in my graduate level classes at the University of Kansas. Going back, I'd like to thank my teachers at my undergraduate university, Truman State University in Kirksville, MO, and all of the wonderful teachers in the Parkway school district in Chesterfield, MO. I'd also like to include my parents and siblings among those teachers, as they provided lessons early on in my life that were necessary for me to excel when I reached the classroom.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction and Overview

1.1 Intrinsically Disordered Proteins

1.1.1 The Protein Quartet

Proteins are a class of biological molecule that many fields of science continue to study in order to characterize, understand, and modify major biological functions in the living cell. Biology has often looked toward the three-dimensional conformations and structural elements to analyze protein functions, from the lock and key model proposed in 1894.[1-2] More recently, there has been a growing interest in the study of proteins that exhibit regions with high flexibility or lack a stable secondary and tertiary structure, but are still able to carry out important biological functions.[3-14] While other terminology has been used, such as "natively disordered proteins," these proteins will be referred to as intrinsically disordered proteins (IDPs)[15] in this dissertation.

Proteins may generally be categorized into a quartet of groups, depending on their degree of disorder: ordered, pre-molten globule, molten globule, and random coil. Ordered proteins possess secondary and tertiary structural elements that have formed from the folding of the protein. Structure formation transitions often follow pathways that reduce the overall Gibbs free energy, *G*, of the protein.[16] By forming rigid structures and folding around a hydrophobic core, these proteins become compact and thermodynamically stable. This stability limits the number and variability of different conformers in the equilibrium ensemble of ordered proteins.[16] The degree of structural disorder in IDPs will differentiate their properties from those of well-structured, ordered proteins, which allows for a rough categorization of the three types of IDPs. The molten globule protein retains secondary structural elements. This property allows some traditional methods of protein structure determination to be applied, such as NMR (see section *1.2*).[17-21] The heterogeneous, inter-converting ensemble of conformers present in most molten globule protein samples excludes the application of conventional X-ray crystallographic studies, however.[22] The pre-molten globule state extends the degree of conformational fluctuation within

a protein, even causing the transient disruption of secondary structure.[23-24] This conformational disorder extends the protein's hydrodynamic radius.[23-24] This larger capture radius allows IDPs an advantage in forming binding interactions with partner molecules. This advantage may counter the slower diffusion of IDPs relative to compact, ordered proteins.[25] An IDP may also be classified as a random coil, a protein with no stable secondary structure and a large hydrodynamic radius. While a complete random coil may not realistically exist,[5, 26] the theoretical state of such a protein can provide a baseline for conformational dynamics studies.

1.1.2 Prevalence within the Proteome

Protein structure characterization methods have been developed mainly to study the three-dimensional conformation of ordered proteins (as will be discussed in section *1.2*). Even with a wide range of methods available to study proteins, it is obvious that attempting to characterize the entire human proteome would be a monumental undertaking. Identification of disordered proteins has often relied on prediction algorithms and the use of genome sequence predicted proteome databases (see section *1.2.1*). These algorithms may be applied to entire databases, categorizing proteins as either ordered or disordered and also correlating disorder frequency with associated cellular processes.[14, 27-29] These predictions reveal that the complex, multi-cellular organisms of the *Eukaryota* domain rely on a their proteomes to be relatively more disordered than the more simple *Archaea* and bacteria.[12, 27, 30] Indeed, of eukaryotic PDB crystal structures surveyed, a predicted 33% of these proteins contain regions of extended disorder.[27] A similar algorithm, based on consensus between cumulative distribution function and charge-hydropathy disorder prediction methods, categorizes 25-30% of eukaryotic proteins as disordered.[31] The high abundance of IDPs in eukaryotic proteomes highlights their importance in the complex cellular processes required of nucleated cells.

1.1.3 Linkages to Disease States and Function

Proteome correlation studies reveal the wide range of cellular processes that IDPs are involved in. Many nucleus- and chromosome-associated proteins are positively correlated with

intrinsic disorder.[29] IDPs are well represented in transcriptional regulation, cellular signaling, post-translational modifications, and cellular recognition.[14, 32-33] A correlation was drawn between proteins with predicted disorder within the yeast proteome and gene ontology annotations.[27] In this study, IDPs are more strongly associated with many nuclear processes such as transcription regulation and DNA packaging and cellular signaling processes such as morphogenesis and endocytosis. Another correlation revealed the strong association between IDPs and human cancer-associated proteins.[34] Other correlations reveal the higher propensity for disordered regions to undergo post-translational modification, particularly signaling/regulation-vital phosphorylation.[8]

1.1.4 Coupled Binding and Folding

While IDPs gain an advantage of higher hydrodynamic radius for protein-binding partner interactions, the specific mechanisms for recognition and binding remain nebulous. IDPs primarily undergo a coupled binding and folding process upon interacting with a target binding partner.[13, 19, 21, 35-39] Folding may also be induced by changes in the chemical environment, such as the addition of heat or protein stabilizers.[40] One popular theory for explaining IDP binding is the "fly-casting" mechanism.[13] A minor portion of an IDP will initiate binding to a structured binding partner and a conformational transition to a more ordered, stable structure will follow. The faster on- and off-rate for protein-partner binding and lower binding free-energy barriers (*i.e.* high specificity / low affinity)[41] also result from greater flexibility found in IDP conformation.[25] IDP flexibility also facilitates binding to multiple target partners.[35] IDPs can adopt different conformations depending on the binding partner.[10, 42] Differentiating these conformational transitions is an ongoing challenge. Validating analytical techniques to study conformational properties and transitions will be a primary focus of this dissertation.

1.1.5 Two Model IDPS: CBP and ACTR

Two model IDPs are studied extensively in chapters 3 and 4 of this dissertation. These IDPs are the nuclear coactivator binding domain (NCBD) of mouse cAMP-responsive element

(CRE)-binding (CREB) binding protein (residues 2059-2117, henceforth abbreviated “CBP”) and a binding domain of human p160 coactivator, activator for thyroid hormone and retinoid receptors (residues 1018-1088, henceforth abbreviated “ACTR”). The free proteins and the complex have been studied extensively by NMR, CD spectroscopy, and small angle X-ray scattering analyses.[19-21, 43] CBP and ACTR are both intrinsically disordered in their free states, but they form a structured complex when bound to each other. The complex formed between CBP and ACTR was the first reported observation of the synergistic folding process between interacting IDPs.[19, 44] Chapter 3 of this dissertation includes a review of these proteins and their properties. With a wealth of biophysical data available, these proteins and their complex serve as model systems to validate the application of hydrogen/deuterium exchange mass spectrometry (H/D-MS) to studying conformational dynamics of IDPs.

1.2 Methods for Characterizing Protein Structure

1.2.1 Disorder Prediction

Amino acid composition contributes to the propensity of proteins to be intrinsically disordered. IDPs tend to be enriched in more polar, disorder inducing amino acid residues and depleted in non-polar, hydrophobic core forming residues. Computerized algorithms predict disordered regions in proteins prior to an experimental characterization. One of the first developed algorithms was the Predictor Of Natural Disordered Regions (PONDR).[4, 45] PONDR was developed based on two training databases. One training set was of ordered proteins and another was of disordered proteins.[4, 46] A PONDR algorithm assigns a score to each residue in a protein based its identity and neighboring residues. IDPs have a greater abundance of A, R, G, Q, S, P, E, and K residues and a depletion in W, C, F, I , Y, V, L and N residues relative to ordered protein.[4, 46] This pattern indicates a general enrichment of charged and polar residues and a depletion of hydrophobic residues. PONDR scores, assigned on a residue by residue basis, give the propensity for a particular amino acid to be disordered. A PONDR score above 0.5 indicates disorder, while a score below 0.5 indicates order. Further iterations of PONDR

have been developed and adapted with expanded database information and improved computational techniques.[47] The PONDR-VL-XT algorithm disorder predictions for ACTR and CBP are shown in figures 1.1 and 1.2, respectively.[4, 48-49] These predictions indicate that most of ACTR will be disordered but CBP will be ordered along the C-terminal half of the protein. This correlates with the experimental evidence that indicates ACTR is a near random coil but CBP is a molten globule. However, when compared with specific NMR results,[19] the predictions do not always agree with residue specific locations of ordered and disordered regions along the sequence. Also, since these predictions are based on protein sequence, they cannot predict folding or unfolding that results from binding or solvent conditions. These shortcomings illustrate the need for experimental methods to study IDP conformations and interactions.

1.2.2 X-ray Crystallography

In an X-ray crystallography experiment, three dimensional protein structure can be determined by measuring the position of electron densities in a protein crystal.[50-51] This technique is limited, however, if certain residues exhibit multiple or low intensity electron densities characteristic of flexible regions.[51] The absence of X-ray diffraction can indicate disorder within a region of the protein. IDPs will have high conformational flexibility in their native state and crystals will be difficult to produce. Those crystals that may be produced from IDPs may not represent the native structure ensemble.[52]

1.2.3 Circular Dichroism

Far-UV circular dichroism (CD) is a useful technique for determining secondary structure content in a protein. By measuring the interaction of circularly polarized light (λ ~250 to 180 nm) with optically active peptide bonds, specific secondary structure content is observed as features along the CD spectrum.[7, 53-57] A characteristic spectrum for α-helical protein shows negative bands at 222 and 208 nm, with a positive band at 193 nm. Protein β-sheet content is indicated by a negative band at 218 nm and a positive band at 195 nm. An extensively disordered protein or random coil will also have a characteristic far-UV CD spectrum, with low ellipticity above 210

nm and negative bands around 195 nm. The relative quantities of these secondary structure elements within a protein can also be determined using CD. Near-UV CD (λ ~340 to 250 nm) may also reveal tertiary structure characteristics by measuring spectral features of aromatic side chains.[58] An additional benefit of using CD is the ability to monitor structural property changes from protein denaturation. Denaturation can be induced with pH change,[58] chemical denaturants,[59] or temperature.[60] CD is an important tool in determining protein structure content. However, localized information cannot be obtained from CD. With this limitation, CD best serves as a complementary technique in the study of localized disorder regions in IDPs.

1.2.4 NMR

Nuclear Magnetic Resonance (NMR) is one of the most useful techniques for localizing conformation and dynamics properties in structured proteins in solution. NMR works by measuring the local magnetic field properties of atomic nuclei that possess a spin, such as ^{1}H, ^{13}C, and ^{15}N.[7, 61] When a magnetic field excites these nuclei, NMR can measure the relaxation resonance frequency and assign the local chemical environment of resonating nuclei within a molecule.[7, 61] This property of NMR allows for single amino acid residue resolution conformation assignments. The ability to differentiate ^{2}H and ^{1}H nuclei makes NMR particularly useful in hydrogen/deuterium exchange experiments, as well.[52, 62] NMR allows the solution-phase study of some proteins that cannot be crystallized.[62-63] NMR has been successfully applied to study residual structure in denatured proteins.[64-65] However, there are several disadvantages to using NMR to study IDPs. NMR resonance signals tend to be highly overlapped with limited shift dispersion for disordered regions.[62] This may flood out important signals originating from structured regions of the protein. IDPs are also more prone to aggregation, complicating NMR experiments which require protein stability and high concentration samples.[62] Studying relatively large proteins using NMR is also challenging. While methods can be adapted to study large proteins with NMR,[66-67] processes requiring routine analysis of large proteins will still face this limitation.[68-69]

1.3 Mass Spectrometry-Based Approaches to Study IDPs

1.3.1 Advantages of MS for IDPs

Mass spectrometry (MS) methods have been used to study IDP conformation.[70-73] Solution-phase analyses, such as with electrospray ionization (ESI) (see section *1.6.1*),[74] are particularly advantageous for protein studies. Such analyses allow proteins to be prepared in physiological buffer prior to analysis, an advantage shared with NMR and CD. Alternative methods such as crystallography may require non-native conditions. MS also allows multi-protein systems to be analyzed, provided an efficient separation method is available. Such systems may be more difficult to analyze with NMR if the target protein is present at a very low relative concentration. Multi-protein systems cannot be analyzed using CD, as CD signals arise from all proteins in the solution, providing only an average signal for the entire ensemble of proteins and their conformational states.

An advantage of applying MS to characterize proteins is the capability and ease in which one may analyze large proteins, such as monoclonal antibodies.[75-78] While relevant to large IDPs, this advantage extends to large proteins in general. Another advantage of using MS techniques is the ability to detect small amounts of protein or peptide. Sub-femtomole MS and picomole tandem MS have been reported as early as 1991.[79] This, again, is important to the study of IDPs, as samples may require preparation at low concentrations to avoid precipitation or aggregation. This dissertation focuses on applications of MS to analyze IDPs. Several MS-based techniques to study protein conformation are discussed in the next three sections.

1.3.2 Approaches for Gas-Phase Analysis

The ESI process, when paired with mass analysis, may also be utilized to serve as an indicator of order or disorder itself. This technique is known as native ESI, as it is used to preserve proteins' native conformations in the gas phase. Certain considerations must be made when designing a native ESI method to best preserve the native conformation through ionization. The ESI solvent should be prepared to provide physiological conditions for an analyte

protein while still being compatible to the desolvation and ionization processes. Solvent properties to consider include pH and solvent volatility.[80] Non-volatile buffer salts can lead to a buildup of desolvated salts in the ESI source. Native ESI techniques are aided by the use of volatile buffer salts, such as ammonium acetate.[80] More compact/structured proteins will generally produce ions with less charge than unfolded/denatured proteins, leading to characteristic charge state distribution profiles.[60, 81] The mass spectra can indicate how folded/unfolded proteins are. Using native ESI, it is also possible to transfer weakly-associated protein complexes into the gas phase during ionization to study complex dissociation, structure, or dynamics.[81] The abilities to preserve complexes in the gas phase and to analyze folded and disordered proteins have allowed native ESI to be applied to analyzing binding and folding processes in IDPs.[82] ESI-MS has also been used to identify individual conformations within an IDP ensemble.[40, 71, 83]

Ion mobility MS (IM-MS) is a specific method of the native ESI technique that also allows gas phase protein analysis. In IM-MS, ion packets are separated based on a combination of cross-sectional area and mass-to-charge ratio (*m*/*z*), in a chamber with an applied an electric field.[81] Ion travel is hampered by the presence of an inert gas, with collisions between the analyte ions and inert gas adding the factor of ion mobility to the electrophoretic separation.[80-81] IM-MS can separate folded conformers from partially unfolded ones because of differences in charge and cross-section between them. Because solution-phase structural characteristics can be partially preserved in IM-MS, this manner of separation allows conformation changes, protein unfolding, and protein binding to be analyzed, even for very large proteins (*i.e.* in the MDa range).[81] Native ESI and IM-MS are limited in their applications to studying intact proteins or large subunits that can present a wide distribution of charge states, which will prevent localization of folding/unfolding without complementary experiments. The actual preservation of solution-phase, native conformations for proteins as they enter the gas phase is still somewhat controversial. Highly charged ions formed during the electrospray process have a higher

likelihood of forming extended conformations due to electrostatic repulsion.[83] While a folded protein may form ions with fewer charges, this problem may be exaggerated in IDPs, which are more likely to be natively extended to begin with. When IDPs become highly charged during ESI, any residual structuring they might have could be disrupted and undetected.

1.3.3 Solution-Based Labeling

Covalent labeling can reveal important dynamic events, if the label is applied on the time scale related to these events. A covalently bound label may provide a unique mass increase that may be measured and quantified by mass spectrometry. The purpose of covalent labeling methods is to provide information about protein conformation and dynamics. Covalent labeling can link the kinetics of label uptake to the structure of the protein, with more flexible regions and/or amino acid residues typically exhibiting faster incorporation rates.

One such labeling technique is the chemical modification of carboxyl-groups in proteins using a 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) or glycine ethyl ester (GEE).[84] The reaction of carboxyl groups with these chemicals selectively labels glutamate and aspartate residues.[84] A similar technique chemically labels exposed arginine residues using *p*-hydroxyphenylglyoxal (HPG).[85-86] Chemical cross-linking has also been useful in determining specific binding locations and affinities in proteins.[81, 87] By binding proteins with targeted ligands or binding partners and analyzing complexes by MS, localized information regarding binding affinities and protein interface properties can be obtained.[87] The integration of large labels onto a protein may disrupt the protein structure. The slow labeling time scales can also allow more time for such disruptions to occur.

Another technique to covalently label protein side chains is fast photochemical oxidation of proteins (FPOP).[88-89] One particular example is hydroxyl radical footprinting. This method pairs a rapid generation of hydroxyl radicals in solution from hydrogen peroxide using a pulsed KrF laser shot (λ ~248 nm) with a dissolved hydroxyl scavenger.[88-89] The concentration of hydroxyl radical falls by several orders of magnitude in under 1 μs in the presence of excess

glutamine, a hydroxyl scavenger.[88] The short lifetime of hydroxyl radical limits the oxidation to surface exposed side chains. The depletion rate of the hydroxyl radical is faster than most protein unfolding events, so the structure disturbing result of the oxidation itself is not measured.[90] However, any secondary labeling pulses would capture structural damage caused by the sterically bulky label. The label can be quantified with mass spectrometry, as the hydroxyl label itself increases the mass of the protein primarily in increments of +16 Da. Amino acids will have different reaction rates with hydroxyl radicals. Many amino acids that are enriched in IDPs relative to ordered proteins, including A, G, Q, S, and E, react more slowly with hydroxyl radicals relative to other amino acids.[4, 46, 91] Amino acids predicted to be depleted in ordered proteins, including W, F, C, and I, generally possess faster hydroxyl radical reaction rate constants.[46, 91] The hydroxyl radical footprinting method might not be ideal for IDPs given a rapid labeling pulse and generally slower amino acid reaction rates.

The FPOP technique has also been adapted to use other radicals to covalently label proteins, such as sulfate anion and iodine radicals.[92-93] Because of the microsecond labeling times provided by FPOP, other analytical techniques may be more suited to monitor dynamic events that occur on the millisecond or slower time scale, such as the transient, localized unfolding of protein regions.[94] One goal in designing a method to label protein in a manner that can be detected by MS should be to make the label itself small and similar in its chemical properties to what it modifies or replaces on the protein molecule. It is this goal that has made the hydrogen/deuterium exchange technique appealing in protein conformation studies.

1.3.4 H/D Exchange

Ideally, covalent label integration can be both initiated and halted with a great deal of control and avoid potentially structure-disrupting factors such as a high energy photochemical pulse, heating, or a sterically bulky label. A growing number of researchers are turning to the method of amide hydrogen/deuterium (H/D) exchange to study conformational dynamics and properties of complex membrane proteins, protein pharmaceuticals, intrinsically disordered

proteins, and proteins simply not amenable to other methods.[30, 94-95] Hydrogen atoms bonded to oxygen, nitrogen, or sulfur in a protein molecule are labile. By immersing a protein in a high D_2O content solution, deuterium on-exchange will progress as these labile hydrogens are replaced with deuterons on the protein molecule.[94] Aliphatic hydrogens do not exchange in H/D exchange experiments. The rate of H/D exchange will depend on the hydrogen environment, specifically the hydrogen's involvement in conformation-stabilizing hydrogen bonding, the fluctuation rate of such bonds, and neighboring amino acid residues.[94, 96-97] Deuterium exchange by most side-chains (histidine being one exception) and N/C-terminal hydrogens will be rapid. These deuterium labels will rapidly back exchange for protons during the exchange quenching and chromatography steps in a typical H/D experiment (see sections *1.4.3* and *1.4.4*). Amide hydrogen exchange will occur more frequently on a time scale that can be conveniently measured in a laboratory. The substitution of deuterium for hydrogen is much less likely to cause any conformational perturbations than the addition of a sterically large, polar label such as a hydroxyl group. The research presented in the following chapters of this dissertation utilizes H/D exchange to study proteins. Thus, a more in depth overview of H/D exchange mechanisms and methods are presented in sections *1.4* and *1.5*.

1.4 Overview of Hydrogen/Deuterium Exchange

1.4.1 Linderstrøm-Lang Scheme and Intrinsic Exchange

The method of H/D exchange for amide hydrogen was first described by Hvidt and Linderstrøm-Lang in 1954.[98] The mechanism of H/D exchange, usually described by the Linderstrøm-Lang model developed in 1966, is shown in equation 1:[94, 99-100]

$$\mathrm{NH}_{\mathrm{closed}} \underset{k_{\mathrm{cl}}}{\overset{k_{\mathrm{op}}}{\rightleftharpoons}} \mathrm{NH}_{\mathrm{open}} \xrightarrow[\mathrm{D_2O}]{k_{\mathrm{int}}} \mathrm{ND}_{\mathrm{open}} \underset{k_{\mathrm{op}}}{\overset{k_{\mathrm{cl}}}{\rightleftharpoons}} \mathrm{ND}_{\mathrm{closed}} \qquad (1)$$

Each amide hydrogen is in dynamic equilibrium between a "closed" (exchange incompetent) state and an "open" (exchange competent) state. The transition between these two states may rely on local fluctuations, slight molecular "breathing" motions, or more global unfolding events.

The opening rate, k_{op}, will determine the frequency at which a residue will enter the exchange competent state before reforming the closed conformation, with the closing rate, k_{cl}. These dynamic events result in the temporary breaking of H-bonds, which exposes the amide hydrogens to exchange with hydrogen or deuterium from the solvent at a rate defined by the intrinsic rate of exchange, k_{int}.[101-102] An alternative theory includes the degree of solvent penetration and exposure among factors affecting H/D exchange. The use of excess D_2O ensures that the exchange goes only toward the right.

The intrinsic rate constant is defined by how quickly an amino acid residue from a random coil will exchange. The residue-specific k_{int} value depends primarily on the identity of the amino acid residue, its neighboring residues, pH, and temperature.[96, 103] Figure 1.3 illustrates the relationship of pH and temperature to forward exchange k_{int} for a mid-sequence alanine residue. The forward exchange process (*i.e.* H→D) itself is both acid and base catalyzed. Acid catalysis is dominant below about pH 3.5 and base catalysis is dominant above pH 3.5 (see figure 1.3).[94] In the base-catalysis dominated pH region above about pH 5, the rate of exchange increases by one order of magnitude for each pH unit increase. In the acid-catalysis dominated pH region below pH 2, the rate of exchange increases by one order of magnitude for each pH unit decrease. A pH-independent water catalysis is more significant around the pH minimum.[96] The forward exchange may be slowed with colder temperatures, as well.

When an amide enters an exchange competent, open state and H/D exchange occurs before the conformation closes, the amide becomes labeled. The observed rate of H/D exchange for each residue, k_{ex}, is defined in equation 2:[104]

$$k_{ex} = \frac{k_{int} k_{op}}{k_{cl} + k_{int} + k_{op}} \quad (2)$$

The approximation shown in equation 3 can be made if the amide is assumed to be mostly in the closed state, *i.e.* $k_{op} << k_{cl}$:[104]

$$k_{ex} \approx \frac{k_{int} k_{op}}{k_{cl} + k_{int}} \tag{3}$$

The approximation of $k_{op} << k_{cl}$ almost always applies, but IDPs will have very fast k_{op}. Chapter 3 discusses how exchange approximations are adjusted for IDPs. The kinetics of exchange can fall under two main regimes, EX1 and EX2, depending on the conditions of exchange. Under EX1 conditions, the structure closing step proceeds slower than the rate of exchange, $k_{cl} < k_{int}$, so exchange is limited by k_{op}:[105]

$$k_{ex,EX1} = k_{op} \tag{4}$$

In other words, observed exchange with EX1 conditions will reveal the kinetics of an unfolding event, as unfolded protein will immediately exchange. The ratio of unlabeled to labeled protein will lessen as exchange progresses. Characteristic EX1 exchange mass spectra will show a simultaneous disappearance of the unlabeled isotopic envelope and growth of the labeled isotopic envelope as a function of labeling time.[104] Under EX2 conditions, $k_{cl} >> k_{int}$ and exchange proceeds as described in equation 5:[105]

$$k_{ex,EX2} = \frac{k_{op}}{k_{cl}} k_{int} \tag{5}$$

Exchange under EX2 conditions may take multiple opening and closing events before exchange actually occurs per exchange site. Measured kinetics of exchange under EX2 conditions can provide information about the thermodynamic stability, $\Delta G_u^{o'}$, of a protein.[104]

$$\Delta G_u^{o'} = -RT \ln\left(\frac{k_{op}}{k_{cl}}\right) = -RT \ln\left(\frac{k_{ex,EX2}}{k_{int}}\right) \tag{6}$$

Characteristic EX2 exchange mass spectra will show an increasing mass shift in the isotopic envelope as a function of labeling time.[104]

1.4.2 Sample Preparation and Labeling

H/D exchange is a useful technique that has been developed for its capabilities in capturing protein conformation and conformational dynamics information.[74, 104, 106-107] There are two types of H/D labeling: pulsed and continuous. In pulsed labeling H/D exchange, a short D_2O labeling pulse is applied after measured time intervals. The progression of conformational transitions can then be analyzed at each time interval. For example, a denatured protein can be introduced to a solution environment that will promote folding over time. The length of the D_2O pulse is kept constant. Information about protein intermediate states is what is captured by the pulse.[94] SUPREX (stability of unpurified proteins from rates of H/D exchange) represents another pulsed labeling technique. In SUPREX, the pulse label is applied to a series of samples that contain increasing amounts of denaturant, such as urea or guanidinium chloride (GdmCl).[108] Using SUPREX, the conformational stability of proteins in the presence and absence of ligands can be measured.[108] Another pulsed labeling technique is PLIMSTEX (protein-ligand interaction by mass spectrometry, titration, and H/D exchange). In a typical PLIMSTEX procedure, a D_2O pulse is applied to samples with varied ligand concentrations.[108] The duration of the pulse may be chosen from a continuous labeling experiment, wherein the difference between bound and unbound protein deuterium uptake is at a maximum.[108] In a typical continuous labeling H/D exchange experiment, deuterium exchange kinetics of a protein or peptide are measured as a function of D_2O exposure time. Proteins or peptides that exhibit slow exchange will require longer exposure times to reach a deuteration maximum. Exchange is slowed as a function of the stability of intramolecular amide hydrogen bonds and the subsequently slower transient bond disruption rates.[101] Faster exchanging regions will reach deuteration maxima relatively quickly. Such rapid exchange is indicative of highly flexible or unstructured regions.

1.4.3 Quenching

Following a set amount of time of D_2O exposure, a quenching step is used to halt the forward exchange reaction. This quenching step typically includes a decrease in pH to about 2.5 and a decrease of the solution temperature. Figure 1.3 illustrates the effects of reduced temperature and pH on the intrinsic exchange rate constant. The quenching step slows intrinsic exchange kinetics by several orders of magnitude and reduces back exchange.[96, 103] Typical labeling conditions and quenched conditions are indicated by black circles in figure 1.3. The curved arrow represents the several order of magnitude slowing of k_{int} in the quenching step. The specific kinetic parameters of the back exchange process (*i.e.* D→H) differ slightly from those of the forward exchange process. These differences result from D_2O having a different autoprotolysis constant than H_2O as well as a difference in the strengths of N-H and N-D bonds.[96] A consequence is the slight adjustment to the pH of minimum exchange. The exchange minimum saddle region is formed by the combined effects of acid-catalysis, base-catalysis, and water-catalysis on exchange. While acidic conditions may denature protein conformation, exchange kinetics are slowed so drastically that deuterium labels are well preserved.

1.4.4 Factors Affecting Loss of Deuterium Label

The final deuterium uptake measurement obtained by mass spectrometry will depend not only on the forward exchange, but also on the additional effects of back exchange.[109] The quenching step and subsequent chromatography steps add H_2O based solutions to the H/D reaction mixture. The added H_2O introduces new equilibrium conditions that result in a loss of deuterium label from the protein or peptides. This label loss is known as back exchange.[109-110] The effects of chromatography temperature, column type, and gradient duration all play critical roles in the controlling of back exchange.[110-111] In studies of proteins under native conditions, deuterated buffers are often prepared to approximate physiological pH and ionic strength. The specific conditions under which different laboratories conduct H/D labeling are highly variable,[112]

including differences in labeling temperature,[113] pH,[96, 103] buffer compositions,[96] deuterium-to-sample ratios, and solution additives such as denaturants.[114] While the purpose of such experiments may be to study the effects these conditions have on protein conformation, it is also important to consider the loss in inter-laboratory comparability due to these conditions' effects on the intrinsic rate of exchange.[96, 103] The different experimental conditions and equipment used in different laboratories will lead to variation in the magnitude of back exchange measured between laboratories.

1.4.5 Steps to Reduce Back Exchange

One continuing goal in H/D exchange experiments is the reduction of back exchange effects.[115] This is often accomplished with the use of refrigeration or ice baths, which help reduce temperature during chromatography. A reduction in pH, both of the quenched sample and of mobile phase, also slows back exchange.[110, 112, 115] Acid additives in the mobile phase aid in the electrospray ionization process (see section *1.6.1*) as well as in the slowing of back-exchange. The type of acid used also plays an important role in deuterium recovery as well as signal intensity. In an H/D exchange study of bovine carbonic anhydrase, formic acid provided better deuterium recovery and electrospray ionization signal intensity than acetic acid.[112] In the works presented in this dissertation, 0.1% formic acid was used as the mobile phase additive in most chromatography methods. The volatility of formic acid gives it an advantage over hydrochloric acid, as HCl may facilitate salt buildup in the electrospray source. Formic acid is also safer to handle and easier to dispose of than trifluoroacetic acid (TFA), another common acid additive in electrospray experiments. TFA also causes ion suppression effects in electrospray ionization, but may improve chromatographic peak shapes.[116]

1.4.6 Back Exchange Controls

Specific control samples can be utilized to assess an amount of back exchange in H/D-MS methods. A common method of performing this assessment involves simply labeling protein or test peptides with D_2O for a much longer time than typical H/D experiment samples.[112] Long

labeling times may also be paired with a denaturant and/or higher temperatures.[117-119] The difference between the average mass of a fully deuterated control and the average mass of an undeuterated control, Δm_{TD}, provides a maximum experimental mass increase. Deuterium recovery, $\%D_{rec}$, can then be defined as the percent quotient of this maximum mass increase divided by the maximum number of exchangeable amide hydrogens, D_{max} (see section *1.5.2* for how D_{max} is determined):

$$\%D_{rec} = \frac{\Delta m_{TD}}{D_{max}} \times 100\% \quad (7)$$

Mean deuterium recovery of around 80%, a value achieved using a refrigerated HPLC system as explained in chapter 2, should be considered excellent for bottom up H/D-MS. Using residue specific properties, some researchers may use theoretical calculations to model back exchange.[119-120]

1.5 H/D Exchange Mass Spectrometry (H/D-MS)

1.5.1 Top-down H/D-MS

H/D exchange mass spectrometry (H/D-MS) represents the combination of an easily applied deuterium label and mass spectrometry instrumentation in order to capture and analyze conformational properties of proteins. As more researchers adopt H/D-MS methods in protein dynamics studies, the method continues to gain optimizations, modifications, and improvements. One common goal is the improvement of sequence spatial resolution to obtain amino acid residue specific uptake kinetics information.

A gas phase fragmentation method to cleave an initially intact protein in a mass spectrometer is referred to as a "top-down" approach.[94] MS/MS techniques such as electron transfer dissociation (ETD) or electron capture dissociation (ECD) are examples of top down approaches. The use of collision induced dissociation (CID) in H/D-MS experiments will likely exhibit intramolecular hydrogen and deuterium migrations. These migrations result in a randomization of deuterium label locations known as scrambling.[121] Scrambling effects are

greatly reduced in optimized ECD and ETD experiments.[122-123] Some top-down H/D-MS methods offer the advantage of bypassing chromatography altogether and using direct infusion into the ionization source to avoid back exchange effects.[124] Experiments using such top-down techniques have even demonstrated spatial resolution as small as one or two residues,[123-124] though sensitivity limitations may exclude such methods from the study of IDPs. Large proteins and multi-protein systems would also be problematic for top-down MS studies, as the number and size of ion fragments will increase. This complicates the ability of MS/MS to isolate precursor ions for further fragmentation, as each analyzed ion requires a narrow *m/z* filter and additional time for fragmentation and analysis. Narrow *m/z* precursor isolation also exacerbates potential hydrogen scrambling effects.[123]

1.5.2 Bottom-up H/D-MS

Many H/D-MS methods include a proteolytic digestion step prior to analysis, which allows information to be localized on a peptide resolved basis. When digesting or chemically cleaving the protein prior to its introduction to the MS, this approach is referred to as a "bottom-up" analysis.[94] Acid-tolerant proteases, such as pepsin and protease type XIII, are required for back exchange minimization because of quench conditions.[125] Improvements to the peptide digestion and chromatographic separation of peptides can lead to better analysis of large proteins.[77, 106] Though peptide resolution may be improved through the use of long gradients, the more time a peptide spends in aqueous mobile phase, the greater its amount of back exchange. This issue of poor resolution in fast gradient separations led to the implementation of ultra performance liquid chromatography (UPLC).[126] In UPLC techniques, small particle size coupled with high pressure improves resolution and reduces gradient times. Other studies have demonstrated a benefit to deuterium recovery through the use of immobilized protease columns for peptide digestions over using solution-phase digests.[106] While some techniques use solution-phase digests to prepare quenched H/D samples, an immobilized protease column

allows rapid, on-line digestion of samples into peptide fragments prior to a chromatographic separation.[106]

Throughout bottom-up H/D-MS experiments, the number of exchangeable amide hydrogens, D_{max}, is limited because the first residue does not have a peptide bond and the second residue undergoes rapid back exchange.[96, 103] This limitation is accounted for in each chapter of this dissertation, with peptide protection maps in chapters 3 and 4 showing the first two residues grayed out. This property also limits H/D-MS analysis to peptides of three amino acids in length or more. Also excluded from contributing to exchangeable amide hydrogen counts are proline residues as proline lacks an amide hydrogen. Determining the maximum number of exchangeable hydrogens is important in discerning if any data is nonsensical (*i.e.* having significantly greater exchange than the maximum) and providing a theoretical maximum if totally deuterated controls were not used or not detected. This dissertation outlines the application of bottom-up H/D-MS to study IDPs. Improvements or optimizations made to the bottom-up H/D-MS technique as a whole, including considerations in proteolysis, HPLC, and data analysis, are also discussed.

1.6 Mass Spectrometry Instrumentation

1.6.1 Electrospray Ionization

Electrospray ionization (ESI) is often applied in H/D-MS, as ESI allows soft ionization of liquid-phase samples.[127] ESI can be used to interface HPLC to a mass spectrometer, with eluent being simultaneously evaporated and ionized. To summarize the process for positive mode ESI briefly;[127] a strong electric field is generated by holding an introduction capillary tip at a high potential relative to the counter-electrode walls of the source chamber. This electric field causes dispersion of eluent into a mist of charged droplets. A heated drying gas will assist in evaporating the solvent from droplets, causing droplet shrinkage. Nitrogen is often used as an economical drying gas. As the droplets become smaller, repulsive forces between charged particles (*i.e.* protons in the initially acidic sample) at the surface of the droplet overcome the

surface tension. This causes the droplets to undergo fission to form smaller droplets. Eventually, solvent evaporation will lead to biomolecules, such as proteins or peptides, gaining a positive charge as they enter the gas phase. The specific mechanism of desolvation and charge acquisition by biomolecules is still under debate.[128-130] Following desolvation, the positively charges ions must be carried to a detector. Ions are guided through a series of filtering stages, such as transfer capillaries, skimmer cones, and multipoles, where vacuum pumps reduce the pressure at various stages. Many parameters regarding electric field strength in the ESI source are controllable. This is beneficial in studying large proteins, as the multiple charging of single molecules produces mass-to-charge ratio (*m/z*) that can be detected by mass analyzers with limited detection ranges.[127]

1.6.2 Time-of-Flight Mass Spectrometer

The choice of mass spectrometer in H/D-MS will vary from one laboratory to another. The MS experiments presented in this dissertation were all performed using an Agilent Technologies Accurate Mass 6220 TOF LC/MS. This instrument is an orthogonal time-of-flight mass spectrometer (TOF-MS).[131] Filtered ions generated by the ESI source are guided into an ion pulser. The ion pulser applies a high voltage pulse to accelerate the ions orthogonally into a flight tube. The pulse gives all ions the same amount of kinetic energy to travel through the flight tube. The flight tube contains an ion mirror, which directs the ions onto a detector.[131] The amount of time it takes an ion to reach the detector will be determined by its mass-to-charge ratio, *m/z*.[127] In a simple linear flight tube, the velocity, υ, of an ion is inversely related to the *m/z* of the ion and the applied ion acceleration potential, *V*:[127]

$$\upsilon = \sqrt{\frac{2zV}{m}} \tag{8}$$

Thus, the time, *t*, required for an ion to travel the length, *L*, of the flight tube to reach a detector is given by equation 9:[127]

$$t = \frac{L}{\upsilon} = L\sqrt{\frac{1}{2V}\left(\frac{m}{z}\right)} \quad (9)$$

Thus, a detector that is able to measure the flight time of ion packets can be used to draw an *m/z* spectrum. Time-of-flight MS instruments exhibit high sensitivity, allowing small amounts of proteins or peptides to be more readily detected in H/D-MS.[127] High mass accuracy and *m/z* peak resolution provided by using TOF-MS are also important in peptide sequence assignments.[127]

1.6.3 CID / MS^E and Peptide Fragmentation

Another capability of some MS instrument is nozzle-skimmer dissociation (NSD), a variant of collision-induced dissociation (CID). This technique allows the formation of peptide ion fragments without the need for an MS/MS instrument, however the single MS instrumentation used does not provide any selectivity in choosing precursor or product ions as a typical tandem MS experiment could.[132] (A similar technique in which a collision cell is present but precursor isolation is not used is also possible.[133]) Despite this limitation, this technique is useful if peptides are initially resolved chromatographically; a chromatogram may be rebuilt in which product ions are matched with precursors. This is necessary to assign sequences to peptides that would otherwise have ambiguous assignments with single MS analysis. While tandem MS techniques are more commonly used in sequence assignments, an NSD technique may provide an expedient alternative. This technique will fail if peptides are not chromatographically resolved efficiently enough. While single MS spectra may resolve co-eluting peptides from each other, the spectral features for product ions resulting from the simultaneous fragmentation of multiple peptides may be too highly overlapped for proper resolution. This technique may also produce insufficient signal intensity, as fragmentation is likely to greatly reduce peak intensity of product ions relative to the precursor ion. The application of such an NSD technique is described in chapter 3 of this work, wherein ion

fragmentor voltage was rapidly cycled to produce three simultaneous ion chromatograms. Low voltage scans measured unfragmented peptide *m/z* values while high voltage scans revealed b/y ion fragments, allowing specific peptide identities to be assigned.

1.7 Data Analysis

1.7.1 Mass Spectra Workflow

This section will briefly describe a general procedure for extracting deuterium uptake data from raw mass spectral data. The first step is to build a peptide database using the spectra obtained from undeuterated control samples. MS software is used to find peptide features within the total ion chromatogram (TIC, figure 1.4A) and isolate each feature's specific monoisotopic *m/z* values in order to build a set of extracted ion chromatograms (EIC, figure 1.4B) and the integrated mass spectra (figure 1.4C). Peptide mass is determined by multiplying the monoisotopic *m/z* value and the charge, *z*, then subtracting the mass of the charge carriers. The value of *z* is indicated by the peptide's isotopic peak spacing. When a protein sequence and predicted proteolysis cleavages are entered, the software can assign peptide identities to each matching EIC based on each peptide's predicted mass. This process is known as peptide mapping. After ambiguous assignments are clarified (such as with NSD, as described in section *1.6.3*), a peptide database can be made that contains information such as retention time, sequence numbering, sequence composition, formula, and predicted mass. The software identifies mass spectral features of peptides based on the expected distribution of isotopes within a peptide population, which is dominated by ^{13}C distribution. When a peptide is deuterated, the isotopic distribution becomes representative of the number of integrated deuterium isotopes, ^{2}H, as well. Figure 1.4D shows the isotopic distribution for a peptide's undeuterated mass spectrum (top) and deuterated mass spectrum (bottom). Because of the atypical distribution of *m/z* values for deuterated peptides, the peptide database is expanded to account for deuteron substitutions. Because the monoisotopic *m/z* peak of a deuterated peptide can be both difficult to determine and of very low relative intensity compared to other *m/z* peaks,

average *m/z* is used instead of monoisotopic *m/z*. The average *m/z* of a peptide is defined as the center of mass of the isotopic distribution. The deuterated peptide database provides the MS software with specific *m/z* values and retention times to search for in the deuterated sample TIC. The software then extracts spectra of deuterated peptides and converts average *m/z* values to average mass values. The difference between average deuterated mass, m_{deut}, and the average mass of undeuterated control, m_{UD}, represents the mass increase, Δm:

$$\Delta m = m_{deut} - m_{UD} \tag{10}$$

Mass increase is displayed by plotting Δm versus D_2O exposure time for a particular peptide or protein (figure 1.4E). Δm is used instead of counting deuterons, #D, to avoid any discrepancy between the integer counting of deuterons and actual mass increase. An addition of a deuteron results in a mass increase slightly higher than 1 Da. Back-exchange effects will also obscure any counting of individual deuterons. Figure 1.4E shows deuterium uptake curves from a protein with and without a mutation. H/D exchange for multiple protein conditions can be compared on a single graph. The mutation causes faster H/D exchange in this peptide. Thus, the mutation condition causes the region to lose conformational stability when compared to wild-type, WT. In this dissertation, some differential conditions explored include mutant versus wild-type, presence of a small molecule to restore function of a mutant, and unbound versus bound. Other examples include exploring the effects of denaturants,[134] chromatography pH or temperature change,[112] post-translational modifications,[135] buffer condition changes,[134] and more.

1.7.2 Quantitative Analysis in H/D-MS

One of the goals of this dissertation, and in the field of H/D-MS methodology in general, is to improve quantitation and comparability of H/D exchange data. Comparisons in H/D exchange data analysis are most often subjective, such as the qualitative comparison made in section *1.7.1*. With deuterium uptake curves plotted in a graphical format, such as in figure

1.4E, curve fitting functions may be calculated using software such as SigmaPlot (Systat Software Inc., San Jose, CA). Rate constants derived from curve fitting allow quantitative comparisons between uptake curves, as discussed in chapters 3 and 4. Other comparison methods are discussed in chapter 5. One of the more common methods for comparing deuterium exchange in NMR is the use of protection factors, *PF*.[101-102, 118] Protection factors are calculated as the ratio between the intrinsic H/D exchange rate constant,[96] k_{int}, and the experimentally measured rate constant, k_{ex}:

$$PF = \frac{k_{\text{int}}}{k_{\text{ex}}} \tag{11}$$

Describing deuterium uptake behavior in terms of protection factors manner is ubiquitous in NMR experiments,[101-102] but still excluded from MS methods without single-residue resolution. Protection factors supply information about the rates of opening and closing of structural conformations in the Linderstrøm-Lang model,[98] presented in equation 1. These opening and closing events limit amide hydrogen exposure to H/D exchange, as described by equations 2-5.

Protection factor determination, however, provides little information about the identity of structured elements, by itself. A specific type of structural feature, such as α-helices, can exhibit wide ranges of *PF*. These *PF* ranges will often overlap between the different features, as well. A specific α-helix may share a similar *PF* value as a β-sheet elsewhere. Recent work has made progress toward relating exchange protection factors to structural elements with residue-specific resolution.[101] The theory that H-bonding properties limits H/D exchange was supported by an NMR analysis of mutated *Staphylococcal* nuclease. The type of H-bond acceptor affects how protected an amide hydrogen will be from exchange, with incorporated water, side chain, and backbone carbonyl H-bond acceptors causing protection factors of almost 10^8. However, protection factor ranges (such as between 10^2 and 10^8 for backbone bonded amide hydrogen)

are still too broad to distinguish one structural element type from another, as both α-helical and β-sheet incorporated amide hydrogens exhibit such ranges in protection.[101]

1.8 Motivation of Dissertation

As it becomes more apparent that IDPs fulfill vital roles in cells, there is a necessity to adapt current methods and develop new ones to study conformational dynamics of IDPs. The purpose of this dissertation research is to demonstrate the application of the H/D-MS method toward better understanding the conformational behavior of IDPs. This research was also conducted to elucidate synergistic folding effects in a pair of model IDPs. In the course of conducting this research, improvements to H/D-MS methods were also made. The following chapters discuss gains not only from a purely scientific standpoint, but also the practical benefits this research offers as H/D-MS becomes more important in applications requiring routine analyses.

Many researchers use bulky and messy ice/water baths or high pressure columns to prevent back exchange and maintain cold temperatures in H/D experiments. The construction of a thermoelectrically refrigerated system for HPLC components is discussed in chapter 2. The inexpensive refrigerated system was constructed to ensure that solvent, enclosure, and column oven temperatures could be controlled to within ±0.1 °C. Such tight temperature control allows back exchange to be both minimal and highly reproducible. Proteolytic digestion efficiency was also improved. The refrigerated system is a solution for online digestion and reverse-phase separation in a bottom-up H/D-MS method.

The synergistic folding in IDP interactions is explored in chapter 3. CBP and ACTR represent two classifications of IDPs in their native, unbound states: the molten globule and random coil, respectively. Together, they form a stable complex that resists H/D exchange. CBP and ACTR were chosen as model systems for the development of H/D-MS to study IDPs. The H/D-MS analysis of CBP and ACTR are compared with NMR structural analyses of these proteins in order to validate the usefulness of H/D-MS to study multiple states of IDPs. This

comparison reveals that unbound CBP exchanges rapidly, despite evidence of a stable structure revealed by NMR. Expression, purification, assay, and preservation procedures are also included to explain how to separate and store IDPs that share similar properties. IDPs such as these may require special handling and preparation procedures, as they may be prone to aggregation or post-translational modification.

As shown in chapter 3, H/D exchange for IDPs, such as ACTR, proceeded too quickly for accurate analysis. That is, many unbound ACTR peptides exhibited full deuteration near or before the minimum D_2O exposure time of 5 seconds. To extend the scope of H/D-MS measurements introduced in chapter 3, chapter 4 introduces a quench flow apparatus for H/D exchange. By extending the time scale from the lower limits imposed by pipetting (~5 seconds) to 42 milliseconds, more complete deuterium uptake curves can be measured. This provides more accurate kinetic parameters for deuterium uptake. The quench flow apparatus can be re-assembled using inexpensive, off-the-shelf parts. This chapter discusses the correlation of subtle exchange protection in specific regions of CBP and ACTR to residual structure measured by NMR.

The applications of H/D-MS are many, as is the range of instrumentation available to conduct such experiments. However, there has yet to be established a standard method by which H/D exchange information from a differential condition study (illustrated by figure 1.4E) may be compared. Typically, data is presented on a peptide level basis, without accounting for peptide length, relative position along an amino acid sequence, or fractional deuterium uptake. Chapter 5 proposes an averaging of overlapping deuterium uptake to better resolve amino acid residue specific information. The extensions discussed in chapter 5 include a normalization of deuterium uptake to quantify exchange as fractional.

References

1. Fischer, E., Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **1894,** *27* (3), 2985-2993.
2. Lemieux, R. U.; Spohr, U., How Emil Fischer Was Led To The Lock and Key Concept for Enzyme Specificity. In *Adv. Carbohydr. Chem. Biochem.* Academic Press: **1994**; Vol. Volume 50, pp 1-20.
3. Huber, R.; Bennett, W. S., Functional significance of flexibility in proteins. *Biopolymers* **1983,** *22* (1), 261-279.
4. Romero, P.; Obradovic, Z.; Li, X.; Garner, E. C.; Brown, C. J.; Dunker, A. K., Sequence complexity of disordered protein. *Proteins* **2001,** *42* (1), 38-48.
5. Dyson, H. J.; Wright, P. E., Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005,** *6* (3), 197-208.
6. Liu, J.; Perumal, N. B.; Oldfield, C. J.; Su, E. W.; Uversky, V. N.; Dunker, A. K., Intrinsic Disorder in Transcription Factors. *Biochemistry* **2006,** *45* (22), 6873-6888.
7. Receveur-Bréchot, V.; Bourhis, J.-M.; Uversky, V. N.; Canard, B.; Longhi, S., Assessing protein disorder and induced folding. *Proteins* **2006,** *62* (1), 24-45.
8. Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N., Functional Anthology of Intrinsic Disorder. 3. Ligands, Post-Translational Modifications, and Diseases Associated with Intrinsically Disordered Proteins. *Journal of Proteome Research* **2007,** *6* (5), 1917-1932.
9. Dunker, A. K.; Silman, I.; Uversky, V. N.; Sussman, J. L., Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008,** *18* (6), 756-764.
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. Bae, S.-H.; Dyson, H. J.; Wright, P. E., Prediction of the Rotational Tumbling Time for Proteins with Disordered Segments. *J. Am. Chem. Soc.* **2009,** *131* (19), 6814-6821.
12. Uversky, V. N.; Dunker, A. K., Understanding protein non-folding. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **2010,** *1804* (6), 1231-1264.
13. Zhou, H.-X.; Pang, X.; Lu, C., Rate constants and mechanisms of intrinsically disordered proteins binding to structured targets. *PCCP* **2012,** *14* (30), 10466-10476.
14. Tantos, A.; Han, K.-H.; Tompa, P., Intrinsic disorder in cell signaling and gene transcription. *Molecular and Cellular Endocrinology* **2012,** *348* (2), 457-465.
15. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z., Intrinsically disordered protein. *J. Mol. Graphics Modell.* **2001,** *19* (1), 26-59.
16. Hilser, V. J.; García-Moreno E, B.; Oas, T. G.; Kapp, G.; Whitten, S. T., A Statistical Thermodynamic Model of the Protein Ensemble. *Chem. Rev.* **2006,** *106* (5), 1545-1558.
17. Baum, J.; Dobson, C. M.; Evans, P. A.; Hanley, C., Characterization of a Partly Folded Protein by NMR Methods: Studies on the Molten Globule State of Guinea Pig .alpha.-Lactalbumin. *Biochemistry* **1989,** *28* (1), 7-13.

18. Hughson, F. M.; Wright, P. E.; Baldwin, R. L., Structural Characterization of a Partly Folded Apomyoglobin Intermediate. *Science* **1990,** *249* (4976), 1544-1548.

19. Demarest, S. J.; Martinez-Yamout, M.; Chung, J.; Chen, H.; Xu, W.; Dyson, H. J.; Evans, R. M.; Wright, P. E., Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **2002,** *415* (6871), 549-553.

20. Demarest, S. J.; Deechongkit, S.; Dyson, H. J.; Evans, R. M.; Wright, P. E., Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci.* **2004,** *13* (1), 203-210.

21. Ebert, M.-O.; Bae, S.-H.; Dyson, H. J.; Wright, P. E., NMR Relaxation Study of the Complex Formed Between CBP and the Activation Domain of the Nuclear Hormone Receptor Coactivator ACTR *Biochemistry* **2008,** *47* (5), 1299-1308.

22. Redfield, C., Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods* **2004,** *34* (1), 121-132.

23. Uversky, V. N.; Ptitsyn, O. B., "Partly Folded" State, a New Equilibrium State of Protein Molecules: Four-State Guanidinium Chloride-Induced Unfolding of beta-Lactamase at Low Temperature. *Biochemistry* **1994,** *33* (10), 2782-2791.

24. Uversky, V. N.; Ptitsyn, O. B., Further Evidence on the Equilibrium “Pre-molten Globule State”: Four-state Guanidinium Chloride-induced Unfolding of Carbonic Anhydrase B at Low Temperature. *J. Mol. Biol.* **1996,** *255* (1), 215-228.

25. Huang, Y.; Liu, Z., Kinetic Advantage of Intrinsically Disordered Proteins in Coupled Folding-Binding Process: A Critical Assessment of the "Fly-Casting" Mechanism. *J. Mol. Biol.* **2009,** *393* (5), 1143-1159.

26. Shortle, D.; Ackerman, M. S., Persistence of Native-Like Topology in a Denatured Protein in 8 M Urea. *Science* **2001,** *293* (5529), 487-489.

27. Ward, J. J.; Sodhi, J. S.; McGuffin, L. J.; Buxton, B. F.; Jones, D. T., Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **2004,** *337* (3), 635-645.

28. Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N.; Obradovic, Z., Functional anthology of intrinsic disorder. 1. biological processes and functions of proteins with long disordered regions. *Journal of Proteome Research* **2007,** *6* (5), 1882-1898.

29. Vucetic, S.; Xie, H.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Obradovic, Z.; Uversky, V. N., Functional Anthology of Intrinsic Disorder. 2. Cellular Components, Domains, Technical Terms, Developmental Processes, and Coding Sequence Diversities Correlated with Long Disordered Regions. *Journal of Proteome Research* **2007,** *6* (5), 1899-1916.

30. Balasubramaniam, D.; Komives, E. A., Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2012,** *in Press, uncorrected proof*.

31. Oldfield, C. J.; Cheng, Y.; Cortese, M. S.; Brown, C. J.; Uversky, V. N.; Dunker, A. K., Comparing and Combining Predictors of Mostly Disordered Proteins. *Biochemistry* **2005,** *44* (6), 1989-2000.

32. Radivojac, P.; Iakoucheva, L. M.; Oldfield, C. J.; Obradovic, Z.; Uversky, V. N.; Dunker, A. K., Intrinsic Disorder and Functional Proteomics. *Biophys. J.* **2007,** *92* (5), 1439-1456.

33. Kovacs, D.; Szabo, B.; Pancsa, R.; Tompa, P., Intrinsically disordered proteins undergo and assist folding transitions in the proteome. *Arch. Biochem. Biophys.* **2012**, Accepted manuscript.

34. Iakoucheva, L. M.; Brown, C. J.; Lawson, J. D.; Obradovic, Z.; Dunker, A. K., Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **2002,** *323* (3), 573-584.

35. Kriwacki, R. W.; Hengst, L.; Tennant, L.; Reed, S. I.; Wright, P. E., Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93* (21), 11504-11509.

36. Uversky, V. N.; Gillespie, J. R.; Fink, A. L., Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **2000,** *41* (3), 415-427.

37. Zhou, P.; Lugovskoy, A. A.; McCarty, J. S.; Li, P.; Wagner, G., Solution structure of DFF40 and DFF45 N-terminal domain complex and mutual chaperone activity of DFF40 and DFF45. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98* (11), 6051-6055.

38. Lacy, E. R.; Filippov, I.; Lewis, W. S.; Otieno, S.; Xiao, L.; Weiss, S.; Hengst, L.; Kriwacki, R. W., p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol. Biol.* **2004,** *11* (4), 358-364.

39. Rumi-Masante, J.; Rusinga, F. I.; Lester, T. E.; Dunlap, T. B.; Williams, T. D.; Dunker, A. K.; Weis, D. D.; Creamer, T. P., Structural Basis for Activation of Calcineurin by Calmodulin. *J. Mol. Biol.* **2012,** *415* (2), 307-317.

40. Brocca, S.; Šamalíková, M.; Uversky, V. N.; Lotti, M.; Vanoni, M.; Alberghina, L.; Grandori, R., Order propensity of an intrinsically disordered protein, the cyclin-dependent-kinase inhibitor Sic1. *Proteins* **2009,** *76* (3), 731-746.

41. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *Journal of Molecular Recognition* **2005,** *18* (5), 343-384.

42. Oldfield, C.; Meng, J.; Yang, J.; Yang, M. Q.; Uversky, V.; Dunker, A. K., Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics* **2008,** *9* (Suppl 1), S1.

43. Kjaergaard, M.; Teilum, K.; Poulsen, F. M., Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (28), 12535-12540.

44. Chen, H.; Lin, R. J.; Schiltz, R. L.; Chakravarti, D.; Nash, A.; Nagy, L.; Privalsky, M. L.; Nakatani, Y.; Evans, R. M., Nuclear Receptor Coactivator ACTR Is a Novel Histone Acetyltransferase and Forms a Multimeric Activation Complex with P/CAF and CBP/p300. *Cell* **1997,** *90* (3), 569-580.

45. Romero, P.; Obradovic, Z.; Kissinger, C.; Villafranca, J. E.; Dunker, A. K., Identifying disordered regions in proteins from amino acid sequence. **1997,** *1*, 90-95 vol.1.

46. Williams, R. M.; Obradović, Z.; Mathura, V.; Braun, W.; Garner, E. C.; J.Young; Takayama, S.; Brown, C. J.; Dunker, A. K., The Protein Non-Folding Problem: Amino Acid Determinants of Intrinsic Order and Disorder. *Pacific Symposium on Biocomputing* **2001,** *6*, 89-100.

47. Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C. J.; Dunker, A. K.; Obradovic, Z., Optimizing long intrinsic disorder predictors with protein evolutionary information. *Journal of Bioinformatics and Computational Biology* **2005,** *03* (01), 35-60.

48. Romero, P.; Obradovic, Z.; Dunker, A. K., Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. *Genome Informatics* **1997,** *8*, 110-124.

49. Li, X.; Romero, P.; Rani, M.; Dunker, A. K.; Obradovic, Z., Predicting protein disorder for N-, C-, and internal regions. *Genome Informatics* **1999,** *10*, 30-40.

50. Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C., A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958,** *181* (4610), 662-666.

51. Davis, A. M.; Teague, S. J.; Kleywegt, G. J., Application and Limitations of X-ray Crystallographic Data in Structure-Based Ligand and Drug Design. *Angew. Chem. Int. Ed.* **2003,** *42* (24), 2718-2736.

52. Dyson, H. J.; Wright, P. E., Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev.* **2004,** *104* (8), 3607-3622.

53. Greenfield, N. J.; Fasman, G. D., Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* **1969,** *8* (10), 4108-4116.

54. Adler, A. J.; Greenfield, N. J.; Fasman, G. D., [27] Circular dichroism and optical rotatory dispersion of proteins and polypeptides. In *Methods Enzymol.* Academic Press: **1973**; Vol. Volume 27, pp 675-735.

55. Greenfield, N. J., Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **2006,** *1* (6), 2876-2890.

56. Whitmore, L.; Wallace, B. A., Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **2008,** *89* (5), 392-400.

57. Venyaminov, S. Y.; Baikalov, I. A.; Shen, Z. M.; Wu, C. S. C.; Yang, J. T., Circular Dichroic Analysis of Denatured Proteins: Inclusion of Denatured Proteins in the Reference Set. *Anal. Biochem.* **1993,** *214* (1), 17-24.

58. Chou, D. K.; Krishnamurthy, R.; Manning, M. C.; Randolph, T. W.; Carpenter, J. F., Physical stability of albinterferon-α2b in aqueous solution: Effects of conformational stability and colloidal stability on aggregation. *J. Pharm. Sci.* **2012,** *101* (8), 2702-2719.

59. Hamid Wani, A.; Udgaonkar, J. B., HX-ESI-MS and Optical Studies of the Unfolding of Thioredoxin Indicate Stabilization of a Partially Unfolded, Aggregation-Competent Intermediate at Low pH. *Biochemistry* **2006,** *45* (37), 11226-11238.

60. Maier, C. S.; Schimerlik, M. I.; Deinzer, M. L., Thermal Denaturation of Escherichia coli Thioredoxin Studied by Hydrogen/Deuterium Exchange and Electrospray Ionization Mass Spectrometry: Monitoring a Two-State Protein Unfolding Transition. *Biochemistry* **1999,** *38* (3), 1136-1143.

61. Bertini, I.; McGreevy, K. S.; Parigi, G., NMR and its Place in Mechanistic Systems Biology. In *NMR of Biomolecules*. Wiley-VCH Verlag GmbH & Co. KGaA: **2012**; pp 1-5.

62. Sharma, S.; Zheng, H.; Huang, Y. J.; Ertekin, A.; Hamuro, Y.; Rossi, P.; Tejero, R.; Acton, T. B.; Xiao, R.; Jiang, M.; Zhao, L.; Ma, L.-C.; Swapna, G. V. T.; Aramini, J. M.; Montelione, G. T., Construct optimization for protein NMR structure analysis using amide hydrogen/deuterium exchange mass spectrometry. *Proteins* **2009,** *76* (4), 882-894.

63. Ferella, L.; Rosato, A.; Turano, P., Determination of Protein Structure and Dynamics. In *NMR of Biomolecules*. Wiley-VCH Verlag GmbH & Co. KGaA: **2012**; pp 51-94.

64. Neri, D.; Billeter, M.; Wider, G.; Wüthrich, K., NMR Determination of Residual Structure in a Urea-Denatured Protein, the 434-Repressor. *Science* **1992,** *257* (5076), 1559-1563.

65. Zhang, O.; Forman-Kay, J. D.; Shortle, D.; Kay, L. E., Triple-resonance NOESY-based experiments with improved spectral resolution: Applications to structural characterization of unfolded, partially folded and folded proteins. *J. Biomol. NMR* **1997,** *9* (2), 181-200.

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

67. Kay, L. E., NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* **2011,** *213* (2), 477-491.

68. Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D., Conformation and Dynamics of Biopharmaceuticals: Transition of Mass Spectrometry-Based Tools from Academe to Industry. *J. Am. Soc. Mass. Spectrom.* **2010,** *21* (3), 323-337.

69. Houde, D.; Berkowitz, S. A.; Engen, J. R., The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J. Pharm. Sci.* **2011,** *100* (6), 2071-2086.

70. Bracken, C.; Iakoucheva, L. M.; Romero, P. R.; Dunker, A. K., Combining prediction, computation and experiment for the characterization of protein disorder. *Curr. Opin. Struct. Biol.* **2004,** *14* (5), 570-576.

71. Frimpong, A. K.; Abzalimov, R. R.; Uversky, V. N.; Kaltashov, I. A., Characterization of intrinsically disordered proteins with electrospray ionization mass spectrometry: Conformational heterogeneity of α-synuclein. *Proteins* **2010,** *78* (3), 714-722.

72. šamalíková, M.; Santambrogio, C.; Grandori, R., Mass Spectrometry Tools for the Investigation of Structural Disorder and Conformational Transitions in Proteins. In *Instrumental Analysis of Intrinsically Disordered Proteins*. John Wiley & Sons, Inc.: **2010**; pp 627-652.

73. Testa, L.; Brocca, S.; Šamalikova, M.; Santambrogio, C.; Alberghina, L.; Grandori, R., Electrospray ionization-mass spectrometry conformational analysis of isolated domains of an intrinsically disordered protein. *Biotechnology Journal* **2011,** *6* (1), 96-100.

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

75. Bondarenko, P. V.; Second, T. P.; Zabrouskov, V.; Makarov, A. A.; Zhang, Z., Mass Measurement and Top-Down HPLC/MS Analysis of Intact Monoclonal Antibodies on a Hybrid Linear Quadrupole Ion Trap–Orbitrap Mass Spectrometer. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (8), 1415-1424.

76. Du, Y.; Wang, F.; May, K.; Xu, W.; Liu, H., LC–MS analysis of glycopeptides of recombinant monoclonal antibodies by a rapid digestion procedure. *J. Chromatogr. B* **2012,** *907* (0), 87-93.

77. Zhang, Q.; Willison, L. N.; Tripathi, P.; Sathe, S. K.; Roux, K. H.; Emmett, M. R.; Blakney, G. T.; Zhang, H.-M.; Marshall, A. G., Epitope Mapping of a 95 kDa Antigen in Complex with Antibody by Solution-Phase Amide Backbone Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2011,** *83* (18), 7129-7136.

78. Fekete, S.; Gassner, A.-L.; Rudaz, S.; Schappler, J.; Guillarme, D., Analytical strategies for the characterization of therapeutic monoclonal antibodies. *TrAC, Trends Anal. Chem.* (0).

79. Strobel, F. H.; Solouki, T.; White, M. A.; Russell, D. H., Detection of femtomole and sub-femtomole levels of peptides by tandem magnetic sector/reflectron time-of-flight mass spectrometry and matrix-assisted laser desorption ionization. *J. Am. Soc. Mass. Spectrom.* **1991,** *2* (1), 91-94.

80. Heck, A. J. R., Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Meth* **2008,** *5* (11), 927-933.

81. Pacholarz, K. J.; Garlish, R. A.; Taylor, R. J.; Barran, P. E., Mass spectrometry based tools to investigate protein-ligand interactions for drug discovery. *Chem. Soc. Rev.* **2012,** *41* (11), 4335-4355.

82. Yi, S.; Boys, B. L.; Brickenden, A.; Konermann, L.; Choy, W.-Y., Effects of Zinc Binding on the Structure and Dynamics of the Intrinsically Disordered Protein Prothymosin α: Evidence for Metalation as an Entropic Switch. *Biochemistry* **2007,** *46* (45), 13120-13130.

83. Brocca, S.; Testa, L.; Sobott, F.; Šamalikova, M.; Natalello, A.; Papaleo, E.; Lotti, M.; De Gioia, L.; Doglia, Silvia M.; Alberghina, L.; Grandori, R., Compaction Properties of an Intrinsically Disordered Protein: Sic1 and Its Kinase-Inhibitor Domain. *Biophys. J.* **2011,** *100* (9), 2243-2252.

84. Zhang, H.; Wen, J.; Huang, R. Y. C.; Blankenship, R. E.; Gross, M. L., Mass spectrometry-based carboxyl footprinting of proteins: Method evaluation. *Int. J. Mass spectrom.* **2012,** *312* (0), 78-86.

85. Fowler, J. D.; Brown, J. A.; Kvaratskhelia, M.; Suo, Z., Probing Conformational Changes of Human DNA Polymerase λ using Mass Spectrometry-Based Protein Footprinting. *J. Mol. Biol.* **2009,** *390* (3), 368-379.

86. McKee, C. J.; Kessl, J. J.; Norris, J. O.; Shkriabai, N.; Kvaratskhelia, M., Mass spectrometry-based footprinting of protein–protein interactions. *Methods* **2009,** *47* (4), 304-307.

87. Mädler, S.; Seitz, M.; Robinson, J.; Zenobi, R., Does Chemical Cross-Linking with NHS Esters Reflect the Chemical Equilibrium of Protein–Protein Noncovalent Interactions in Solution? *J. Am. Soc. Mass. Spectrom.* **2010,** *21* (10), 1775-1783.

88. Hambly, D.; Gross, M., Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J. Am. Soc. Mass. Spectrom.* **2005,** *16* (12), 2057-2063.

89. Hambly, D.; Gross, M., Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin. *Int. J. Mass spectrom.* **2007,** *259* (1–3), 124-129.

90. Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L., Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* **2009,** *81* (16), 6563-6571.

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

92. Gau, B. C.; Chen, H.; Zhang, Y.; Gross, M. L., Sulfate Radical Anion as a New Reagent for Fast Photochemical Oxidation of Proteins. *Anal. Chem.* **2010,** *82* (18), 7821-7827.

93. Chen, J.; Cui, W.; Giblin, D.; Gross, M., New Protein Footprinting: Fast Photochemical Iodination Combined with Top-Down and Bottom-Up Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2012,** *23* (8), 1306-1318.

94. 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.

95. Iacob, R.; Engen, J., Hydrogen Exchange Mass Spectrometry: Are We Out of the Quicksand? *J. Am. Soc. Mass. Spectrom.* **2012,** *23* (6), 1003-1010.

96. Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W., Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 75-86.

97. Tsutsui, Y.; Wintrode, P. L., Hydrogen/Deuterium Exchange-Mass Spectrometry: A Powerful Tool for Probing Protein Structure, Dynamics, and Interactions. *Curr. Med. Chem.* **2007,** *14*, 2344-2358.

98. Hvidt, A.; Linderstrøm-Lang, K., Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim. Biophys. Acta* **1954,** *14* (0), 574-575.

99. Hvidt, A.; Nielsen, S. O., Hydrogen Exchange in Proteins. In *Adv. Protein Chem.* Academic Press: **1966**; Vol. Volume 21, pp 287-386.

100. Baldwin, R. L., Early days of protein hydrogen exchange: 1954–1972. *Proteins* **2011,** *79* (7), 2021-2026.

101. 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.

102. Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W., Protein hydrogen exchange: Testing current models. *Protein Sci.* **2012,** *21* (7), 987-995.

103. Connelly, G. P.; Bai, Y.; Jeng, M.-F.; Englander, S. W., Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 87-92.

104. Konermann, L.; Tong, X.; Pan, Y., Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **2008,** *43* (8), 1021-1036.

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

106. Wang, L.; Pan, H.; Smith, D. L., Hydrogen Exchange-Mass Spectrometry: Optimization of Digestion Conditions. *Mol. Cell. Proteomics* **2002,** *1* (2), 132-138.

107. Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L., Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100* (12), 7057-7062.

108. Fitzgerald, M. C.; West, G. M., Painting Proteins with Covalent Labels: What's In the Picture? *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (6), 1193-1206.

109. Zhang, Z.; Smith, D. L., Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* **1993,** *2* (4), 522-531.

110. Keppel, T.; Jacques, M.; Young, R.; Ratzlaff, K.; Weis, D., An Efficient and Inexpensive Refrigerated LC System for H/D Exchange Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2011,** *22* (8), 1472-1476.

111. Zhang, H.-M.; Bou-Assaf, G. M.; Emmett, M. R.; Marshall, A. G., Fast Reversed-Phase Liquid Chromatography to Reduce Back Exchange and Increase Throughput in H/D Exchange Monitored by FT-ICR Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (3), 520-524.

112. Burkitt, W.; O'Connor, G., Assessment of the Repeatability and Reproducibility of Hydrogen/Deuterium Exchange Mass Spectrometry Measurements. *Rapid Commun. Mass Spectrom.* **2008,** *22* (23), 3893-3901.

113. Zhang, A.; Qi, W.; Singh, S.; Fernandez, E., A New Approach to Explore the Impact of Freeze-Thaw Cycling on Protein Structure: Hydrogen/Deuterium Exchange Mass Spectrometry (HX-MS). *Pharm. Res.* **2011,** *28* (5), 1179-1193.

114. Mazon, H.; Marcillat, O.; Forest, E.; Vial, C., Denaturant sensitive regions in creatine kinase identified by hydrogen/deuterium exchange. *Rapid Commun. Mass Spectrom.* **2005,** *19* (11), 1461-1468.

115. Walters, B.; Ricciuti, A.; Mayne, L.; Englander, S. W., Minimizing Back Exchange in the Hydrogen Exchange-Mass Spectrometry Experiment. *J. Am. Soc. Mass. Spectrom.* **2012**, 1-8.

116. Shou, W. Z.; Naidong, W., Simple means to alleviate sensitivity loss by trifluoroacetic acid (TFA) mobile phases in the hydrophilic interaction chromatography–electrospray tandem mass spectrometric (HILIC–ESI/MS/MS) bioanalysis of basic compounds. *J. Chromatogr. B* **2005,** *825* (2), 186-192.

117. Kipping, M.; Schierhorn, A., Improving hydrogen/deuterium exchange mass spectrometry by reduction of the back-exchange effect. *J. Mass Spectrom.* **2003,** *38* (3), 271-276.

118. Chetty, P. S.; Mayne, L.; Lund-Katz, S.; Stranz, D.; Englander, S. W.; Phillips, M. C., Helical structure and stability in human apolipoprotein A-I by hydrogen exchange and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (45), 19005-19010.

119. Zhang, Z.; Zhang, A.; Xiao, G., Improved Protein Hydrogen/Deuterium Exchange Mass Spectrometry Platform with Fully Automated Data Processing. *Anal. Chem.* **2012,** *84* (11), 4942-4949.

120. Feng, L.; Orlando, R.; Prestegard, J. H., Amide Proton Back-Exchange in Deuterated Peptides: Applications to MS and NMR Analyses. *Anal. Chem.* **2006,** *78* (19), 6885-6892.

121. Jørgensen, T. J. D.; Gårdsvoll, H.; Ploug, M.; Roepstorff, P., Intramolecular Migration of Amide Hydrogens in Protonated Peptides upon Collisional Activation. *J. Am. Chem. Soc.* **2005,** *127* (8), 2785-2793.

122. Rand, K. D.; Adams, C. M.; Zubarev, R. A.; Jørgensen, T. J. D., Electron Capture Dissociation Proceeds with a Low Degree of Intramolecular Migration of Peptide Amide Hydrogens. *J. Am. Chem. Soc.* **2008,** *130* (4), 1341-1349.

123. Zehl, M.; Rand, K. D.; Jensen, O. N.; Jørgensen, T. J. D., Electron Transfer Dissociation Facilitates the Measurement of Deuterium Incorporation into Selectively Labeled Peptides with Single Residue Resolution. *J. Am. Chem. Soc.* **2008,** *130* (51), 17453-17459.

124. Pan, J.; Han, J.; Borchers, C. H.; Konermann, L., Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* **2009,** *131* (35), 12801-12808.

125. Mazon, H.; Marcillat, O.; Forest, E.; Vial, C., Local dynamics measured by hydrogen/deuterium exchange and mass spectrometry of creatine kinase digested by two proteases. *Biochimie* **2005,** *87* (12), 1101-1110.

126. Wu, Y.; Engen, J. R.; Hobbins, W. B., Ultra Performance Liquid Chromatography (UPLC) Further Improves Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2006,** *17* (2), 163-167.

127. Dass, C., *Fundamentals of Contemporary Mass Spectrometry*. John Wiley & Sons, Inc.: Hoboken, New Jersey, **2007**.

128. Mora, J. F. d. l.; Van Berkel, G. J.; Enke, C. G.; Cole, R. B.; Martinez-Sanchez, M.; Fenn, J. B., Electrochemical processes in electrospray ionization mass spectrometry. *J. Mass Spectrom.* **2000,** *35* (8), 939-952.

129. Konermann, L., A Simple Model for the Disintegration of Highly Charged Solvent Droplets during Electrospray Ionization. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (3), 496-506.

130. 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.

131. *Agilent 6200 Series TOF and 6500 Series Q-TOF LC/MS System Concepts Guide*. 4th Ed. ed.; Agilent Technologies, Inc.: Santa Clara, CA, **2007**.

132. Gabelica, V.; Pauw, E. D., Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrom. Rev.* **2005,** *24* (4), 566-587.

133. Plumb, R. S.; Johnson, K. A.; Rainville, P.; Smith, B. W.; Wilson, I. D.; Castro-Perez, J. M.; Nicholson, J. K., UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun. Mass Spectrom.* **2006,** *20* (13), 1989-1994.

134. Uversky, V., Intrinsically Disordered Proteins and Their Environment: Effects of Strong Denaturants, Temperature, pH, Counter Ions, Membranes, Binding Partners, Osmolytes, and Macromolecular Crowding. *The Protein Journal* **2009,** *28* (7), 305-325.

135. Houde, D.; Peng, Y.; Berkowitz, S. A.; Engen, J. R., Post-translational Modifications Differentially Affect IgG1 Conformation and Receptor Binding. *Mol. Cell. Proteomics* **2010,** *9* (8), 1716-1728.

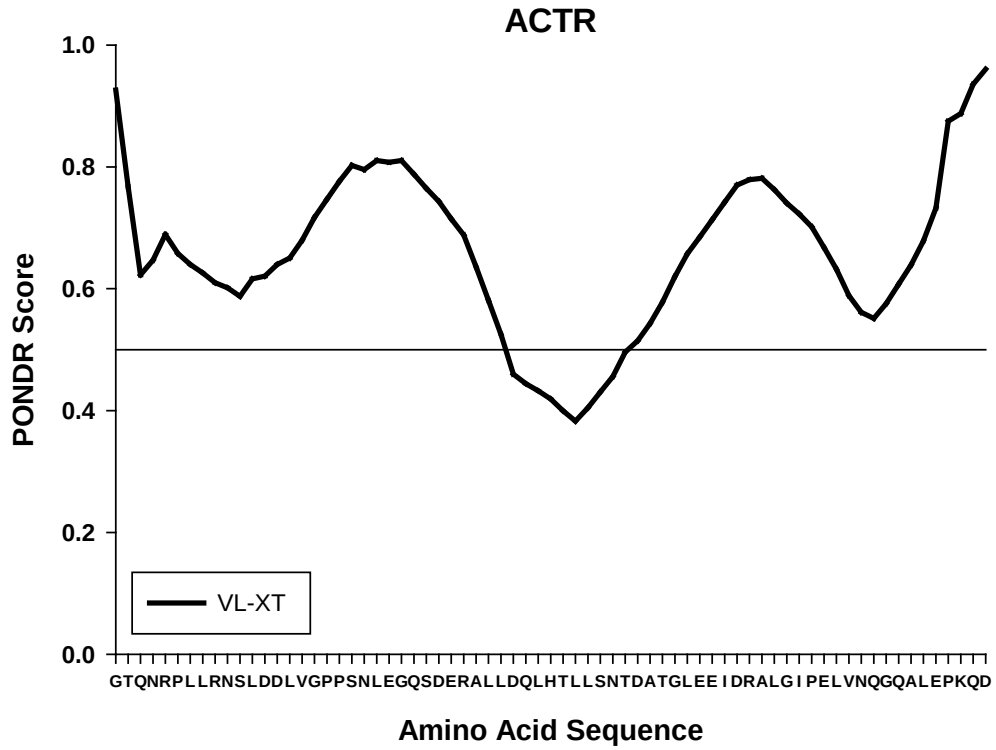


Figure 1.1. The PONDR-VL-XT disorder prediction for ACTR domain sequence.

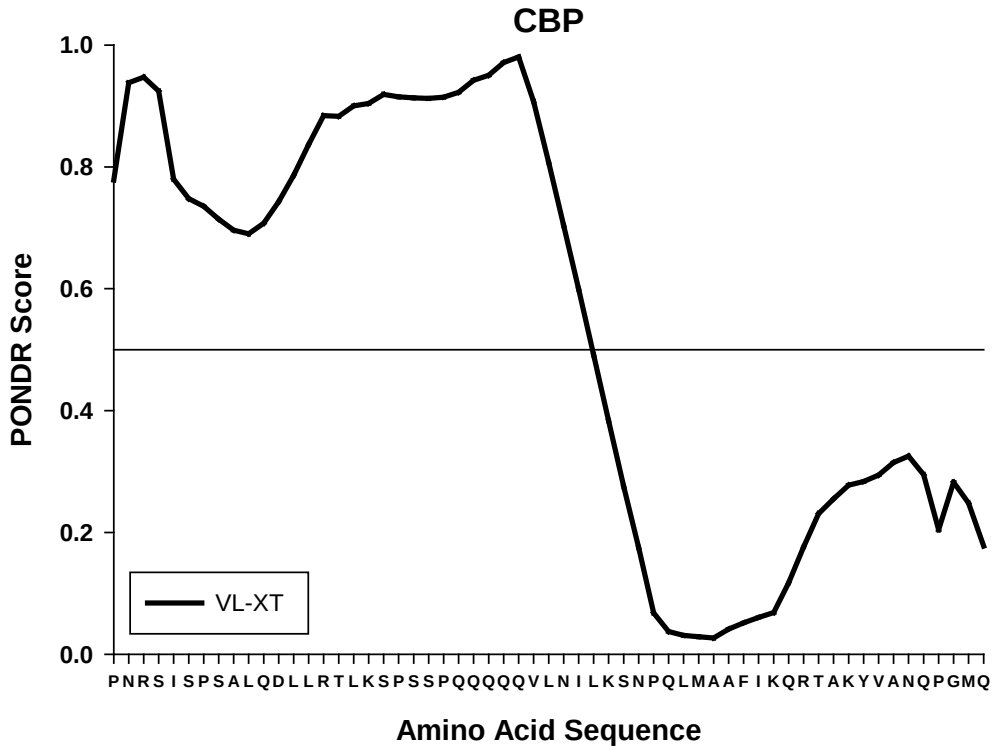


Figure 1.2. The PONDR-VL-XT disorder prediction for CBP domain sequence.

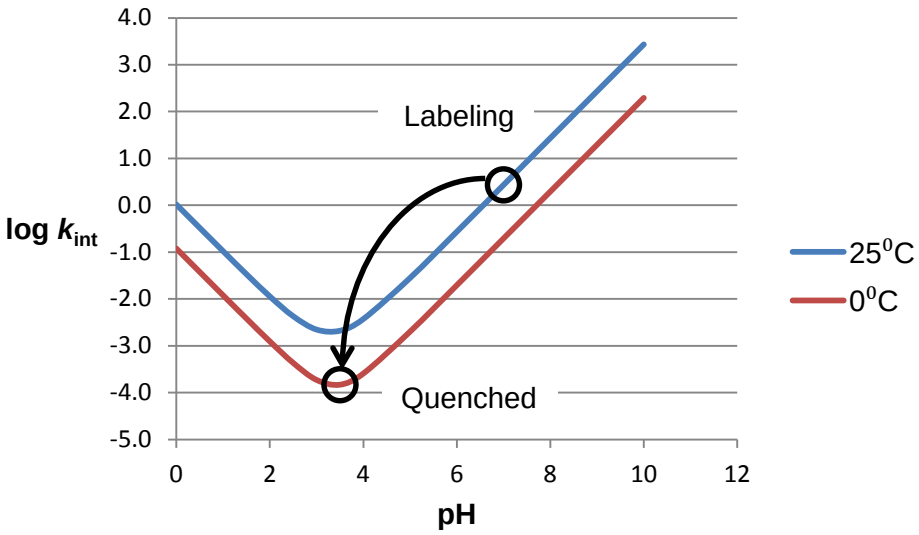


Figure 1.3. This graph shows the calculated[96, 103] forward exchange intrinsic rate constant for the 4th residue in a $(Ala)_7$ peptide chain as a function of pH at two different temperatures.

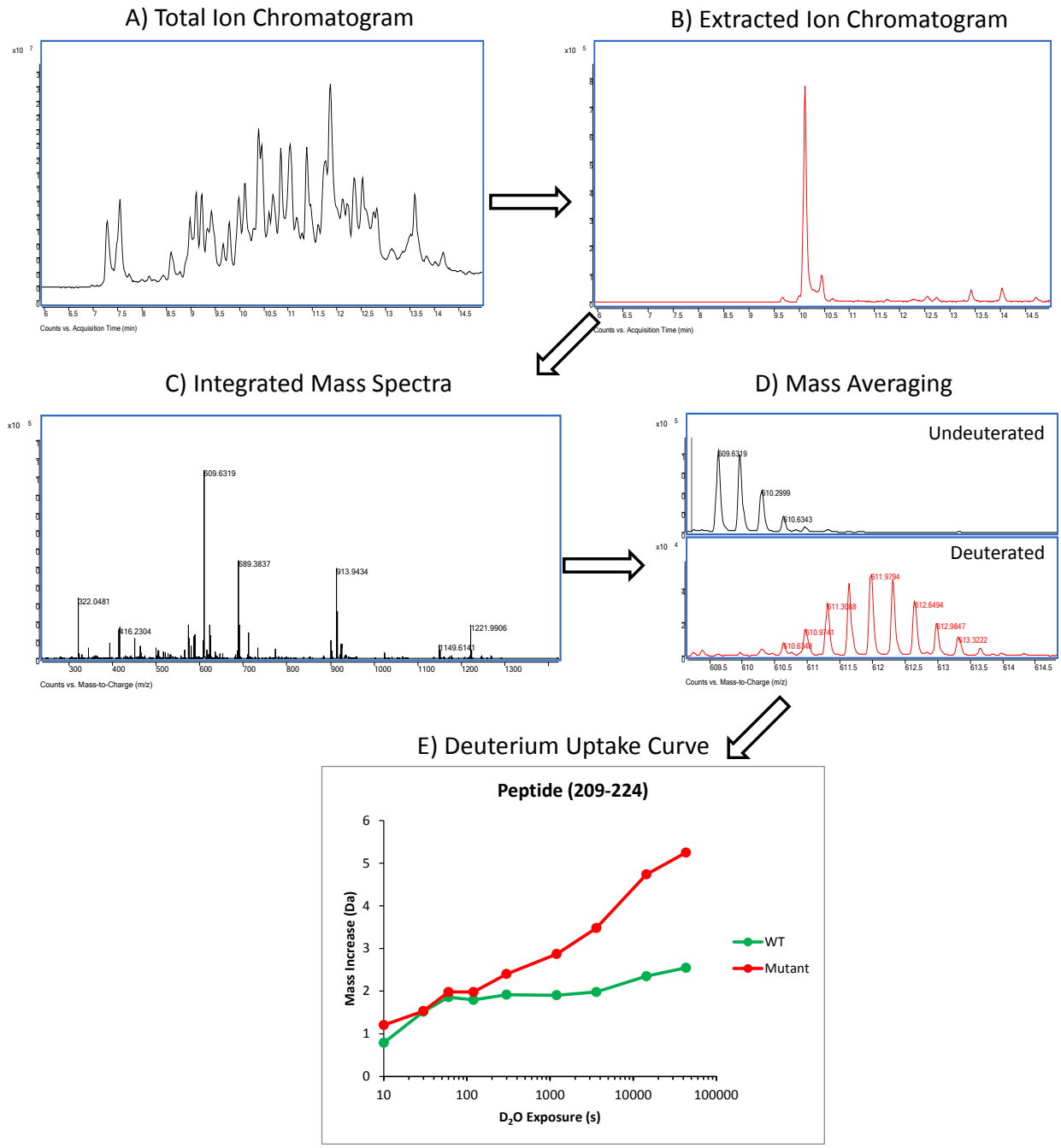


Figure 1.4. Example mass spectra workflow. **A)** Initial TIC measuring total ion intensity over time. **B)** EICs show intensity of specific *m/z* values. This EIC shows the intensity of the 609.63 *m/z* peak and reveals that this peptide elutes at about 10.1 minutes. **C)** Mass spectra integrated from the peak revealed by an EIC. Here, the high intensity 609.63 *m/z* peak can be seen in the mass spectrum extracted from a window around 10.1 minutes. **D)** Clipped mass spectra showing the distribution of *m/z* values for an undeuterated and deuterated peptide. The average mass of the undeuterated peptide is 1826.98 Da. The average mass of the deuterated peptide is 1830.46 Da. The measured mass increase is 3.48 Da. **E)** Deuterium uptake curves showing average mass increase at each D_2O exposure time point for a peptide from both the mutant protein and WT protein.

Chapter 2: An Efficient and Inexpensive Refrigerated LC System for H/D-MS

2.1 Introduction

Hydrogen/deuterium exchange mass spectrometry (H/D-MS) has become widespread in studying protein conformation and dynamics in solution.[1-4] In a typical experiment, a protein is exposed to a large excess of deuterium oxide (D_2O) so that amide hydrogens will exchange with solvent deuterium. The rate of deuterium uptake is limited by hydrogen-bonding and secondary structure; measurement of the rate provides information regarding structural fluctuations within a protein.[2, 5-6] H/D exchange is carried out for set periods of time so that the progression of deuterium uptake can be measured over time. The H/D exchange reaction is then quenched by mixing samples with acid to decrease the pH to ~2.5, followed by rapid cooling of the samples. Deuterium uptake information may also be localized using digestion of deuterated proteins with acid-tolerant proteases followed by HPLC separation.[7-8] Retention of the deuterium label is pivotal to the success and usefulness of H/D-MS studies. Unfortunately, once the reaction has been quenched, the deuterium label slowly back-exchanges with solvent hydrogen.[8] Immediate and stable cooling is necessary to slow the rate of back-exchange during a short LC run.

A typical approach is to immerse all LC components in an ice/water[8-10] or temperature-controlled water bath.[11] Baths are inexpensive and offer efficient cooling of LC components. However, troubleshooting leaks is nearly impossible and water can damage submerged components. Baths also pose the risk of spills or leaks into the LC/MS. Hence, iceless cooling systems for refrigerated HPLC have been developed for H/D-MS.[12-14] In our lab, conventional refrigerators had unacceptable variation in temperature, typically at 4±2 ^{o}C, and long re-equilibration when solvent reservoirs were refilled. Furthermore, even with forced-air convection, cooling is inefficient since heat must be transferred from the mobile phase to the LC components and then dissipated into the air. A system featuring direct mobile phase cooling using thermoelectrically-cooled mobile phase heat exchangers could only reach 4 ^{o}C and suffered from long gradient delays.[14] A thermal chamber to cool the HPLC components to 1.25

^{o}C has been reported, but little detail was provided.[12] A UHPLC system, designed specifically for the Waters nanoAcquity, has also been reported but requires specialized LC pumps and bulky water chillers for heat dissipation.[13]

Here, we detail the performance of a compact "all-in-one" refrigerated unit. The system features an internal protease column oven to enhance digestion efficiency. The unit is constructed from off-the-shelf components, requires minimal specialized assembly, and has a small footprint. A novel single valve design simplifies sample injection and HPLC separation procedures. We also provide an empirical demonstration of cooling efficiency, chromatographic performance, and deuterium recovery.

2.2 Experimental

2.2.1 Reagents

Optima LC/MS grade acetonitrile and water were obtained from Fisher Scientific (Hanover Park, IL). 99+% formic acid in 1 mL ampules was obtained from Thermo Scientific (West Palm Beach, FL). Deuterium oxide (99.9% D) and myoglobin from horse skeletal muscle were obtained from Sigma (St. Louis, MO).

2.2.2 Columns and Traps

The 50 × 2.1 mm immobilized protease column was prepared in-house, as previously described,[15-16] using pepsin from porcine gastric mucosa (Sigma, St. Louis, MO), and POROS 20 AL column packing material (Applied Biosystems, Carlsbad, CA). A 10 × 1.0 mm C12 trap, packed in-house with Jupiter 4u Proteo 90 Å stationary phase from Phenomenex (Torrance, CA), was used for peptide trapping. Myoglobin digests were separated on a 50 × 1.0 mm Ascentis Express Peptide ES-C18 column with 2.7 µm superficially porous particles (Supelco Analytical, Bellefonte, PA).

2.2.3 Construction of the Refrigerated System

For the housing, we used a prefabricated moisture-tight steel box with a hinged lid (35 cm × 30 cm × 20 cm, National Electrical Manufacturers Association standard 13). The housing is

insulated with hand-cut panels made from 2.5 cm-thick high density closed-cell PVC foam insulation (R value of 5). The side and bottom panels were press-fitted into place while the top panel is removable. In some parts of the housing, the insulation was thinned to ~ 1 cm to accommodate fittings. Liquids feed through the walls through 1/16 inch stainless steel bulkhead unions with 0.25 mm bore (VICI-Valco, Houston, TX). The unions are sealed to the outer face of the housing with pressure-sealing washers. Gaskets to seal the openings for the thermoelectric coolers and the switch valve standoff were hand-cut from 1/16 inch thick Buna-N rubber. Except where noted, all basic construction materials were obtained from McMaster-Carr (Elmhurst, IL). To prevent condensation and frost accumulation, the housing is briefly purged with dry air before cooling is started.

The enclosure air is cooled by a thermoelectric air conditioner (76 W capacity, MAA260T-24, MELCOR, Trenton, NJ) (see figure 2.1). The enclosure air temperature is measured using a 15 kΩ thermistor (TS136-76, McShane, Medina, OH) placed at the point where air flow enters the air conditioner heat exchanger. The mobile phase cooler consists of two 1.6 µL chromatographic heat exchangers (G1316-80003, Agilent, Santa Clara, CA) mounted to a thermoelectric cold plate (36 W capacity, CP-036, TE Technology, Traverse City, MI) (see figure 2.1). Heat transfer between the cold plate and heat exchangers is mediated by a thin layer of zinc oxide thermal paste (TP-1, TE Technology, Traverse City, MI) applied to the heat exchangers. The cold plate temperature is measured using a 15 kΩ thermistor (MP-2996, TE Technology, Traverse City, MI). The thermoelectric coolers are powered by two 300 W DC programmable power supplies (RKW 24-14K for the air conditioner and RKW 12-27K for the cold plate, Kepko, Flushing, NY). The power supply voltages are controlled by separate proportional-integral-derivative (PID) temperature controllers operating in cool-only proportional output mode (TC-48-20, TE Technology, Traverse City, MI). An integrated assembly comprising the temperature controllers, power supplies, and system power was constructed in-

house. The controller features switches that allow the thermoelectric cooler fans to be run independently from the thermoelectric elements.

The protease column oven consists of a 3 inch long × 0.563 inch ID ceramic tube (99.5% high purity alumina, Sentro Tech Corp., Berea, OH) with nichrome wire (28 gage, 60% nickel / 16% chromium, Omega Engineering, Inc., Stamford, CT) coiled around the outer surface and affixed with epoxy. The column oven is insulated with two PVC foam insulation blocks. The column oven is powered by a 1000 mA / 12 VDC output plug-in voltage transformer (McMaster-Carr, Elmhurst, NJ) regulated by a CSi32 Series RTD-input PID bench-top controller (Omega Engineering, Inc., Stamford, CT). The bench-top controller measures the temperature of the protease column oven via a PT100 class “B” 4-wire RTD surface sensor (Omega) placed between the protease column and inner wall of the ceramic tube. While placing the oven outside the enclosure would decrease the heat load, the long runs of tubing required would introduce undesirable system volume.

2.2.4 Chromatographic System

The chromatographic system consists of a capillary binary gradient pump with high pressure mixing (G1376A, Agilent, Santa Clara, CA), used for elution flow, and an isocratic pump (G1310A, Agilent, Santa Clara, CA), used for the digestion flow (see figure 2.2). All mobile phases were degassed by an in-line vacuum degasser (G1379B, Agilent, Santa Clara, CA). All tubing that comes into contact with sample is 1/16 inch outer diameter PEEK-clad fused silica capillary with inner diameters ranging from 75 µm to 150 µm (SGE, Austin, TX). Upstream solvent delivery lines are 1/16 inch outer diameter PEEK.

The switch valve is a two-position 12 port valve with 0.25 mm ports (D6C12CW-RAD, VICI-Valco, Houston, TX). The valve is attached to a right angle drive gearbox and mounted to a 15 cm stand-off assembly. Switching of the valve is actuated by a microelectric actuator and controller (EP, VICI-Valco, Houston, TX). The actuator is mounted to the outside of the enclosure (see figure 2.3). The valve is actuated by software-controlled contact closure outputs

sent to the valve controller by a BCD interface board (G1351-68701, Agilent, Santa Clara, CA) installed in the isocratic pump (see figure 2.2). A manual push-button is used to trigger the LC/MS run start after manual sample loading. The push-button start and valve control buttons are incorporated into a simple one-piece assembly constructed in-house.

2.2.5 Temperature Measurements

Fluid temperature measurements were measured with an insulated ultra fine type T thermocouple (PY2 72-7567, Harvard Apparatus, Holliston, MA) The air temperature sensor was a glass braid insulated type K thermocouple with a bare junction (SC-TT-K, Omega Engineering, Stamford, CT). Temperatures were recorded at 1 sec intervals using a logging dual-channel temperature meter with a wireless computer interface (HH806W, Omega Engineering, Stamford, CT) and computed as ten second averages.

Our mobile phase temperature measurements were restricted to pure water to prevent damage to the thermocouple. While there are differences in the thermal properties of water and acetonitrile, a simple thermal analysis[17] shows that these differences are negligible. Mobile phase heat exchange in our system can be approximated as a fluid undergoing laminar flow through a tube with an isothermal surface. Under these conditions, heat transfer out of the fluid is dominated by heat transfer from the fluid to the tube wall. The fluid temperature at the exit (T_e) of a heat exchanger with a wall surface temperature of T_s is calculated by equation 1.

$$T_e = T_s - \left(T_s - T_i\right) e^{-\mathrm{NTU}} \tag{1}$$

Here, T_i is the inlet temperature and NTU is the *number of transfer units,* a dimensionless measure of heat transfer. The number of transfer units is calculated by equation 2.

$$\mathrm{NTU} = \frac{3.66\pi Lk}{\dot{V}\rho C_P} \tag{2}$$

Here, L is the length of the tubing, k is the thermal conductivity of the fluid, $\dot{V}$ is the volume flow rate, ρ is the bulk density of the fluid, and C_P is the specific heat capacity of the fluid. The factor of 3.66 arises under the condition of laminar flow in a cylindrical tube. Examination of equation 1 shows that the exit temperature asymptotically reaches the surface temperature above ~5 NTUs. Although water has a higher thermal conductivity than acetonitrile (k = 0.606 W m^{-1} K^{-1} for H_2O *vs.* k = 0.188 W m^{-1} K^{-1} for CH_3CN at 25 °C), the product of specific heat and density is also considerably higher (ρC_P = 4.2 J cm^{-3} K^{-1} for H_2O at 25 °C *vs.* ρC_P = 1.8 J cm^{-3} K^{-1} for CH_3CN at 20 °C).[18] Hence, in the asymptotic limit of NTU > 5, the differences in the thermal properties are negligible. On this basis, our temperature measurements, based only on water can be extrapolated to organic solvents.

Fluid temperature was measured by diverting flow through a short piece of PEEK tubing at various points. An insulated thermocouple was fed into the open end of the PEEK tubing. Water from this deliberate leak was collected in a beaker. Air temperature was measured with a bare junction thermocouple. Both thermocouples were calibrated against melting ice using a single-point offset correction.

2.2.6 LC/MS Measurements

Injected peptide samples were passed through the pepsin column and loaded onto the C12 trap by 0.1% formic acid in water at 0.2 mL min^{-1} for 2.5 minutes to desalt and focus samples. Solvent A was 0.1% formic acid in water, and solvent B was 90% acetonitrile / 10% water / 0.1% formic acid. The peptides were separated on the ES-C18 column using a 15% to 70% B gradient in 5 minutes at 50 µL min^{-1}. The optimized gradient omitted a few outlying peptides that were not retained on the column or that failed to elute until the column washing step. The refrigerated enclosure temperature was maintained at 0 ^{o}C and the protease column was unheated. Eluted peptides were ionized by electrospray and analyzed by a time of flight mass spectrometer (Agilent model 6220) operating in positive ion mode. Nitrogen drying gas

was heated to 275 ^{o}C at a flow rate of 10 L min^{-1}, with a nebulizer pressure of 30 psig. Fragmentor, skimmer, and capillary voltages were set to 150 V, 50 V, and 4000 V, respectively. In measuring deuterium recovery, the cold protease column was used in the system to keep flow path volumes and conditions constant, despite the fact that deuterated myoglobin was pepsin-digested prior to injection.

2.2.7 H/D Exchange

To ensure complete deuteration, peptic peptides were prepared by digesting myoglobin with pepsin *before* labeling.[19] 5 µL aliquots of myoglobin peptides or angiotensin I were deuterated with 95 µL of deuteration buffer (10 mM phosphate, 50 mM sodium chloride dissolved, pD 6.9). Exchange was quenched after at least 89 minutes with 100 µL of ice-cold 200 mM sodium phosphate (pH 2.6). To maximize reproducibility, quenched samples were loaded into the 100 µL sample loop (figure 2.3) 60 seconds after quench followed by sample injection after an additional 30 seconds. 19-fold dilution with D_2O gives a maximum possible deuteration of 95%. Our deuteration values were normalized as a percentage of this maximum. Fully deuterated angiotensin I was prepared as described above except it was not digested with pepsin.

2.2.8 Data Processing

Undeuterated samples of myoglobin peptic peptides were analyzed first, and identified peptides were matched to the sequence of myoglobin using MassHunter Qualitative Analysis software with BioConfirm (version B.03.01, Agilent Technologies, Santa Clara, CA). Only non-ambiguous peptides identified with a mass tolerance of ±20 ppm were used in this study. The raw spectra of deuterated peptides were analyzed with HX-Express software[20] to find the average mass of each peptide. The difference between the average deuterated mass and the theoretical undeuterated mass was used as the measurement of deuterium uptake. The maximum number of exchangeable amide hydrogens was calculated by taking the number of amino acid residues in a peptide, subtracting two for the rapid back-exchange of the first two

residues, and subtracting the number of proline residues (since proline lacks an amide hydrogen).

2.3 Results and Discussion

2.3.1 Description of the Refrigerated System

Back-exchange is minimized when LC conditions are maintained at the lowest possible temperature. To meet this need, we constructed a thermoelectrically cooled (*i.e.*, Peltier effect) compartment with precise mobile phase and independent protease column temperature controls (see figure 2.1). The enclosure is cooled using an air-to-air thermoelectric cooler module. Mobile phases are cooled by heat exchangers (2 µL internal volume) mounted to a thermoelectric cold plate assembly. Heat is dissipated to the surroundings by forced air convection, eliminating the need for secondary heat transfer to a circulating chiller.[13] Since proteolysis can be more efficient at elevated temperatures,[13] we incorporated an insulated, low power oven (< 5W) so that the protease column can be warmed up to 20 °C above the enclosure temperature. Temperatures are regulated by independent PID (proportional integral derivative) temperature controllers (see figure 2.1). One innovative feature of our system is immediate cooling of peptides back to 0 °C at the outlet of the protease column oven (see figure 2.3). The efficiency of the heat exchangers increases with increasing flow rate (see figure 2.4 and table 2.1). By setting the heat exchangers to –4 °C, the mobile phase cools to at least 0 °C at all flow rates. The unit was fabricated almost entirely from off-the-shelf assemblies. The total cost of materials was less than US$5,000. Assembly required only simple machining of metal parts and hand cutting of foam insulation. The chromatographic system consists of two pumps and online digestion using an immobilized protease column,[16] but we use a single twelve port valve to handle the entire analysis (see figure 2.3). Simple contact closures programmed in the gradient time-table trigger valve actuation (see figure 2.2). The loop is manually loaded with the valve in the *elute* position (figure 2.3, lower panel). When triggered by an external push-button controller, the valve switches to the *digest and trap* position (figure 2.3, middle panel).

The digestion flow (0.1% formic acid) carries the sample through the loop and protease column. The resultant peptides are trapped, concentrated, and desalted on a reversed-phase peptide trap. (Our system is readily converted to analysis of intact proteins by replacing the protease column with a union.) When the valve is switched back to the *elute* position, a water/acetonitrile gradient elutes the peptides from the trap and separates them on a reversed-phase column. Simultaneously, the protease column is flushed to waste. By forming the gradient outside the housing, our refrigerated system is readily interfaced with any gradient HPLC system without re-plumbing. In particular, our capillary HPLC pump includes a mixer, pulse damper, proportioning valve, and flow sensor that remain under ambient conditions. Although our system is not compatible with 600 bar pressures, simple upgrades would enable UHPLC operation.

2.3.2 Dependence of the Cooling Efficiency on Flow Rate

To determine the efficiency of the Peltier-cooled mobile phase heat exchangers, we first measured the temperature of water immediately downstream of the heat exchanger. In our initial experiments, we set the enclosure air temperature to 0 °C and the heat exchanger cold plate to −4 °C. We measured the temperature of water exiting the heat exchanger as a function of flow rate (see figure 2.4). As the flow rate was increased, the fluid temperature asymptotically approached the −4 °C temperature of the cold plate, showing that heat exchange becomes more efficient as the flow rate increases. The cold plate set point temperature that was required to reach 0 °C at various flow rates is listed in table 2.1. While we use the 50 µL min^{-1} flow rate for the gradient separation of the peptides, we use a flow rate of 200 µL min^{-1} for the digestion flow (see figure 2.3). Significant under-cooling of the heat exchanger cold plate was required to reach ~0 °C at the 50 µL min^{-1} flow rate (see figure 2.4). Setting the cold plate to –4 °C caused water, flowing at 200 µL min^{-1}, to supercool to −1.8 °C. Supercooling of water in small volumes in the absence of ice-nucleating surfaces is a very well-known phenomenon with a body of literature that dates back almost two hundred years.[21] As a recent example, NMR spectra of proteins in liquid water at –15 °C were acquired from samples in 1.0 mm diameter

glass capillaries.[22] Given the narrow dimensions of our tubing (< 200 µm) and the short residence time of the fluid in the heat exchanger, it is not at all surprising that we observed limited supercooling. Since our flow path should be in thermal equilibrium with the air inside the chamber, held at 0 °C, we expected that water would not remain supercooled once it had passed through the tubing and valve downstream from the heat exchanger. Indeed, when we measured the temperature of water flowing at 200 µL min^{-1} as it exited port #8 (see figure 2.3), we found that the water temperature had increased to 0.1 °C. We have used this system routinely for over one year and have never encountered any blockages or clogging due to the instantaneous nucleation of supercooled water.

2.3.3 Temperature Stability

We measured air temperature in the enclosure and the mobile phase temperature (water, 200 µL min^{-1}) over a period of nine hours from a room temperature start. The system reached a stable operating temperature of ~0 °C in less than one hour (see figure 2.5). With the protease column oven set to 15 °C, the enclosure air temperature was stable at 0.5±0.1 °C. (With the protease column oven turned off, the enclosure temperature was –0.1±0.1 °C.) We evaluated the stability of the fluid temperature under several different scenarios. Water exiting the heat-exchanger positioned at the outlet of an unheated protease column (see figure 2.3) was –2.2±0.1 °C (see figure 2.5, Heater Off). With the protease column oven heated to 15 °C, the protease column effluent remained near 13.0 °C (figure 2.5, Protease Out), demonstrating both its stability and efficiency. Heating the protease column increases digestion efficiency, but would accelerate back-exchange. A heat exchanger located immediately downstream from the protease column cools the effluent back to –2.2±0.1 °C (figure 2.5, Exchanger Out). The rapid, but small fluctuations are on the order of the precision limit of our temperature meters. Under quench conditions (pH 2.5), these fluctuations would result in approximately a 1% decrease in the half-life of amide D/H back-exchange,[23-24] hence this degree of temperature stability is suitable for carrying out H/D-MS.

2.3.4 Chromatographic Performance

We assessed chromatographic performance under refrigerated conditions using an optimized water/acetonitrile gradient and a column with superficially porous C18 particles. For complex digests, chromatographic resolution is critical for minimizing MS overlap and ion suppression. For the myoglobin digest, peak widths at half maximum ranged from 3.6 to 7.5 seconds, with a median of 4.4 seconds (see table 2.2 and figure 2.6). For comparison, the median peak width obtained using UHPLC for the room temperature separation of a phosphorylase B digest was 2.7 seconds.[13] Hence the chromatographic performance remains commendable. In addition, since our system does not operate at the high system pressures required in UHPLC, we minimize the potential of extra back-exchange caused by frictional heating in the column.[25]

2.3.5 Deuterium Recovery

For fully deuterated myoglobin peptides, the deuterium recovery ranged from 63.2% to 97.7% with a mean of 80.2% (table 2.2 and figure 2.7). Back-exchange calculations (at pH 2.5, 0 °C)[23-24] show that the wide range of deuterium recoveries is not surprising. For example, for peptide 110-134, which has a calculated half-life of 46 minutes, we measured 63.2% deuterium recovery, while for peptide 8-13, which has a half-life of 128 minutes, we measured 97.7% deuterium recovery. Heating the protease column from 0 °C to 15 °C resulted in a slight decrease in deuterium recovery, from 71.7% to 70.5%, with fully deuterated angiotensin I (see table 2.2).

Our deuterium recovery results are at least as good as recovery obtained using other systems. An average of 70.0% recovery for cytochrome *c* peptides was obtained with a UHPLC system.[13] With an ultrafast elution gradient 78% recovery for the test peptide LHRH was obtained.[26] Recovery using an HPLC system in a 0 °C water bath was measured using angiotensin I that had been labeled for over one month[11] appears to be ~84% (~5.9 Da increase), but when we account for slow H/D exchange at the C2 carbons of the imidazole rings

of two histidine residues,[27] the amide deuterium recovery was ~56%. Hence, our recovery of 71.7% for this same peptide, labeled for only 1-2 hours, was considerably higher. In addition to excellent recovery, our system demonstrated outstanding reproducibility, with a mean relative standard deviation of ±0.91% for measured deuterium uptake in myoglobin peptides (table 2.2). For the ultrafast elution gradient of test peptide LHRH, 0.71% error was reported in deuterium uptake measurements.[26] While these results may not be directly comparable because of differences between methods, deuterium recovery using our refrigerated system and a typical gradient is more than satisfactory for H/D-MS analyses.

2.4 Conclusions

We have constructed an inexpensive refrigerated system and evaluated its performance for H/D-MS. The temperatures of the mobile phases were directly measured to validate the effectiveness of cooling, something that has not been previously reported. Small volume heat exchangers efficiently cooled the mobile phases, allowing gradients to be formed at room temperature. This efficiency eliminates the need for large dwell volume for cooling that can delay or broaden gradients, a significant advantage when many samples need to be run in succession. Efficient and stable cooling of the system resulted in excellent deuterium recovery and reproducibility. Use of a superficially porous reversed-phase column gave excellent chromatographic results with short gradients under quenched conditions (pH 2.5, 0 °C) without the need for ultra-high pressures. Finally, our refrigerated LC system can be readily interfaced with any LC/MS system since only simple contact closures are required to drive valve actuation.

References

1. Fitzgerald, M. C.; West, G. M., Painting Proteins with Covalent Labels: What's In the Picture? *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (6), 1193-1206.
2. Konermann, L.; Tong, X.; Pan, Y., Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **2008,** *43* (8), 1021-1036.
3. Marcsisin, S.; Engen, J., Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Anal. Bioanal. Chem.* **2010,** *397* (3), 967-972.
4. Tsutsui, Y.; Wintrode, P. L., Hydrogen/Deuterium Exchange-Mass Spectrometry: A Powerful Tool for Probing Protein Structure, Dynamics, and Interactions. *Curr. Med. Chem.* **2007,** *14*, 2344-2358.
5. Englander, S. W.; Kallenbach, N. R., Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quarterly Review of Biophysics* **1984,** *16* (4), 521-655.
6. Hvidt, A.; Nielsen, S. O., Hydrogen exchange in proteins. *Adv. Protein Chem.* **1966,** *21*, 287-385.
7. Rosa, J. J.; Richards, F. M., An experimental procedure for increasing the structural resolution of chemical hydrogen-exchange measurements on proteins: Application to ribonuclease S peptide. *J. Mol. Biol.* **1979,** *133* (3), 399-416.
8. Zhang, Z.; Smith, D. L., Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* **1993,** *2* (4), 522-531.
9. Brudler, R.; Gessner, C. R.; Li, S.; Tyndall, S.; Getzoff, E. D.; Woods, J. V. L., PAS Domain Allostery and Light-induced Conformational Changes in Photoactive Yellow Protein upon I2 Intermediate Formation, Probed with Enhanced Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Mol. Biol.* **2006,** *363* (1), 148-160.
10. Slysz, G. W.; Percy, A. J.; Schriemer, D. C., Restraining expansion of the peak envelope in H/D exchange-MS and its application in detecting perturbations of protein structure/dynamics. *Anal. Chem.* **2008,** *80*, 7004.
11. Burkitt, W.; O'Connor, G., Assessment of the Repeatability and Reproducibility of Hydrogen/Deuterium Exchange Mass Spectrometry Measurements. *Rapid Commun. Mass Spectrom.* **2008,** *22* (23), 3893-3901.
12. Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; He, Y.; Hendrickson, C. L.; Marshall, A. G.; Griffin, P. R., Probing Protein Ligand Interactions by Automated Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2006,** *78* (4), 1005-1014.
13. Wales, T. E.; Fadgen, K. E.; Gerhardt, G. C.; Engen, J. R., High-Speed and High-Resolution UPLC Separation at Zero Degrees Celsius. *Anal. Chem.* **2008,** *80* (17), 6815-6820.
14. Weis, D. D.; Engen, J. R., Microflow HPLC/ESI enhances hydrogen exchange mass spectrometry. *Protein Sci.* **2005,** *14(s1)*, 213.
15. Busby, S. A.; Chalmers, M. J.; Griffin, P. R., Improving digestion efficiency under H/D exchange conditions with activated pepsinogen coupled columns. *Int. J. Mass spectrom.* **2007,** *259* (1-3), 130-139.
16. Wang, L.; Pan, H.; Smith, D. L., Hydrogen Exchange-Mass Spectrometry: Optimization of Digestion Conditions. *Mol. Cell. Proteomics* **2002,** *1* (2), 132-138.

17. Çengel, Y., *Heat Transfer: A Practical Approach*. 2nd ed.; McGraw-Hill: Boston, **2003**.

18. Haynes, W. M., *CRC Handbook of Chemistry and Physics*. 91st Edition, Internet Version 2011 ed.; Taylor Francis Group: Boulder, Colorado, 2010.

19. Resing, K. A.; Ahn, N. G., Deuterium Exchange Mass Spectrometry as a Probe of Protein Kinase Activation. Analysis of Wild-Type and Constitutively Active Mutants of MAP Kinase Kinase-1. *Biochemistry* **1998,** *37* (2), 463-475.

20. Weis, D. D.; Engen, J. R.; Kass, I. J., Semi-Automated Data Processing of Hydrogen Exchange Mass Spectra Using HX-Express. *J. Am. Soc. Mass. Spectrom.* **2006,** *17* (12), 1700-1703.

21. Angell, C. A., Supercooled Water. *Annu. Rev. Phys. Chem.* **1983,** *34*, 593-630.

22. Skalicky, J. J.; Mills, J. L.; Sharma, S.; Szyperski, T., Aromatic ring-flipping in supercooled water: implications for NMR-based structural biology of proteins. *J. Am. Chem. Soc.* **2001,** *123*, 388-397.

23. Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W., Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 75-86.

24. Connelly, G. P.; Bai, Y.; Jeng, M.-F.; Englander, S. W., Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 87-92.

25. Gritti, F.; Martin, M.; Guiochon, G., Influence of viscous friction heating on the efficiency of columns operated under very high pressures. *Anal. Chem.* **2009,** *81*, 3365-3384.

26. Zhang, H.-M.; Bou-Assaf, G. M.; Emmett, M. R.; Marshall, A. G., Fast Reversed-Phase Liquid Chromatography to Reduce Back Exchange and Increase Throughput in H/D Exchange Monitored by FT-ICR Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2009,** *20* (3), 520-524.

27. Miyagi, M.; Nakazawa, T., Determination of pKa Values of Individual Histidine Residues in Proteins Using Mass Spectrometry. *Anal. Chem.* **2008,** *80* (17), 6481-6487.

Figures and Tables

Table 2.1. Heat exchanger set points required to reach 0 °C mobile phase at different flow rates.

flow rate (µL min^{-1})	heat exchanger set point (°C)	fluid temperature (°C)
50	−4.0	−0.1
100	−2.0	−0.1
200	−0.8	−0.1

The enclosure set point was 0 °C for these measurements.

Table 2.2. Deuterium recovery for deuterated peptides.

Peptide[1]	Sequence	Average deuterated mass (Da)[2]	Theoretical undeuterated mass (Da)	Mass Increase (Da)[3]	Precision RSD (%)[4]	Exchangeable amide H[5]	Deuteration (%)	Peak width[6] (s)
1-7	GLSDGEW	766.82 ± 0.05	762.76	4.06	1.2	5	85.4	3.56 ± 0.03
8-13	QQVLNV	703.51 ± 0.02	699.80	3.71	0.66	4	97.7	3.76 ± 0.03
8-29	QQVLNVWGKVEADIAGHGQEVL	2407.27 ± 0.16	2390.65	16.62	0.93	20	87.5	5.78 ± 0.25
14-29	WGKVEADIAGHGQEVL	1719.69 ± 0.10	1708.87	10.82	0.94	14	81.3	4.38 ± 0.06
56-69	KASEDLKKHGTVVL	1533.16 ± 0.10	1524.76	8.40	1.2	12	73.7	7.48 ± 0.54
84-96	AELKPLAQSHATK	1400.95 ± 0.08	1393.59	7.36	1.0	10	77.5	4.44 ± 0.16
110-134	AIIHVLHSKHPGDFGADAQGAMTKA	2586.11 ± 0.12	2572.90	13.21	0.88	22	63.2	4.18 ± 0.28
125-139	ADAQGAMTKALELFR	1631.09 ± 0.07	1621.86	9.23	0.77	13	74.7	6.12 ± 0.31
137-153	LFRNDIAAKYKELGFQG	1981.75 ± 0.08	1970.23	11.52	0.67	15	80.9	4.56 ± 0.12
Mean					0.91		80.2	
Angiotensin I[7]	DRVYIHPFHL	1301.24 ± 0.02	1296.48	4.76	0.34	7	71.7	
Angiotensin I[8]	DRVYIHPFHL	1301.16 ± 0.02	1296.48	4.69	0.44	7	70.5	

[1]Peptides with residue numbers are from myoglobin.
[2]Mean value of 3 replicates ± standard deviation.
[3]Deuterated mass – undeuterated mass.
[4]Relative standard deviation (%) calculated as standard deviation of the average deuterated mass divided by mass increase.
[5]For a peptide with a free N-terminus, the number of exchangeable amides = number of residues – number of prolines not in the first two residues – 2.
[6]Chromatographic peak full width at half-maximum, mean value of 3 replicates ± standard deviation.
[7-8]Angiotensin I with 0 °C[7] and 15 °C[8] protease column.

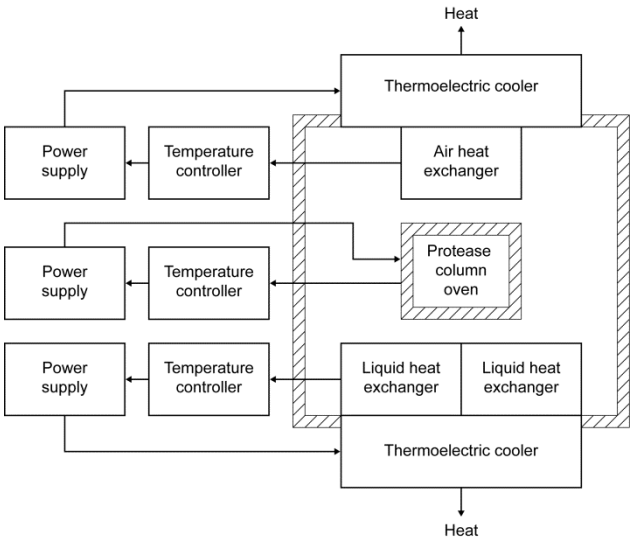


Figure 2.1. The mobile phase temperature, enclosure air temperature, and protease column oven temperature are independently controlled by three different thermoelectric coolers with independent temperature sensors. Further details are provided in the Experimental section.

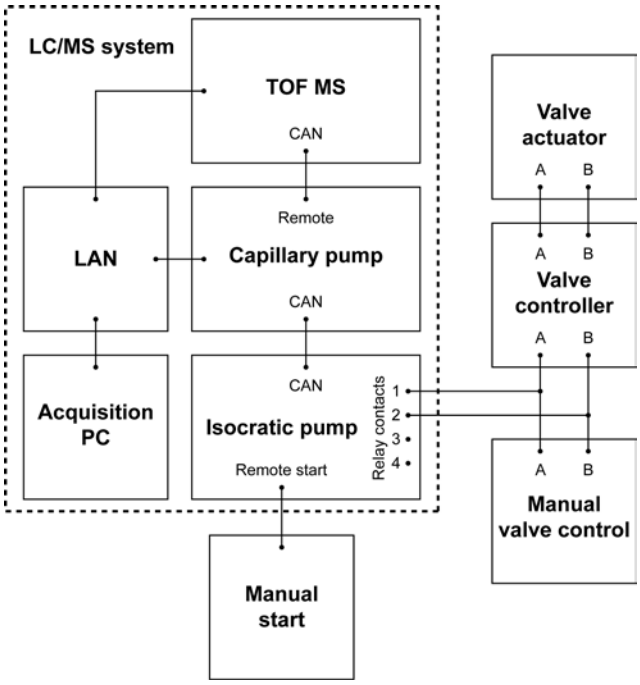


Figure 2.2. The entire refrigerated system is interfaced to the LC/MS (enclosed in the dashed line) using a single input and two outputs, as shown in this schematic diagram of the valve control and instrument interface. A single manual input signal triggers the start of the LC/MS run. The two position switching valve is interfaced to the LC/MS system via two contact closures. (The CAN connectors are part of the Agilent inter-module communication network.)

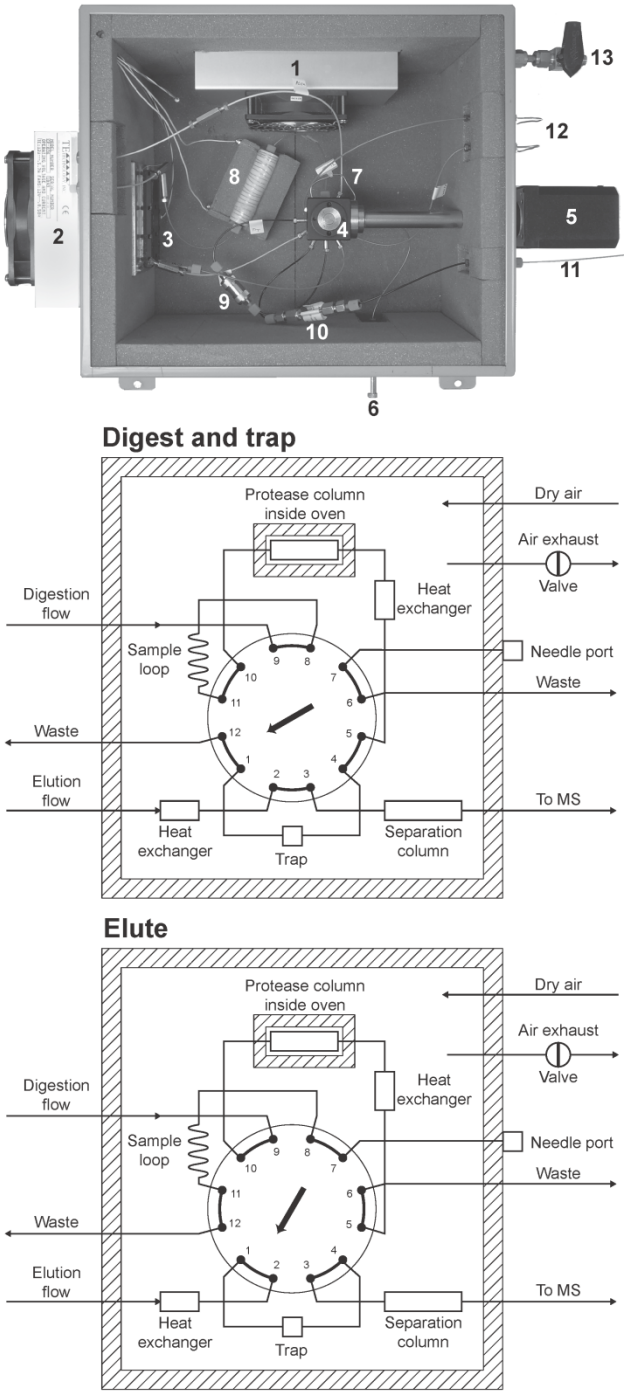


Figure 2.3. Liquid handling in the refrigerated system. The top panel is a photograph of the refrigerated system viewed from above with the lid open. The labeled parts are (1) thermoelectric air cooler module, (2) thermoelectric cold plate module, (3) mobile phase heat exchangers, (4) twelve port switching valve, (5) valve actuator, (6) needle port, (7) sample loop, (8) protease column inside oven, (9) peptide trap, (10) peptide separation column, (11) MS inlet line, (12) waste lines, and (13) inlet and outlet for dry air purge. For visual presentation, the cover of the protease column oven (8) has been removed. The temperature controllers are not shown. The image was processed using Photoshop (Adobe Systems Inc., San Jose, CA) to

correct for lens distortion, to remove glare spots, and to remove the background, cover, and some wiring. The lower panels show schematic diagrams of the flow paths with the switching valve in the digest and trap position (middle) and the elute position (bottom). For clarity, not all components are oriented as shown in the photograph.

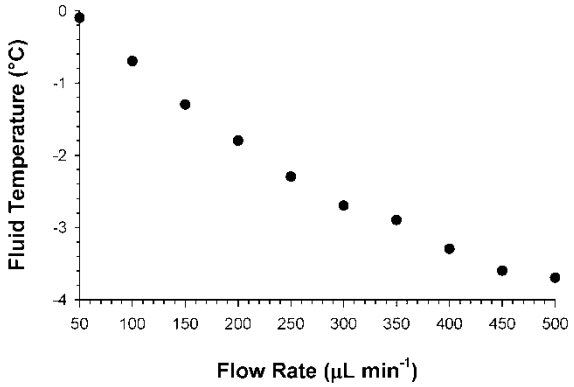


Figure 2.4. The efficiency of the liquid heat exchangers depends on the fluid flow rate, but even at the lowest flow rate, water can be cooled from ambient inlet temperature to 0 °C. Here, the liquid heat exchanger set point was −4 °C and the enclosure air temperature set point was 0 °C. Between flow rate changes, the temperature was allowed to stabilize for 5 min before the temperature was recorded.

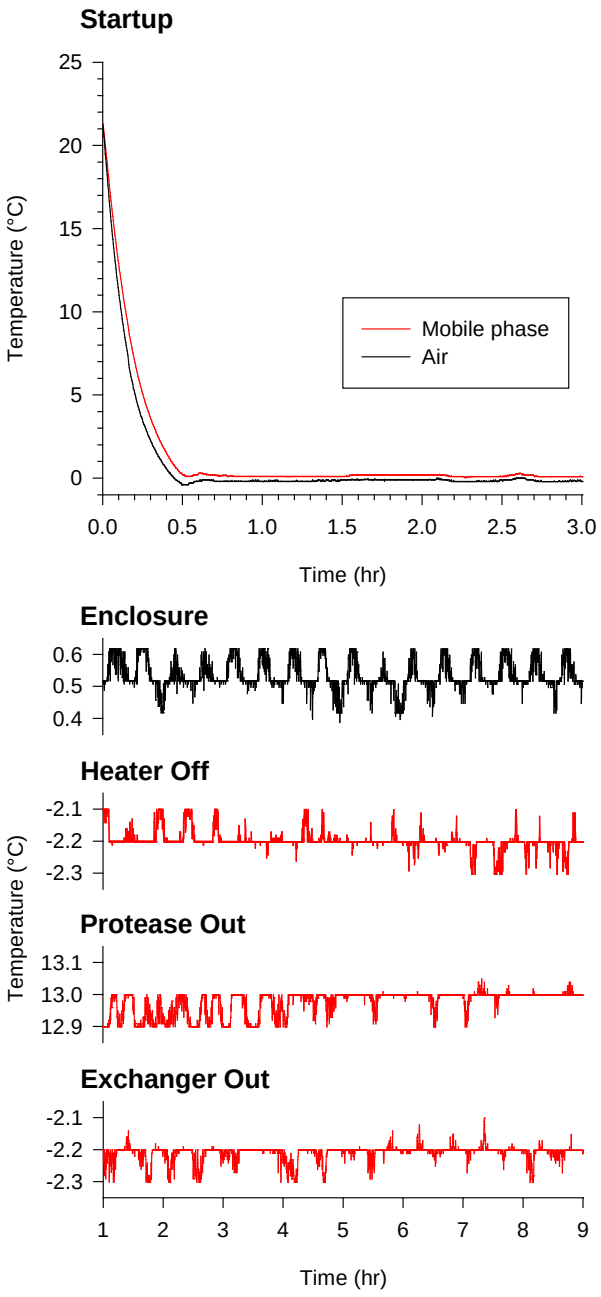


Figure 2.5. Refrigeration of mobile phases and enclosure air from the initial startup is rapid and temperatures are stable once set points are reached. The top panel (**Startup**) shows the amount of time required for both temperatures to reach their respective set points. Here, the water temperature was measured at the inlet of the sample loop. The lower panels show temperature stability after the set points had been reached. (**Enclosure**) shows the air temperature with the protease column oven operating at a set point of 15 ^{o}C. (**Heater Off**) shows the temperature of water emerging from the heat exchanger immediately downstream of the protease column with the column oven turned off. (**Protease Out**) shows the temperature of water emerging from the protease column operating at a set point of 15 ^{o}C. (**Exchanger Out**) shows the temperature of water after being cooled by the post-oven heat exchanger. The flow

rate for all measurements was 200 µL min^{-1}. Temperature measurements were started when the cooling system was turned on at t = 0 hr.

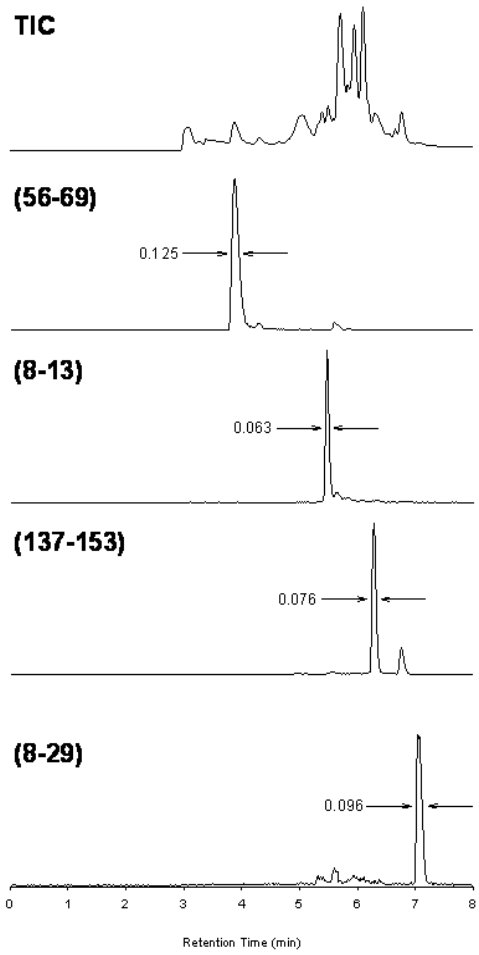


Figure 2.6. The high resolution of our chromatography system is shown by these representative chromatograms from myoglobin peptides. The total ion chromatogram is shown at the top (TIC). Extracted ion chromatograms for select peptides are identified by their amino acid residue numbers. The full widths at half-maximum are the averages of three successive samples (see table 2.2). Each chromatogram has been normalized to 100% maximum relative intensity.

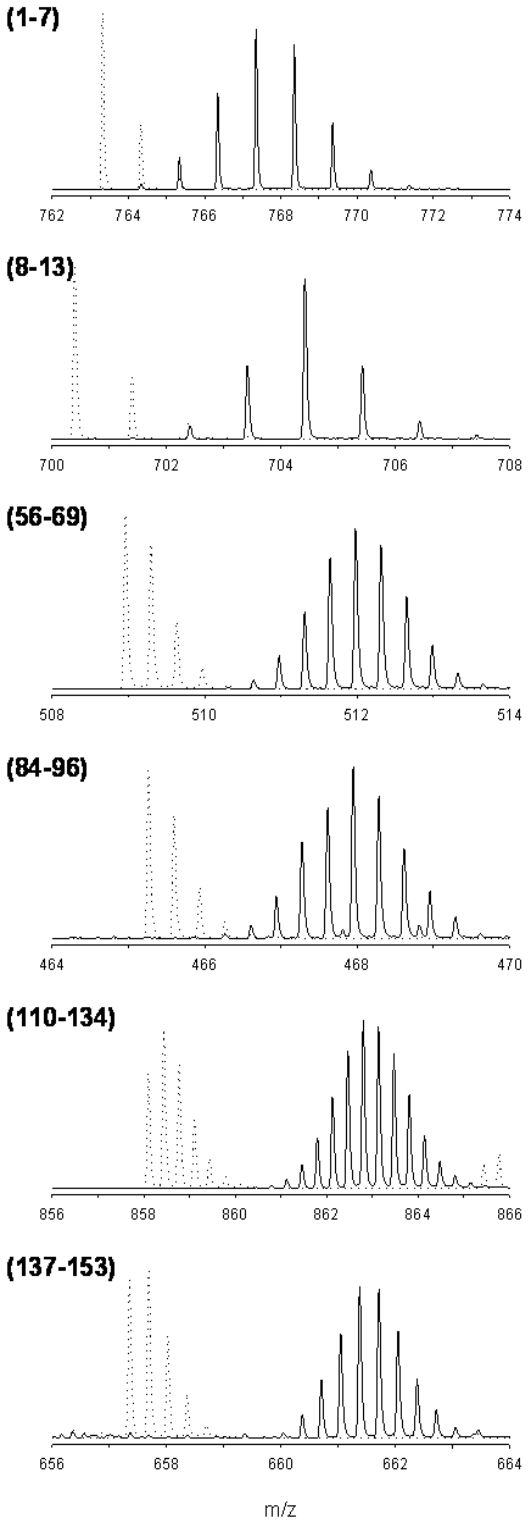


Figure 2.7. Representative mass spectra show deuterium uptake by several myoglobin peptides. The undeuterated isotope envelope spectrum for each peptide is shown with dotted lines. The mass spectra of 95% deuterated peptides are shown with overlaid solid lines. Each

spectrum was normalized to 100% intensity. Sequential amino acid residue numbers are given in parentheses. The sequences of the peptides can be found in table 2.2.

Chapter 3: Mapping Unstructured Regions and Synergistic Folding in IDPs with H/D-MS

3.1 Introduction

Intrinsically disordered proteins (IDPs) contain one or more regions that exhibit a high degree of conformational flexibility and lack a structured conformation. Many IDPs use disordered regions to carry out function.[1-11] Intrinsically disordered proteins differ from structured proteins by low sequence complexity and a bias against hydrophobic residues; these difference have become the basis of a variety of algorithms to predict disorder.[12] Disorder prediction algorithms have shown that IDPs are ubiquitous in the proteomes of higher organisms and comprise roughly one third of all proteins.[13-14] Bioinformatic approaches have demonstrated a clear distinction between the functions of structured proteins and disordered proteins. While protein structure is associated primarily with functions such as enzyme catalysis, ion channels, and cellular structure, protein disorder is associated primarily with functions that require protein-protein interactions such as signaling.[14]

In many cases, the disordered regions that mediate protein-protein interactions undergo coupled binding and folding.[15-16] It has been speculated that the affinity of many IDP-mediated protein-protein interactions is significantly decreased by the strong entropic driving force for the free unstructured protein.[8] At the same time, the conformational flexibility of the unstructured protein allows it to sample many different conformational states. Higher chain flexibility also produces greater binding rates in IDPs due to a larger target capture radius in solution.[17] Low affinity, coupled with high specificity, is an essential element for the transient protein-protein interactions required in cellular signaling. In many cases a single disordered protein can recognize many different binding partners with the same disordered region, often adopting different folds upon binding.[18] Understanding the mechanisms of disorder-mediated protein interactions requires information about the unstructured regions of the proteins and their disorder-to-order transitions.

Many different experimental techniques are available to characterize disordered proteins.[19-21] The techniques can be divided into those that provide only a global picture of the extent of disorder (*e.g.,* size exclusion chromatography, intrinsic fluorescence, circular dichroism, small angle x-ray scattering, atomic force microscopy) and a more limited list of those that provide localized to atomic-scale resolution (*e.g.*, NMR, limited proteolysis, FRET). In order to map regions of a protein that undergo coupled binding and folding, localized information is essential. One technique that has seen only limited application to the characterization of disordered proteins is amide H/D exchange. Amide hydrogens along the protein backbone become susceptible to exchange with deuterium from D_2O during structural fluctuations that disrupt amide hydrogen bonding.[22-24] By diluting a protein with a large excess of D_2O, amide H/D exchange can be used as a probe of the backbone flexibility. By measuring the rate at which deuterium is incorporated, the relative flexibility of different backbone segments can be determined. When H/D exchange is measured by NMR,[25-26] by mass spectrometry of proteolytic fragments[27] or by a top-down fragmentation approach[28] one can obtain localized information about the backbone flexibility of the protein of interest. The MS approach offers several advantages including the ability to work with large proteins at low protein concentrations in physiologically relevant buffers and the ability to detect all proteins in a complex without selective isotopic labeling.

There are many examples of disordered segments in otherwise well-folded proteins identified by H/D-MS.[29-30] The technique has also been used to characterize chemically unfolded proteins.[31-32] Yet the use of H/D-MS to obtain local flexibility data on natively unstructured proteins is considerably more limited.[33-35] The tyrosine kinase interacting protein (Tip) from *Herpesvirus saimiri* was predicted to be disordered and exhibited two characteristics of an intrinsically disordered protein: a far UV CD spectrum and aberrantly high apparent molecular weight by gel filtration chromatography.[33] However, the analysis of the H/D exchange kinetics was limited to rather long D_2O exposure times and the effects of binding to the

structured target, lymphocytic cell kinase (Lck), were not measured. To establish the utility of H/D exchange measurements for the characterization of IDPs and their interactions, we have undertaken an H/D exchange study of the nuclear co-activator binding domain of the CREB binding protein (CBP, CBP_MOUSE, residues 2059-2117) and the activation domain of the activator of thyroid and retinoid receptors (ACTR, NCOA3_HUMAN, residues 1018-1088). The interacting domains of CBP and ACTR are intrinsically disordered and undergo coupled binding and folding.[15] Both the isolated disordered domains and the folded complex have been well-characterized by a variety of biophysical techniques.[15, 36-39] Here, we have used isolated CBP and ACTR and their complex as model systems to explore the capability of the H/D exchange MS technique to characterize intrinsically disordered proteins and coupled binding and folding.

3.2 Materials and Methods

3.2.1 Materials

Ampicillin and isopropyl β-D-1-thiogalactopyranoside (IPTG) were obtained from Research Products International Corp (Mt. Prospect, IL). LB Medium used for cell cultures was prepared with Bacto Tryptone and Bacto Yeast Extract obtained from Becton, Dickinson, and Company (Franklin Lakes, NJ). Pepsin from porcine gastric mucosa, piperazine, sodium hydroxide, DL-dithiothreitol, bestatin hydrochloride, E-64, 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF), pepstatin A2, and deuterium oxide (99.9% D) were obtained from Sigma (St. Louis, MO). Benzonase nuclease was obtained from EMD Chemicals (Gibbstown, NJ). 99+% formic acid, used as an additive in prepared mobile phases, was obtained from Thermo Scientific (West Palm Beach, FL). Sodium phosphate monobasic, sodium phosphate dibasic, sodium hydroxide hydrochloric acid, boric acid, tris base, and sodium chloride, were obtained from Fisher Scientific (Hanover Park, IL). Optima LC/MS grade acetonitrile and water used in chromatography were both obtained from Fisher Scientific. POROS 20 AL column packing material was obtained from Applied Biosystems (Carlsbad, CA).

3.2.2 Protein Expression

CBP and ACTR domains were co-expressed from a pET22B co-expression vector[15] in *E. coli* strain BL21 Star(DE3) (Invitrogen, Carlsbad, CA). Glycerol stock was streaked onto agar plates containing 0.1 g L^{-1} ampicillin and grown overnight at 37 °C. All cell cultures were grown in LB medium (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl, pH 7.0) containing 0.1 g L^{-1} ampicillin. Cells were grown at 37 °C in a rotary shaker at 225 rpm. 25 mL starter cultures were grown overnight following inoculation with single colonies. 10 mL of the starter culture was used to inoculate a 500 mL culture. Cell cultures were induced with 1 mM IPTG after reaching OD_{600} = 1.0 and incubated for an additional 4 hours. The cells were pelleted at 5000 × g for 10 minutes at 4 ^{o}C and stored at −20 ^{o}C.

3.2.3 Protein Purification

Cell pellets were resuspended in 10 mL of lysis buffer (20 mM Tris, 1 mM EDTA, 1 g L^{-1} lysozyme, 5 units benzonase nuclease, pH 7.5) to which was added 250 µL of protease inhibitor cocktail per gram of pellet. The protease inhibitor cocktail contained 23 mM AEBSF, 2 mM bestatin, 0.3 mM pepstatin A2, 0.3 mM E-64. The resuspended pellet was incubated at room temperature for 10 minutes with gentle shaking. Cellular debris was removed by centrifugation (15000 × g for 10 min at 4 °C). Protein purification chromatography steps were performed using a fast protein liquid chromatography (FPLC) system (ÄKTAPrime plus, GE Healthcare, Piscataway, NJ) at 4 °C using 214 nm detection. While both proteins were co-expressed in *E. coli* and share similar molecular mass, ACTR and CBP have drastically different theoretical pI values of 4.2 and 11.1, respectively.[40] Because of this, we chose ion exchange to isolate the two proteins. ACTR was purified using a buffer exchange step followed by anion exchange and gel filtration at 4 ^{o}C. The lysis buffer was exchanged for anion exchange buffer (50 mM piperazine, pH 9.0) using a desalting column (HiPrep 26/10 Desalting, GE Healthcare). Collected protein fractions containing ACTR were identified by apparent mass using SDS-PAGE, pooled, and passed through an anion exchange column (HiTrap Q XL, GE Healthcare)

in 50 mM piperazine using a 0 to 1 M NaCl gradient over 70 min at 1 mL min^{-1}. Fractions containing ACTR were identified by SDS-PAGE, pooled, then loaded onto a gel filtration column (HiPrep 16/60 Sephacryl S-100 HR, GE Healthcare) and eluted into 10 mM phosphate, 50 mM NaCl buffer, pH 6.9. ACTR stock concentration was determined to be 14 μM using a BCA assay[41] with 595 nm detection and calibration with BSA standards (Thermo Fisher Scientific).

To purify CBP, we used a buffer exchange step followed by cation exchange and a second buffer exchange step at 4 ^{o}C. Clarified lysate was passed through the desalting column with 50 mM boric acid, pH 9.0, with 3 mM dithiothreitol (DTT) to prevent oxidation of methionine residues. Collected protein fractions containing CBP were identified by apparent mass using SDS-PAGE, pooled, and passed through cation exchange column (HiTrap SP FF, GE Healthcare) in 50 mM boric acid with 0.03 M DTT using a 0 to 1 M NaCl gradient over 70 min at 1 ml min^{-1}. Fractions containing ACTR were identified by SDS-PAGE, then loaded onto the desalting column and eluted into 10 mM phosphate, 50 mM NaCl, 3 mM DTT buffer, pH 6.9. DTT was removed from an aliquot of CBP stock by using acetone precipitation and reconstitution of protein in buffer without DTT prior to the protein assay. CBP stock concentration was determined to be 17 μM using a BCA assay[41] with 595 nm detection and calibration with BSA standards. Both ACTR and CBP protein identities were confirmed by mass spectrometry (see section *3.2.6*). Protein stocks were split into aliquots, frozen with liquid nitrogen, and stored at −74^{o}C.

3.2.4 H/D Exchange Labeling

Based on the reported dissociation constant of 34 nM for the ACTR/CBP complex,[15] a 5-fold molar excess of CBP will bind greater than 90% of ACTR, and *vice versa* under labeling conditions (see below). Four sample pools were prepared to represent both ACTR and CBP in free and complex states. A 1:4.9 molar ratio of ACTR to CBP was prepared by mixing 19 μL of ACTR stock was mixed with 81 μL CBP stock (complex). Under labeling conditions (1:20 dilution, see below), 94% of ACTR is bound. Similarly, a 1:4.7 molar ratio of CBP to ACTR was

prepared by mixing 15 µL of CBP stock with 85 µL ACTR stock (complex) yielding 93% bound CBP under labeling conditions. For the individual free proteins, an equal volume phosphate buffer was substituted for the binding partner. All pools were incubated at room temperature for an hour prior to deuterium labeling.

Samples were labeled by rapidly diluting 5 µL of a sample pool with 100 µL of 10 mM phosphate, 50 mM NaCl in D_2O, pD 6.9. After labeling intervals ranging between 5 sec and 12 hr, the labeling reaction was quenched with 17 µL 0.1 M HCl to decrease the pH to 2.6. Undeuterated controls were prepared for each sample pool by mixing 5 µL of a sample with 100 µL phosphate buffer in H_2O, pH 6.9, and 17 µL 0.1 M HCl. Totally deuterated controls were prepared for each sample by diluting 5 µL with 100 µL phosphate buffer in D_2O, pH 6.9, followed by quench after 24 hours with 17 µL 0.1 M HCl. After quenching, all samples were immediately flash frozen with liquid nitrogen and stored at −74 ^{o}C. Individual samples were thawed by hand in approximately 2 min just prior to analysis.

3.2.5 Chromatography

An Agilent Technologies 1200 series isocratic pump and a 12-port injector valve (Valco Instruments Co. Inc.) were used to inject the loaded sample through a 50 × 2.1 mm immobilized pepsin column prepared in-house as described previously[42] at 0.2 mL min^{-1} flow rate of 0.1% formic acid in water onto a self-packed C_{12} trap (10 mm × 1 mm, Jupiter 4 µm Proteo 90 Å, Phenomenex, Torrence, CA) for 2.5 minutes to desalt and concentrate the samples. An Agilent Technologies 1200 series binary pump was used to carry out a high performance liquid chromatography separation of peptic fragments. Solvent A was 0.1% formic acid in water and solvent B was 90% acetonitrile / 10% water / 0.1% formic acid. Peptic peptides were separated on a C_{12} column (50 mm × 1 mm Jupiter 4 µm Proteo 90 Å, Phenomenex) using water/acetonitrile gradient (10% to 35% B in 5 minutes at 50 µL min^{-1}). A thermoelectrically cooled system, constructed in-house, kept the loop, valve, columns, and solvent lines at 0 ^{o}C (see chapter 2).

3.2.6 MS Analysis

Following separation, peptides were eluted into the electrospray ionization source of a time of flight mass spectrometer (Model 6220, Agilent Technologies, Santa Clara, CA) operating in positive ion mode. Nitrogen at 325 ^{o}C and 10 L/min flow rate was used as the drying gas, with a nebulizer pressure of 30 psig. Fragmentor, skimmer, and capillary voltages were 150 V, 50 V, and 4000 V, respectively. Intact protein samples were checked using accurate mass measurements. These mass checks ensured the absence of any post-translational modifications to the proteins and measured the extent of oxidation of stored protein, which was determined to be low enough so as not to cause interference in further experiments.

3.2.7 Peptide Database

Once chromatographically resolved peptide mass spectral data were collected, MassHunter Qualitative Analysis with Bioconfirm software (Agilent Technologies, Santa Clara, CA) was used to build a peptide database. The MassHunter software provided a suite of methods to perform calculations and analyze mass-to-charge spectral features from mass spectrometer data. The “Find by Molecular Feature” and “Define and Match Sequences” methods allowed for peptide features to be isolated and for clipped *m/z* spectra and extracted ion chromatograms specific to each peptide to be extracted and graphed. Peak height minima, retention time windows, and peak spacing tolerances were defined by the user. A pre-calculated mass list or protein sequence with selected digest parameters can be used to include or exclude specific masses, with mass matching tolerances also defined by the user. We built a peptide database from spectral features that matched a non-specific protease digestion of our undeuterated target protein sequence. A mass matching tolerance of ±20 ppm was defined. Ambiguous peptide matches were identified using a rough CID technique, or nozzle-skimmer dissociation (NSD), as explained in section *1.6.3*. Protein sequences longer than those of CBP or ACTR require more complex MS/MS techniques, using instrumentation beyond TOF-MS, due

to the greater likelihood of an ambiguous assignment and the greater difficulty in resolving peptides.

Peptic peptides of undeuterated ACTR and CBP were identified by matching accurate mass measurements with a non-specific protease digest in the MassHunter software (version B.03.01). To resolve ambiguous near-isobaric peptide assignments, NSD product ion spectra were obtained by cycling the fragmentor voltage cycle every 0.5 seconds between 150 V, 200 V, and 250 V. Observed fragment ion peaks were matched against b/y ion fragmentation predictions (Protein Prospector, MS-Product, http://prospector.ucsf.edu). The final list of identified peptides can be found in table 3.1. Using the expected retention times and molecular feature extraction, mass spectra were extracted for each deuterated peptide sample. HX-Express software[43] was used to find the average mass of each peptide using the portion of the isotopic distribution that was at least 20% of the base peak. Deuterium uptake for each peptide at each time point was calculated as the mass difference between the average peptide mass and the average mass of the undeuterated control. Our deuterium uptake results are presented without back-exchange correction.[27]

Although the MassHunter software can extract specific mass-to-charge peak envelopes that match a given mass list, it was not able to identify deuterated peptide peak envelopes as peptide features, due to the presence of deuterons upsetting the distribution of *m/z* values within the envelope. A deuterated peptide mass database was built by calculating the expected average mass of an undeuterated peptide, then adding new entries that replace amide hydrogen from the formula of the peptide with deuterium atoms one by one. 1.006277 Da was added to the average mass for each deuteron replacement. In analyzing mass spectra, we measured the *m/z* values for the detected population of the peptide. The distribution of deuterons will not be uniform, so deuterium uptake was measured as the average mass increase of the population relative to the average mass of undeuterated control for EX2 exchange kinetics. While current software packages such as HDExaminer (Sierra Analytics,

Modesto, CA) are able to extract deuterated peptide features and obtain average mass increase, other software packages were used to perform this calculation separately from MassHunter. A manual survey of extracted spectra for deuterated peptides was also necessary to differentiate between actual peptides and small molecule interferences.

3.2.8 Intrinsic Exchange Calculations

The progress of intrinsic amide H/D exchange for unstructured peptides was calculated according to equation 1:

$$D(t) = \sum_{i=3}^{m} \left(1 - e^{-k_{\text{int}}^{(i)} t}\right) \tag{1}$$

where *m* is the number of residues, $k_{\text{int}}^{(i)}$ is the rate constant for exchange for residue *i* of the peptide.[44] Proline residues were assigned a rate constant of zero since proline has no amide hydrogen. Upon proteolysis, the amide hydrogen *i* = 1 becomes a primary amine and hence back-exchanges rapidly. Back-exchange at *i* = 2 is also generally quite rapid.[44-45] Thus, index *i* begins at 3, since there will be no deuterium remaining at *i* = 1,2. The calculations were implemented using an Excel spreadsheet (Microsoft, Redmond, WA) adapted from a spreadsheet available on Prof. Walter Englander's website (http://hx2.med.upenn.edu). Our adapted version is available upon request. The intrinsic exchange calculations were normalized to the maximum observed deuterium uptake.

3.2.9 Kinetic Analysis

In this work, we have quantified the kinetics of H/D exchange using the method described by Englander and co-workers[46] where the deuterium uptake has been fit to either a single or double stretched exponential function[46-47]:

$$\begin{aligned} D(t) &= N\left[1 - e^{-(\langle k \rangle t)^{\beta}}\right] \\ D(t) &= N_1\left[1 - e^{-(\langle k_1 \rangle t)^{\beta}}\right] + N_2\left[1 - e^{-(\langle k_2 \rangle t)^{\beta}}\right] \end{aligned} \tag{2}$$

D(*t*) is the extent of deuteration as a function of D_2O exposure time, *t*; *N* is the number of exchangeable amides; $\langle k \rangle$ represents the segment-averaged rate constant for exchange (either for intrinsic exchange $\langle k_{\mathrm{int}} \rangle$ or observed exchange $\langle k_{\mathrm{HX}} \rangle$); and *β* is an exponential stretching factor that accounts for the distribution in the intrinsic rates of H/D exchange of the individual amides. *β* was determined using a fit to the intrinsic exchange calculation (see equation 1). The fit to intrinsic exchange also yields the segment-averaged rate constant for intrinsic exchange, $\langle k_{\mathrm{int}} \rangle$. The use of the stretched exponential function requires fewer adjustable parameters than multi-exponential[48-49] fits commonly used to describe H/D exchange kinetics and thus the non-linear least-squares fit is more robust. In either case, since the resulting rate constants are segment-averaged, they are best considered "phenomenological" [50] rate constants. The ratio of the intrinsic exchange rate constant, $\langle k_{\mathrm{int}} \rangle$, to the measured rate constant, $\langle k_{\mathrm{HX}} \rangle$, can be used to estimate a segment-averaged protection factor, $\langle PF \rangle$, a measure of the slow-down in exchange induced by the structure of the protein:

$$\langle PF \rangle = \frac{\langle k_{\mathrm{int}} \rangle}{\langle k_{\mathrm{HX}} \rangle} \tag{3}$$

An average protection factor of unity indicates a segment that is completely unstructured, *i.e.,* the segment exchanges at its intrinsic rate, while a protection factor greater than 1 indicates the region is protected from H/D exchange. Because the measured H/D exchange is segment-averaged, these protections factors should not be interpreted thermodynamically as can be done for measurements obtained from single-residue resolution measurements.[26]

3.3 Results

3.3.1 Ligand Binding Equilibrium for ACTR and CBP

We initiated amide H/D exchange by rapidly diluting the individual proteins or the complex with a nineteen-fold excess of deuterated buffer. After intervals ranging from 5 s to 12

hr, the exchange reaction was quenched by acidification. Under quench conditions, labile D^+ is rapidly back-exchanged to H^+, but amide D back-exchanges much more slowly. To localize the exchange measurements to short linear segments of the proteins, the quenched samples were digested with pepsin and then deuterium uptake as a function of D_2O exposure time was quantified with mass spectrometry. To compare the free and bound states of the proteins, H/D exchange was carried out for each protein separately and for each protein in the presence of a five-fold molar excess of its binding partner (*i.e.,* 1:5 CPB:ACTR and 5:1 CBP:ACTR).

3.3.2 H/D-MS of ACTR and CBP

We quantified deuterium uptake by measuring the increase in the average mass of each segment. The mass increases for the segments shown in figure 3.1, plotted as a function of time, are shown in figure 3.2. Similar plots for all segments are provided in figure 3.3. For ease of viewing, the results are plotted in figure 3.2 on both a logarithmic scale and on a linear scale for the early exchange times. The maximum number of exchangeable amides is indicated by the limit of the vertical axis and the horizontal dashed lines on the plots represents the maximum observed exchange, measured after 12 hr of D_2O exposure. Also plotted in figure 3.2 is the intrinsic exchange for each segment calculated using equation 1. In the free state, both CBP 31-40 and CBP 43-59 are protected from exchange, as indicated by slower exchange than predicted using the intrinsic calculation. In the free state, ACTR 33-47 exchanges at the intrinsic rate, but ACTR 48-59 is slightly protected from exchange. In the complex, all four segments become substantially more protected from exchange.

To quantify the extent of protection, we derived segment-averaged protection factors, $\langle PF \rangle$, by fitting the kinetic data to stretched exponential or bi-exponential functions, see equation 2.[46] The results of the kinetic analysis are presented in table 3.2 and mapped onto the sequences of the proteins in figure 3.4. For free ACTR, all segments exchange at or near the rate predicted for intrinsic exchange (figure 3.2) with protection factors ranging from 1.0 to 2.4

(table 3.2, figure 3.4). In the deuterium uptake plots, however, there is a clear distinction between segments that exchange at the calculated intrinsic rate and those that exchange more slowly. Deuterium uptake plots for segments 1-12, 33-42, 33-47, 36-47, and 60-71 capture only the plateau of H/D exchange (see figures 3.2 and 3.3). In these segments, the observed H/D exchange is consistent with the intrinsic exchange calculation. In contrast, for the 13-22, 29-35, and 48-59 segments, our measurements at 5 and 10 sec of D_2O exposure capture the knee of the H/D exchange curve and show clearly that these segments exchange more slowly than predicted by intrinsic exchange. Unlike ACTR, all segments in free CBP exchange more slowly than predicted by the intrinsic exchange calculation. Notably, all CBP segments follow similar H/D exchange kinetics leading to a narrow range in the protection factors of between 4.1 and 5.7 (see table 3.2 and figure 3.4). Hence, there is very little contrast between segments that span the helical core and segments that span the tails.

In the complex, there are substantial changes in protection in many, but not all ACTR segments. Binding to CBP has no effect on exchange in the N- and C-terminal regions of ACTR (1-22 and 60-71). These regions exchange at the intrinsic rate, indicating that they remain unprotected. In the middle regions of ACTR (29-59), however, there is a substantial slow-down in H/D exchange. Several segments, for example ACTR 48-59, show clear bi-exponential H/D exchange kinetics (see figure 3.2). In the middle regions of ACTR, segment-averaged protection factors are as high as 180. The bimodal exchange kinetics allowed us to subdivide several of the segments,[48] wherein one part exchanges rapidly while other regions are more protected from exchange. We have estimated the number of amides involved in the bi-exponential fits as

$$\#\text{ of amides with } \langle PF \rangle_{1,2} = \left(\frac{N_{1,2}}{N_1 + N_2} \right) N_{\max} \qquad (4)$$

where N_{max} is the maximum number of detectable amide exchanges. This essentially represents a back-exchange correction[27] for the protection factors. By this means we have been able to differentiate regions of high and low protection within a single peptide segment. For example, the 33-42 region of ACTR in the complex encompasses both a highly protected ($\langle PF_2 \rangle$ = 180, 4 residues) and a weakly protected ($\langle PF_1 \rangle$ = 4.4, 4 residues) sub-segments. This segment spans four residues of helix 1 and four residues of the linker between helix 1 and helix 2. Although the resolution of H/D-MS data does not define the exact locations of the two amide populations, we have presented the data in figure 3.4 aligned with the known secondary structure of ACTR in the complex. Several other segments of ACTR spanning the helix 1 to helix 2 region (33-42, 33-47, and 36-47) also exhibit bi-exponential kinetics. The 62-71 segment spans both helix 3 and the unstructured C-terminal tail of ACTR, but there is no increase in protection in this region of ACTR in the complex.

The H/D exchange kinetics for CBP also change significantly upon complex formation. There are substantial increases in protection all along the length of CBP (table 3.2, figure 3.4). The largest increases in protection are in the Cα2 and Cα3 regions and in the linker connecting the two (30-38, 31-40, and 40-43). Here, segment-averaged protection factors increase from ~6 to 59-84. Protection also increases in the segments covering the N-terminal half of Cα1 and the C-terminal half of Cα3, however the increases are not as large because these segments also cover the N- and C-terminal tails, respectively.

3.4 Discussion

3.4.1 Application of H/D-MS to IDPs

Amide H/D exchange is used widely to characterize ligand-induced changes in the conformation and dynamics of proteins. Our objective in this work was to explore the capabilities and limitations of equilibrium H/D exchange to characterize intrinsically disordered proteins and reveal the details of their unstructured-to-structured transitions. The intrinsically

disordered interaction domains of CBP and ACTR[37] serve as a useful model system since both the free domains and their co-folded complex have been well-characterized.[15, 36-39]

On the basis of NMR, CD, and urea unfolding, CBP has been described as a molten globule, but its ensemble is sufficiently well-defined that structures have been solved.[38, 51] NMR secondary chemical shift measurements show well-formed helical structure in free CBP.[37] The recent ensemble of structures[38] shows that CBP has a fairly well-defined hydrophobic core comprising all three helices that aligns well with the structure of CBP in the complex,[15] but the N- and C-terminal regions are less well-defined. CBP is only marginally stable and undergoes two-state unfolding with $\Delta G_u = 6.1 \pm 0.4$ kJ mol^{-1} [38] from which it follows that approximately 8% of the CBP population occupies conformers belonging to the unfolded sub-population at 300 K. Absence of additional peaks in the NMR spectrum of free CBP led to the proposal that CBP inter-converts between different conformers slightly faster than intermediate exchange (10^4 s^{-1}).[38]

CBP and ACTR form a tight 1:1 complex (K_d = 34 nM) with well-defined tertiary structure. The structure consists of a helical bundle in which a series of leucine side chains in helix 1 of ACTR lie in a hydrophobic groove between helices 1 and 3 of CBP (see figure 3.5).[15] Secondary chemical shift measurements show dramatic increases in secondary structure across all of Aα1 and Aα2 but only a slight increase in secondary structure on the N-terminal side of helix Cα1 and the C-terminal end of helix Cα3.[37] Relaxation measurements show that both proteins become less dynamic in the complex.[37]

3.4.2 Mechanism of Amide H/D Exchange

The simplest representation of the mechanism of amide H/D exchange is the Linderstrøm-Lang scheme:

$$P^{\mathrm{cl}}_{\mathrm{H}^{(i)}} \underset{k^{(i)}_{\mathrm{cl}}}{\overset{k^{(i)}_{\mathrm{op}}}{\rightleftharpoons}} P^{\mathrm{op}}_{\mathrm{H}^{(i)}} \xrightarrow[\mathrm{D_2O}]{k^{(i)}_{\mathrm{int}}} P^{\mathrm{op}}_{\mathrm{D}^{(i)}} \qquad (5)$$

Within the context of understanding proteins as a rapidly inter-converting ensemble of conformations,[52] amide H/D exchange at residue *i* occurs when the protein enters an exchange-competent, so-called open conformation(s), collectively represented here as $P^{\mathrm{op}}_{\mathrm{H}^{(i)}}$. In the open conformers, the hydrogen bond for amide *i* is transiently broken through some sort of dynamic process. The opening reaction is followed by H/D exchange determined by the intrinsic rate constant, $k^{(i)}_{\mathrm{int}}$. The intrinsic rate of exchange depends on the amino acid sequence, temperature, and pH and is readily calculated from well-established kinetic parameters.[44] The extent to which the conformation of the protein slows exchange ($k^{(i)}_{\mathrm{op}}$ and $k^{(i)}_{\mathrm{cl}}$) can be quantified using the protection factor, *PF*:

$$PF^{(i)} = \frac{k^{(i)}_{\mathrm{int}}}{k^{(i)}_{\mathrm{HX}}} \tag{6}$$

where $k^{(i)}_{\mathrm{HX}}$ is the observed rate constant. An important distinction, that we shall return to later, between H/D measurements by NMR[53] and MS is that MS measurements are averaged across a short segment of the protein whereas NMR measurements are made with single amide resolution. To distinguish between single-amide and segment-averaging, we have adopted the convention of by applying a superscripted *i* to denote values at single-residue resolution and enclosing segment-averaged values within brackets.

3.4.3 The 31-40 Region of CBP as an Illustrative Example

We begin by using the 31-40 segment of CBP to illustrate how we have analyzed our data. Mass spectra of the 31-40 peptide of CBP show that the region undergoes rapid amide H/D exchange, reaching saturation within approximately 20 s of D_2O exposure (figure 3.1). In the CBP/ACTR complex, exchange in the 31-40 region of CBP is much slower, taking approximately 1 hr to reach saturation (figure 3.1). We quantified H/D exchange by measuring the center of mass of the isotopic distributions shown in figure 3.1. The results are plotted as a

function of D_2O exposure time in figure 3.2 (black and grey symbols in the lower left-hand panel). Following proteolysis under quench conditions (pH = 2.5, 0 °C), the first two residues undergo rapid D/H back-exchange so that measured deuteration at these two residues is zero. The 31-40 segment of CBP also includes one proline (see table 3.2). Hence, there are a total of seven residues in the 31-40 segment of CBP in which deuterium uptake can be measured. The maximum number of exchangeable amides is indicated in figure 3.2 by the upper limit of the vertical axis. However, even under quenched conditions some D/H back-exchange still occurs,[27, 45] so the horizontal dashed line on the plot represents the maximum observed exchange, +6.4 Da, measured after 24 hr of D_2O exposure. Hence, the deuterium recovery in this segment was 90%. Also plotted in figure 3.2 is the intrinsic exchange for the 31-40 segment of CBP (depicted as a dotted curve) calculated using equation 1. That the 31-40 segment of free CBP is protected from exchange is evident from its slower deuterium uptake relative to the intrinsic exchange. Formation of the CBP/ACTR complex further increases protection.

The kinetic analysis reveals that the 31-40 region of CBP is loosely structured and that the region undergoes coupled binding and folding in the presence of ACTR. The segment-averaged H/D exchange rate constants, $\langle k_{HX} \rangle$, for the 31-40 segment of CBP is 0.22 s^{-1} for free CBP. The rate constant decreases to 0.017 s^{-1} in the CBP/ACTR complex. The intrinsic rate constant for exchange, $\langle k_{int} \rangle$, for this segment of CBP is 1.3 s^{-1} (see table 3.2). Thus, in the 31-40 segment of CPB, the average protection factor increased from 5.7 for free CBP to 73 in the CBP/ACTR complex (table 3.2, figure 3.4). While the deuterium uptake data for bound CBP (see figure 3.2) suggests a stretched bi-exponential, such fits were unsatisfactory. It appears that the first exponential phase is nearly complete by the earliest time-point, 5 s, hence obtaining kinetic parameters for this first, rapid phase is difficult.

3.4.4 Limitations of the Kinetic Analysis

There are two important limitations to interpreting the results obtained using the analysis described above. The first limitation is that our measurements are segment-averaged. The simplified scheme for H/D exchange depicted in equation 5 must be revised to incorporate the averaging of multiple rate constants. Segment-averaged H/D exchange is not by itself difficult to describe mathematically and such simulations have been described.[24, 54] Rather, the limitations lie with the inversion of segment-averaged H/D exchange kinetic data to obtain mechanistic insight into the dynamics of the protein, as would be represented by a set of $k_{\mathrm{op}}^{(i)}$ and $k_{\mathrm{cl}}^{(i)}$ values. Such an inversion appears to be intractable since the kinetic models would be over-parameterized. Furthermore, it is not clear what the physical significance or reasonableness of using segment-averaged values would be. Nevertheless, semi-quantitative interpretation of H/D-MS data is still possible as we shall address in subsequent sections of this paper.

The second limitation with our approach is that the physical significance of the protection factor is unclear because for intrinsically disordered proteins, such as the interacting domains of CBP and ACTR, the conventional EX1 and EX2 kinetics limits may not hold. Under the conventional EX2 kinetic limit, $PF^{(i)}$ values are directly proportional to the ratio $k_{\mathrm{op}}^{(i)} \big/ k_{\mathrm{cl}}^{(i)}$, equal to the equilibrium constant for entry into the exchange competent state, $P_{\mathrm{H}^{(i)}}^{\mathrm{op}}$. Such a relationship follows directly from the EX2 kinetic limit that requires $k_{\mathrm{cl}}^{(i)} \gg k_{\mathrm{op}}^{(i)} \gg k_{\mathrm{int}}^{(i)}$. Under the EX2 limit, protection factors can be related to the Gibbs free energy cost for entering the exchange-competent state ($\Delta G_{\mathrm{op}}^{(i)} = RT \ln PF^{(i)}$). The EX1 kinetic limit requires $k_{\mathrm{int}}^{(i)} \gg k_{\mathrm{op}}^{(i)} \gg k_{\mathrm{cl}}^{(i)}$ and leads to $k_{\mathrm{HX}}^{(i)} = k_{\mathrm{op}}^{(i)}$. The EX1 and EX2 kinetic limits, though reasonable for structured proteins, do not necessarily hold for intrinsically disordered proteins. For example, a mostly unstructured protein

that inter-converts between protected and unprotected states with a kinetic limit of $k_{\mathrm{op}}^{(i)} > k_{\mathrm{cl}}^{(i)} \gg k_{\mathrm{int}}^{(i)}$ is easily envisioned. Thus for disordered proteins it may be necessary to consider the full treatment of H/D exchange kinetics[22] without assuming either EX1 or EX2 kinetic limits (*vide infra*).

3.4.5 Residual Structure in ACTR

For free ACTR, all segments exchange at or near the rate predicted for intrinsic exchange (figure 3.2), yielding protection factors ranging from 1.0 to 2.4 (table 3.2, figure 3.4). Limited protection in ACTR is consistent with previous measurements by NMR and circular dichroism showing that free ACTR is unstructured.[15, 37] In ACTR segments covering residues 1-12 and 33-47 our H/D exchange measurements only captured the plateau of the H/D exchange (see figure 3.2), hence the protection factor of 1.6 in the 33-47 region is attributable to experimental error. However, in the segments 13-22, 29-35, and 48-59, our measurements at 5 and 10 sec of D_2O exposure captured the knee of the H/D exchange curve (see figure 3.2). In these regions, the differences between observed deuterium uptake and the deuterium uptake, predicted on the basis of intrinsic exchange, do appear to be significant (figure 3.2, dotted curve *vs.* black circles). Hence, protection in these regions is significantly greater than unity. Based on extensive NMR relaxation data and secondary chemical shifts, Ebert and co-workers[37] concluded that ACTR had the characteristics of an unstructured random coil. More recently, however, the conformation of free ACTR was interrogated using CD, small-angle x-ray scattering (SAXS), and multidimensional NMR.[39] In particular, secondary chemical shift measurements using reference shifts obtained from ACTR in urea indicated that a small percentage of alpha-helical propensity was present in the three regions of ACTR that form helices in complex with CBP, particularly in the Aα1. Thus, our H/D-MS measurements support the existence of transient helical character in an otherwise unstructured protein.

While our H/D exchange data shows evidence of residual structure, we are unable to determine the timescale for the formation/loss of this structure. For a highly unstructured protein such as ACTR, it is reasonable to project that the life-time of the exchange competent state is long relative to the time-scale of H/D exchange (*i.e.,* $k_{\text{int}}^{(i)} > k_{\text{cl}}^{(i)}$). Furthermore, we expect that such a relationship would extend to the segment level: $\langle k_{\text{int}} \rangle > \langle k_{\text{cl}} \rangle$. At the extreme, the rate constant for closing would be zero and the segment would exchange at the intrinsic rate. Under this limit, complete exchange will take place before the segment has time to adopt an exchange-protected conformation. For segment-averaged kinetics, exchange under the limit $\langle k_{\text{int}} \rangle > \langle k_{\text{cl}} \rangle$ would produce a bimodal isotope profile in which a population of a fully deuterated segment grows in at the expense of the undeuterated population. Such an isotopic profile would resemble the isotopic profile that can be observed under the EX1 kinetic limit,[24, 54-57] even though the EX1 kinetic limit would not necessarily apply here. The lengthy dead-time of our experiments (5 sec) prevented us from detecting bimodal isotopic profiles since all of the ACTR peptides are nearly fully deuterated at 5 sec. Equilibrium H/D exchange measured on the millisecond timescale using quench-flow[58-59] may yield valuable insight into the nature of the transient helical structure in ACTR.

3.4.6 Limited, Uniform Protection Reveals Conformational Inter-conversion by Free CBP

Unlike ACTR, all segments in free CBP exchange more slowly than predicted for intrinsic exchange (see figure 3.2). However, the protection factors span a narrow range between 4.1 and 5.7 (see figure 3.4 and table 3.2). One interpretation could be that the increased protection means that CBP is simply more structured than ACTR. Indeed, free CBP has been described as a molten globule on the basis of NMR, CD, and urea unfolding, and is sufficiently well-ordered that its structure has been solved.[15, 38, 51] Secondary chemical shift NMR measurements clearly show well-formed helical structure in CBP.[37] However, the nearly-

uniform protection factors that we observe are not consistent with a molten globular state in which three alpha helical segments are connected by flexible linkers. In addition, the magnitude of the protection factors is too small. For example, when apolipoprotein A-I was analyzed by H/D-MS using the segment-averaged stretched exponential fitting, alpha helical segments of were found to have segment-averaged protection factors ranging from 10^1 to greater than 10^4.[46] In molten globular apomyoglobin at pH 4.2, protection factors ranged from 5-200 in helical regions while in the native state at pH 6.0, protection factors ranged between 10^3 to greater than 10^5.[60] Hence, the protection factors for CBP of 4-6 appear to be inconsistent with the available structural data for CBP.

The inconsistency between our measurements and the available structural data can be resolved if we interpret our data in terms of the more general two-process model of H/D exchange.[61-62] According to the generalized two-process model, the protein conformational ensemble consists of two macroscopic states: a folded population (*F*) and an unfolded population (*U*) that inter-convert with global folding and unfolding rate constants of k_F and k_U, respectively. Within these macroscopic states are a collection of rapidly inter-converting conformers, some protected from exchange (denoted with the superscripted cl) and some locally exchange competent (denoted by the superscripted op). The exchange process is depicted in the following scheme and also in figure 3.6.

$$\begin{array}{ccc} \boxed{U^{\mathrm{cl}}_{\mathrm{H}^{(i)}} \underset{k^{(i)}_{\mathrm{cl},U}}{\overset{k^{(i)}_{\mathrm{op},U}}{\rightleftharpoons}} U^{\mathrm{op}}_{\mathrm{H}^{(i)}}} & \xrightarrow[\mathrm{D_2O}]{k^{(i)}_{\mathrm{int}}} & U^{\mathrm{op}}_{\mathrm{D}^{(i)}} \\ k_U \uparrow\downarrow k_F & & \\ \boxed{F^{\mathrm{cl}}_{\mathrm{H}^{(i)}} \underset{k^{(i)}_{\mathrm{cl},F}}{\overset{k^{(i)}_{\mathrm{op},F}}{\rightleftharpoons}} F^{\mathrm{op}}_{\mathrm{H}^{(i)}}} & \xrightarrow[\mathrm{D_2O}]{k^{(i)}_{\mathrm{int}}} & F^{\mathrm{op}}_{\mathrm{D}^{(i)}} \end{array} \qquad (7)$$

In the two-process model, exchange at amide *i* can occur either through an exchange competent conformer of the folded $\left(F^{\mathrm{op}}_{\mathrm{H}^{(i)}}\right)$ or unfolded $\left(U^{\mathrm{op}}_{\mathrm{H}^{(i)}}\right)$ ensemble. As well-illustrated though exchange flux calculations by Henkels and Oas,[63] the observed H/D exchange kinetics

can be dominated by very poorly populated states when the rate of exchange by these states is much faster than exchange by majority states. Since the usual kinetic limits for H/D exchange may not be relevant for intrinsically disordered proteins, the full kinetic treatment of H/D exchange for the two-process model[62] is required. In the limit of $k_F \gg k_{int}$, the observed rate constant for exchange $k_{HX}^{(i)}$ will be:

$$k_{HX}^{(i)} \approx \left[f_F f_F^{op(i)} + f_U f_U^{op(i)} \right] k_{int}^{(i)} \quad (8)$$

where f_F and f_U represent the fraction of molecules occupying the folded and unfolded ensembles, respectively ($f_F + f_U = 1$); and $f_F^{op(i)}$ and $f_U^{op(i)}$ represent the fractions of exchange-competent molecules in the folded and unfolded conformational ensembles.. It follows that

$$PF^{(i)} \approx \left[f_F f_F^{op(i)} + f_U f_U^{op(i)} \right]^{-1} \quad (9)$$

It is obvious by inspection that the protection factors of the folded and unfolded states are the reciprocals of the fractions of molecules in exchange competent conformations (*i.e.*, $PF_{F,U} = 1 / f_{F,U}^{op(i)}$). At the end of this section we will establish the relevance of the kinetic limit $k_F \gg k_{int}$ that we have invoked here.

Equations 8 and 9 and figure 3.6 provide an intuitive description of H/D exchange kinetics for IDPs. The first term represents the fraction of molecules in the folded ensemble that are in exchange-competent conformations and the second term represents the fraction of molecules in the unfolded ensemble that are in exchange-competent conformations. Hence, the H/D exchange flux[63] occurs through two separate channels: exchange-competent conformers of the folded ensemble and exchange-competent conformers of the unfolded ensemble. We can anticipate that for a marginally stable protein like CBP, the second term in equation 9 may be of the same order of magnitude or even larger than the first. In other words, the CBP conformational ensemble could consist of inter-converting molten globular and unfolded ensembles with most unfolded conformers being exchange-competent (*i.e.,* unprotected). For

illustrative purposes, consider a protein that is 95% molten globular with $f_F^{\mathrm{op}(i)} = 0.001$ (*PF*= 1000) and 5% unfolded $f_U^{\mathrm{op}(i)} = 0.80$ (*PF* = 1.25). Under this scenario, the observed protection factor of 24 is primarily dictated by the unstructured state. While there is a relatively low probability of being in the unfolded state (f_U = 0.05), exchange flux[63] through this pathway is much greater because most of the conformers are exchange competent ($f_U^{\mathrm{op}(i)} = 0.80$), as denoted by the heavy arrow for exchange flux in figure 3.6. Although extension of the two-process model to segment-averaged exchange is mathematically intractable, we assert that, qualitatively at least, the segment-averaged H/D exchange works in a similar manner.

Based on the two-process model, the small, near-uniform protection factors for CBP shown in figure 3.4 are consistent with a well-folded molten globular form of CBP that is in equilibrium with an unstructured state (see figure 3.6). Both the apparent lack of secondary structure and the small magnitude of the protection factors arise because the majority of H/D exchange occurs during intervals when CBP molecules are in the unfolded conformational ensemble $\left(U_{\mathrm{H}^{(i)}}^{\mathrm{op}}\right)$. Our observations are consistent with recent work by Kjaergaard and co-workers who characterized free CBP using NMR, CD, and small angle X-ray scattering.[38] The authors determined that free CBP is marginally stable and undergoes two-state unfolding (ΔG_u = 6.1 ± 0.4 kJ mol^{-1}, K_U = 0.085), hence under native conditions, ~8% of the molecules are in unfolded conformations at 300 K. The accepted model of the molten globular state has well-defined secondary structure, but fluctuating tertiary structure. High protection factors have often been taken as evidence of the existence of well-defined secondary structure in molten globules. The preceding treatment offers an explanation for why our observed protection factors are so much less than those expected for a molten globule and they support the conclusions of Kjaergaard *et al.* that CBP occupies both a molten globular state and an unfolded state.[38]

The absence of bimodal isotopic profiles in spectra of the CBP segments indicates that the unstructured conformers are only transiently unprotected such that either the kinetic limits $k_f \gg k_{\text{int}}^{(i)}$ and/or $k_{\text{cl},U}^{(i)} \gg k_{\text{int}}^{(i)}$ apply. A bimodal isotopic profile would be evident in our spectra if the limits did not hold since unstructured peptide segments would become fully deuterated during its time in the unprotected state. Since there is no evidence of such an EX1-like isotope profile in any of the CBP peptides (see figure 3.1), we conclude that CPB exchanges under one or both of the limits $k_f \gg k_{\text{int}}^{(i)}$ and/or $k_{\text{cl},U}^{(i)} \gg k_{\text{int}}^{(i)}$. Hence, the lifetime of the unstructured state must be msec or shorter (*i.e.,* much less than k_{int}^{-1} which is on the order of 1 s at pD 7[44]). The transient nature is consistent with the exchange constant ($k_F + k_U$) of 10^4 s^{-1} estimated by NMR[38] and supports our use of the kinetic limit $k_F \gg k_{\text{int}}$ in equations 8 and 9.

3.4.7 Formation of the ACTR-CBP Complex

Our H/D exchange data can be used to map regions of ACTR and CBP that undergo coupled binding and folding in the complex. For comparison, we have mapped the observed protection factors onto the primary sequences of the proteins and the secondary structural elements in the CBP/ACTR complex[15] (see figure 3.4). Binding to CBP has no effect on exchange kinetics in the N- and C-terminal regions of ACTR (1-22 and 60-71). Even in the complex, these regions exchange at the intrinsic rate, indicating that these regions remain disordered (see figures 3.2 and 3.4, top). In the middle regions of ACTR (29-59), however, there is a substantial slow-down in H/D exchange. Several segments, for example ACTR 48-59, show clear bimodal H/D exchange kinetics (see figure 3.2). In the middle regions of ACTR, segment-averaged protection factors are as high as 180. The bimodal exchange kinetics allowed us to subdivide several of the segments,[48] wherein one part exchanges rapidly while other regions are more protected from exchange. For example, the 33-42 region of ACTR in the CBP/ACTR complex encompasses both a highly protected sub-segment ($\langle PF_2 \rangle$ = 180, 4

residues) and a weakly protected sub-segment ($\langle PF_1 \rangle$ = 4.4, 4 residues). Hence, the observed protection factors are consistent with this segment spanning four residues of helix 1 and four residues of the linker between helix 1 and helix 2. Although the resolution of our H/D-MS data does not allow us to define the exact locations of the two amide populations, we have presented the data in figure 3.4 aligned with the previously-assigned secondary structure of ACTR in the complex. Several other segments of ACTR spanning the helix 1 to helix 2 region (33-42, 33-47, and 36-47) also exhibit bi-exponential kinetics. The 62-71 segment spans both helix 3 and the unstructured C-terminal tail of ACTR, but there is no observed increase in protection in this region of ACTR in the complex. We attribute this to both the difficulty of resolving bimodal kinetics involving only two residues and also to helix fraying at the end of helix three.

H/D exchange kinetics for CBP also change significantly when CBP binds to ACTR. There are substantial increases in protection all along the length of CBP (table 3.2, figure 3.5). The largest increases in protection are in the Cα2 and Cα3 regions and in the linker connecting the two (30-38, 31-40, and 40-43). Here, segment-averaged protection factors increase from 5.5-5.7 to 59-84. Protection also increases in the segments covering the N-terminal half of Cα1 and the C-terminal half of Cα3, however the increases are not as large because these segments also cover the N- and C-terminal tails, respectively. The observed increases in protection are consistent with NMR measurements showing a decrease in backbone motions in CBP upon binding to ACTR.[37]

Figures 3.5 summarizes the effects of coupled binding and folding on the exchange kinetics of ACTR. The figure shows the fold-changes in segment-averaged protection, $\langle PF_{\text{complex}} \rangle / \langle PF_{\text{free}} \rangle$, that accompany complex formation. The observed fold-increases in protection in ACTR in the complex range from ~1 in the terminal regions to ~100 in helix 2 of ACTR (see table 3.2 and figure 3.4). We have found that the C-terminal half of helix Aα1 and all of helix Aα2 become highly protected in the complex (see figure 3.4). Helices Aα1 and Aα2 lie

in a hydrophobic groove formed by CBP helices Cα1 and Cα3.[15] NMR measurements of secondary chemical shifts showed dramatic increases in secondary structure across all of Aα1 and Aα2.[37] We found that the linker between Aα1 and Aα2 remains unprotected from H/D exchange in the ACTR/CBP complex indicating that this region remains unstructured, consistent with NMR relaxation measurements showing that the linker is highly dynamic in the complex.[37] The largest increase in protection was in the N-terminal half of Aα3 (55-59) while the 60-71 segment that covers the last two residues of Aα3 (62-63) has no increase in protection. Only the N-terminal half of (residues 55-59 in our nomenclature) had significant increases in chemical shift upon complex formation.[37]

The effects of complex formation on protection in CBP are summarized in figure 3.5. Compared with ACTR, our H/D exchange data shows only modest increases in protection (9-17-fold) in the complex. This effect can be attributed to higher protection in free CBP that increases the denominator in the protection ratio. Our H/D exchange measurements show that the C-terminal half of Cα2 and the N-terminal half of Cα3 (residues 30-43) become the most highly protected regions of bound CBP. There are only small increases in protection in bound CBP in the segments that span the N-terminal of Cα1 and the C-terminal of Cα3 (1-14 and 43-59, respectively). Two factors limit the increase in protection in these regions in the complex. First, fraying of the ends of the helices is evident in the ensemble of structures of free CBP.[38] Second, the 1-14 and 43-59 segments encompass unstructured terminal tails. Protection in the 17-30 segment of bound CBP is greater than in the tail regions, but less than in the central core of CBP. Since the 17-30 segment spans the long linker between Cα1 and Cα2 and the flexible polyQ region, limited increase in protection here is consistent with flexibility in this region of bound CBP. Our results may at first appear to contradict the secondary chemical shift data[37] showing that CBP was well-structured and that binding to ACTR induced negligible changes in secondary chemical shifts across all of Cα1 and Cα2, with minor increases in secondary chemical shift only in Cα3. This apparent contradiction arises because H/D exchange

measurements on free CBP reflect the inter-conversion between folded and unstructured conformers (see previous section). Without this inter-conversion, the H/D exchange protection in free CBP would be much higher and the fold-increases upon binding would be much smaller.

3.5 Conclusions

In this work, we have tested the capabilities of H/D-MS on the model IDPs CBP and ACTR and their synergistically folded complex. Our results demonstrate that a great deal can be learned about IDPs using the technique. We have demonstrated the ability to detect traces of residual structure in the unstructured protein ACTR. In the case of molten globular CBP, limited, but uniform protection revealed rapid inter-conversion between ensembles of relatively well-protected conformers and unstructured conformers. An important point that emerges from this work is that the seemingly rapid, uniform exchange, as we observed with CBP, could be easily misinterpreted as indicating an unstructured protein. With the folded complex, we have identified short segments of CBP and ACTR that undergo coupled binding and folding and our results show that the inter-conversion between free and bound forms of the proteins is fast on the H/D exchange timescale. Overall, the picture of ACTR and CBP that emerges from our H/D exchange measurements is consistent with previous biophysical measurements.

Our work also underscores the point that the effects of ligand binding need to be interpreted carefully.[63-64] This is particularly important with IDPs since the free form of the IDP will often be weakly protected. A small population of rapidly exchanging free IDP can dilute the protection observed in the complex. Limited protection in the complex might be misinterpreted as a loosely folded complex.

H/D-MS holds the promise of making good progress in the difficult terrain of intrinsically disordered proteins. While peptide-level H/D exchange lacks the single amide resolution afforded by NMR, mass spectrometry offers two distinct advantages. First, mass spectrometry affords the ability to measure H/D exchange using only picomole quantities of protein under dilute solution conditions. Given the difficulty that many laboratories have encountered in the

expression and purification of recombinant IDPs, there may simply never be enough protein for NMR or other biophysical methods. The second advantage of mass spectrometry is that protein mass is nearly unlimited. Recent work with ~150 kDa monoclonal antibodies[65] and viral assemblies[66] has clearly demonstrated the ability of H/D-MS to handle large proteins. In large proteins, identification of short unstructured molecular recognition features (MoRFs) that undergo coupled binding and folding[67] would be difficult by any other technique.

Our work also reveals a number of important unmet needs for continued development of H/D-MS as a tool to probe intrinsically disordered proteins. While the stretched exponential equation[46-47] does seem to empirically describe H/D exchange kinetics, the theoretical basis for its use and for the interpretation of the derived parameters is currently lacking. In the absence of a good model to justify use of the stretched exponential, we believe derivation of thermodynamic values from segment-averaged protection should be approached with caution. There is clearly a need for an extension of the two-process model of H/D exchange to explicitly address both segment-averaging and the departure of IDPs from exchange under the conventional kinetic limits. We speculate that such an extension might lend new insight into the analysis of H/D exchange kinetics in IDPs. Finally, measurements at shorter dead times (msec-sec) would also be beneficial for characterizing residual structure in unstructured proteins such as ACTR.

References

1. Sigler, P. B., Acid blobs and negative noodles. *Nature* **1988,** *333* (6170), 210-212.
2. Kriwacki, R. W.; Hengst, L.; Tennant, L.; Reed, S. I.; Wright, P. E., Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93* (21), 11504-11509.
3. Garner, E.; Cannon, P.; Romero, P.; Obradovic, Z.; Dunker, A. K., Predicting Disordered Regions from Amino Acid Sequence: Common Themes Despite Differing Structural Characterization. *Genome Informatics* **1998,** *9*, 201-213.
4. Wright, P. E.; Dyson, H. J., Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999,** *293* (2), 321-331.
5. Uversky, V. N.; Gillespie, J. R.; Fink, A. L., Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **2000,** *41* (3), 415-427.
6. Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradovic, Z., Intrinsic disorder and protein function. *Biochemistry* **2002,** *41* (21), 6573-6582.
7. Uversky, V. N., Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* **2002,** *11* (4), 739-756.
8. Dyson, H. J.; Wright, P. E., Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell. Biol.* **2005,** *6* (3), 197-208.
9. Tompa, P.; Fuxreiter, M.; Oldfield, C. J.; Simon, I.; Dunker, A. K.; Uversky, V. N., Close encounters of the third kind: disordered domains and the interactions of proteins. *BioEssays* **2009,** *31* (3), 328-335.
10. Gsponer, J.; Madan Babu, M., The rules of disorder or why disorder rules. *Prog. Biophys. Mol. Biol.* **2009,** *99* (2-3), 94-103.
11. Uversky, V. N.; Dunker, A. K., Understanding protein non-folding. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **2010,** *1804* (6), 1231-1264.
12. He, B.; Wang, K.; Liu, Y.; Xue, B.; Uversky, V. N.; Dunker, A. K., Predicting intrinsic disorder in proteins: an overview. *Cell Res* **2009,** *19* (8), 929-949.
13. 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.
14. Xie, H.; Vucetic, S.; Iakoucheva, L. M.; Oldfield, C. J.; Dunker, A. K.; Uversky, V. N.; Obradovic, Z., Functional anthology of intrinsic disorder. 1. biological processes and functions of proteins with long disordered regions. *Journal of Proteome Research* **2007,** *6* (5), 1882-1898.
15. Demarest, S. J.; Martinez-Yamout, M.; Chung, J.; Chen, H.; Xu, W.; Dyson, H. J.; Evans, R. M.; Wright, P. E., Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **2002,** *415* (6871), 549-553.
16. Lacy, E. R.; Filippov, I.; Lewis, W. S.; Otieno, S.; Xiao, L.; Weiss, S.; Hengst, L.; Kriwacki, R. W., p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol. Biol.* **2004,** *11* (4), 358-364.

17. Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G., Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97* (16), 8868-8873.

18. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Intrinsically disordered proteins in human diseases: Introducing the D-2 concept. *Annu. Rev. Biophys.* **2008,** *37*, 215-246.

19. Receveur-Bréchot, V.; Bourhis, J.-M.; Uversky, V. N.; Canard, B.; Longhi, S., Assessing protein disorder and induced folding. *Proteins* **2006,** *62* (1), 24-45.

20. Eliezer, D., Biophysical characterization of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2009,** *19* (1), 23-30.

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

22. Hvidt, A.; Nielsen, S. O., Hydrogen exchange in proteins. *Adv. Protein Chem.* **1966,** *21*, 287-385.

23. Englander, S. W.; Kallenbach, N. R., Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quarterly Review of Biophysics* **1984,** *16* (4), 521-655.

24. Konermann, L.; Tong, X.; Pan, Y., Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **2008,** *43* (8), 1021-1036.

25. 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.

26. Bai, Y.; Sosnick, T. R.; Leland, M.; Englander, S. W., Protein Folding Intermediates: Native-State Hydrogen Exchange. *Science* **1995,** *269* (5221), 192-197.

27. Zhang, Z.; Smith, D. L., Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* **1993,** *2* (4), 522-531.

28. Pan, J.; Han, J.; Borchers, C. H.; Konermann, L., Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Amer. Chem. Soc.* **2009,** *131*, 12801–12808.

29. Sharma, S.; Zheng, H.; Huang, Y. J.; Ertekin, A.; Hamuro, Y.; Rossi, P.; Tejero, R.; Acton, T. B.; Xiao, R.; Jiang, M.; Zhao, L.; Ma, L.-C.; Swapna, G. V. T.; Aramini, J. M.; Montelione, G. T., Construct optimization for protein NMR structure analysis using amide hydrogen/deuterium exchange mass spectrometry. *Proteins* **2009,** *76* (4), 882-894.

30. Pantazatos, D.; Kim, J. S.; Klock, H. E.; Stevens, R. C.; Wilson, I. A.; Lesley, S. A.; Woods, V. L., Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101* (3), 751-756.

31. Mazon, H.; Marcillat, O.; Forest, E.; Vial, C., Denaturant sensitive regions in creatine kinase identified by hydrogen/deuterium exchange. *Rapid Commun. Mass Spectrom.* **2005,** *19* (11), 1461-1468.

32. Raza, A. S.; Dharmasiri, K.; Smith, D. L., Identification of non-covalent structure in apocytochrome c by hydrogen exchange and mass spectrometry. *J. Mass Spectrom.* **2000,** *35* (5), 612-617.

33. Mitchell, J. L.; Trible, R. P.; Emert-Sedlak, L. A.; Weis, D. D.; Lerner, E. C.; Applen, J. J.; Sefton, B. M.; Smithgall, T. E.; Engen, J. R., Functional characterization and conformational analysis of the *Herpesvirus saimiri* Tip-C484 protein. *J. Mol. Biol.* **2007,** *366* (4), 1282-1293.

34. Hansen, J. C.; Wexler, B. B.; Rogers, D. J.; Hite, K. C.; Panchenko, T.; Ajith, S.; Black, B. E., DNA Binding Restricts the Intrinsic Conformational Flexibility of Methyl CpG Binding Protein 2 (MeCP2). *J. Biol. Chem.* **2011,** *286* (21), 18938-18948.

35. Croy, C. H.; Bergqvist, S.; Huxford, T.; Ghosh, G.; Komives, E. A., Biophysical characterization of the free IκBα ankyrin repeat domain in solution. *Protein Sci.* **2004,** *13* (7), 1767-1777.

36. Demarest, S. J.; Deechongkit, S.; Dyson, H. J.; Evans, R. M.; Wright, P. E., Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci.* **2004,** *13* (1), 203-210.

37. Ebert, M.-O.; Bae, S.-H.; Dyson, H. J.; Wright, P. E., NMR Relaxation Study of the Complex Formed Between CBP and the Activation Domain of the Nuclear Hormone Receptor Coactivator ACTR *Biochemistry* **2008,** *47* (5), 1299-1308.

38. Kjaergaard, M.; Teilum, K.; Poulsen, F. M., Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (28), 12535-12540.

39. Kjaergaard, M.; Nørholm, A.; Hendus-Altenburger, R.; Pedersen, S. F.; Poulsen, F. M.; Kragelund, B. B., Temperature-dependent structural changes in intrinsically disordered proteins: Formation of α-helices or loss of polyproline II? *Protein Sci.* **2010,** *19* (8), 1555-1564.

40. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A., ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **2003,** *31* (13), 3784-3788.

41. Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C., Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **1985,** *150* (1), 76-85.

42. Wang, L.; Pan, H.; Smith, D. L., Hydrogen exchange-mass spectrometry: optimization of digestion conditions. *Mol. Cell. Proteomics* **2002,** *1* (2), 132-138.

43. Weis, D. D.; Engen, J. R.; Kass, I. J., Semi-Automated Data Processing of Hydrogen Exchange Mass Spectra Using HX-Express. *J. Am. Soc. Mass. Spectrom.* **2006,** *17* (12), 1700-1703.

44. Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W., Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 75-86.

45. Connelly, G. P.; Bai, Y.; Jeng, M.-F.; Englander, S. W., Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 87-92.

46. Chetty, P. S.; Mayne, L.; Lund-Katz, S.; Stranz, D.; Englander, S. W.; Phillips, M. C., Helical structure and stability in human apolipoprotein A-I by hydrogen exchange and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (45), 19005-19010.

47. Dewey, T. G., Fractal analysis of proton exchange kinetics in lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91* (25), 12101-12104.

48. Resing, K.; Hoofnagle, A.; Ahn, N., Modeling deuterium exchange behavior of ERK2 using pepsin mapping to probe secondary structure. *J. Am. Soc. Mass. Spectrom.* **1999,** *10* (8), 685-702.

49. Codreanu, S. G.; Ladner, J. E.; Xiao, G.; Stourman, N. V.; Hachey, D. L.; Gilliland, G. L.; Armstrong, R. N., Local Protein Dynamics and Catalysis: Detection of Segmental Motion Associated with Rate-Limiting Product Release by a Glutathione Transferase. *Biochemistry* **2002,** *41* (51), 15161-15172.

50. 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.

51. Lin, C. H.; Hare, B. J.; Wagner, G.; Harrison, S. C.; Maniatis, T.; Fraenkel, E., A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Mol. Cell* **2001,** *8* (3), 581-590.

52. Hilser, V. J.; Freire, E., Structure-based Calculation of the Equilibrium Folding Pathway of Proteins. Correlation with Hydrogen Exchange Protection Factors. *J. Mol. Biol.* **1996,** *262* (5), 756-772.

53. Milne, J. S.; Mayne, L.; Roder, H.; Wand, A. J.; Englander, S. W., Determinants of protein hydrogen exchange studied in equine cytochrome c. *Protein Sci.* **1998,** *7* (3), 739-745.

54. Xiao, H.; Hoerner, J. K.; Eyles, S. J.; Dobo, A.; Voigtman, E.; Mel'čuk, A. I.; Kaltashov, I. A., Mapping protein energy landscapes with amide hydrogen exchange and mass spectrometry: I. A generalized model for a two-state protein and comparison with experiment. *Protein Sci.* **2005,** *14* (2), 543-557.

55. Englander, S. W.; Mayne, L.; Bai, Y.; Sosnick, T. R., Hydrogen exchange: The modern legacy of Linderstrøm-Lang. *Protein Sci.* **1997,** *6* (5), 1101-1109.

56. Kreshuk, A.; Stankiewicz, M.; Lou, X.; Kirchner, M.; Hamprecht, F. A.; Mayer, M. P., Automated detection and analysis of bimodal isotope peak distributions in H/D exchange mass spectrometry using HeXicon. *Int. J. Mass spectrom. In Press, Corrected Proof*.

57. Weis, D. D.; Wales, T. E.; Engen, J. R.; Hotchko, M.; Ten Eyck, L. F., Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J. Am. Soc. Mass. Spectrom.* **2006,** *17* (11), 1498-1509.

58. Dharmasiri, K.; Smith, D. L., Mass Spectrometric Determination of Isotopic Exchange Rates of Amide Hydrogens Located on the Surfaces of Proteins. *Anal. Chem.* **1996,** *68*, 2340-2344.

59. Truhlar, S. M. E.; Croy, C. H.; Torpey, J. W.; Koeppe, J. R.; Komives, E. A., Solvent Accessibility of Protein Surfaces by Amide H/2H Exchange MALDI-TOF Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2006,** *17* (11), 1490-1497.

60. Hughson, F. M.; Wright, P. E.; Baldwin, R. L., Structural characterization of a partly folded apomyoglobin intermediate. *Science* **1990,** *249*, 1544-1548.

61. Woodward, C. K.; Hilton, B. D., Hydrogen isotope exchange kinetics of single protons in bovine pancreatic trypsin inhibitor. *Biophys. J.* **1980,** *32* (1), 561-575.

62. Qian, H.; Chan, S. I., Hydrogen exchange kinetics of proteins in denaturants: a generalized two-process model. *J. Mol. Biol.* **1999,** *286* (2), 607-616.

63. Henkels, C. H.; Oas, T. G., Ligation-State Hydrogen Exchange: Coupled Binding a Folding Equilibria in Ribonuclease P Protein. *J. Am. Chem. Soc.* **2006,** *128* (24), 7772-7781.

64. Wildes, D.; Marqusee, S., Hydrogen exchange and ligand binding: Ligand-dependent and ligand-independent protection in the Src SH3 domain. *Protein Sci.* **2005,** *14* (1), 81-88.

65. Houde, D.; Arndt, J.; Domeier, W.; Berkowitz, S.; Engen, J. R., Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2009,** *81*, 2644-2651.

66. Monroe, E. B.; Kang, S.; Kyere, S. K.; Li, R.; Prevelige Jr, P. E., Hydrogen/Deuterium Exchange Analysis of HIV-1 Capsid Assembly and Maturation. *Structure* **2010,** *18* (11), 1483-1491.

67. Mohan, A.; Oldfield, C. J.; Radivojac, P.; Vacic, V.; Cortese, M. S.; Dunker, A. K.; Uversky, V. N., Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.* **2006,** *362* (5), 1043-1059.

Figures and Tables

Table 3.1. Peptide Assignments.

Peptide	Sequence	# of exchangable amides	RT (min)	Monoisotopic m/z (observed)	z	Monoisotopic Mass (observed)	Monoisotopic Mass (calculated)	Mass error (ppm)
ACTR(1-12)	GTQNRPLLRNSL	9	5.0	456.93	3	1367.77	1367.76	2.21
ACTR(13-22)	DDLVGPPSNL	6	6.5	1026.51	1	1025.50	1025.50	-0.12
ACTR(29-35)	RALLDQL	5	6.2	414.75	2	827.49	827.49	1.17
ACTR(33-42)	DQLHTLLSNT	8	6.6	571.30	2	1140.58	1140.58	0.67
ACTR(33-47)	DQLHTLLSNTDATGL	13	7.2	799.91	2	1597.80	1597.79	1.43
ACTR(36-47)	HTLLSNTDATGL	10	5.9	621.82	2	1241.63	1241.63	1.03
ACTR(48-59)	EEIDRALGIPEL	9	7.2	677.87	2	1353.72	1353.71	1.68
ACTR(60-71)	VNQGQALEPKQD	9	3.3	663.84	2	1325.66	1325.66	2.04
CBP(1-10)	PNRSISPSAL	7	5.1	521.29	2	1040.56	1040.56	3.01
CBP(1-14)	PNRSISPSALQDLL	11	7.1	755.92	2	1509.82	1509.82	3.52
CBP(17-30)	LKSPSSPQQQQQVL	10	4.9	784.43	2	1566.84	1566.84	1.97
CBP(30-38)	LNILKSNPQ	6	5.5	513.80	2	1025.59	1025.59	1.36
CBP(31-40)	NILKSNPQLM	7	5.8	579.32	2	1156.63	1156.63	2.97
CBP(40-43)	MAAF	2	5.9	439.20	1	438.19	438.19	1.17
CBP(43-59)	FIKQRTAKYVANQPGMQ	14	4.1	660.69	3	1979.05	1979.04	3.48

Table 3.2. Peptide deuterium uptake kinetics.

			Intrinsic		Free			Complex							
Protein	Segment	Residues	β	k_{int} (s^{-1})	N	$k_{HX,free}$ (s^{-1})	PF	N_1	N_2	$k_{HX,1}$ (s^{-1})	$k_{HX,2}$ (s^{-1})	PF_1	PF_2	PF_1/PF_{free}	PF_2/PF_{free}
ACTR	1-12	1018-1029	0.60	1.8	8.3	1.8	1.0	8.0		3.0		0.60		0.60	
ACTR	13-22	1030-1039	0.61	1.1	5.0	0.44	2.4	5.1		0.54		1.9		0.80	
ACTR	29-35	1046-1052	0.90	0.61	4.4	0.34	1.8	4.5		0.023		27		15	
ACTR	33-42	1050-1059	0.66	1.9	7.0	1.2	1.6	3.2	3.7	0.43	0.010	4.4	181	2.7	113
ACTR	33-47	1050-1064	0.72	1.7	11	1.3	1.4	7.3	3.9	0.56	0.011	3.1	154	2.3	114
ACTR	36-47	1053-1064	0.73	1.5	7.9	0.93	1.7	5.4	2.6	0.66	0.012	2.3	131	1.4	79
ACTR	48-59	1065-1076	0.77	0.63	7.0	0.31	2.0	1.6	5.5	1.1	0.0046	0.59	136	0.29	67
ACTR	60-71	1077-1088	0.57	1.0	7.4	0.67	1.6	7.7		0.59		1.8		1.1	
CBP	1-10	2059-2068	0.75	2.0	6.5	0.50	4.1	6.4		0.50		4.0		0.99	
CBP	1-14	2059-2072	0.70	1.3	9.3	0.30	4.2	8.9		0.11		12		2.7	
CBP	17-30	2075-2088	0.74	2.1	9.1	0.37	5.7	8.8		0.044		48		8.4	
CBP	30-38	2088-2096	0.64	1.3	5.7	0.24	5.4	5.5		0.022		59		11	
CBP	31-40	2089-2098	0.68	1.3	6.3	0.22	5.7	6.1		0.017		73		13	
CBP	40-43	2098-2101	0.97	1.4	2.1	0.28	5.0	2.0		0.017		84		17	
CBP	43-59	2101-2117	0.75	1.6	12	0.37	4.3	11		0.081		20		4.6	

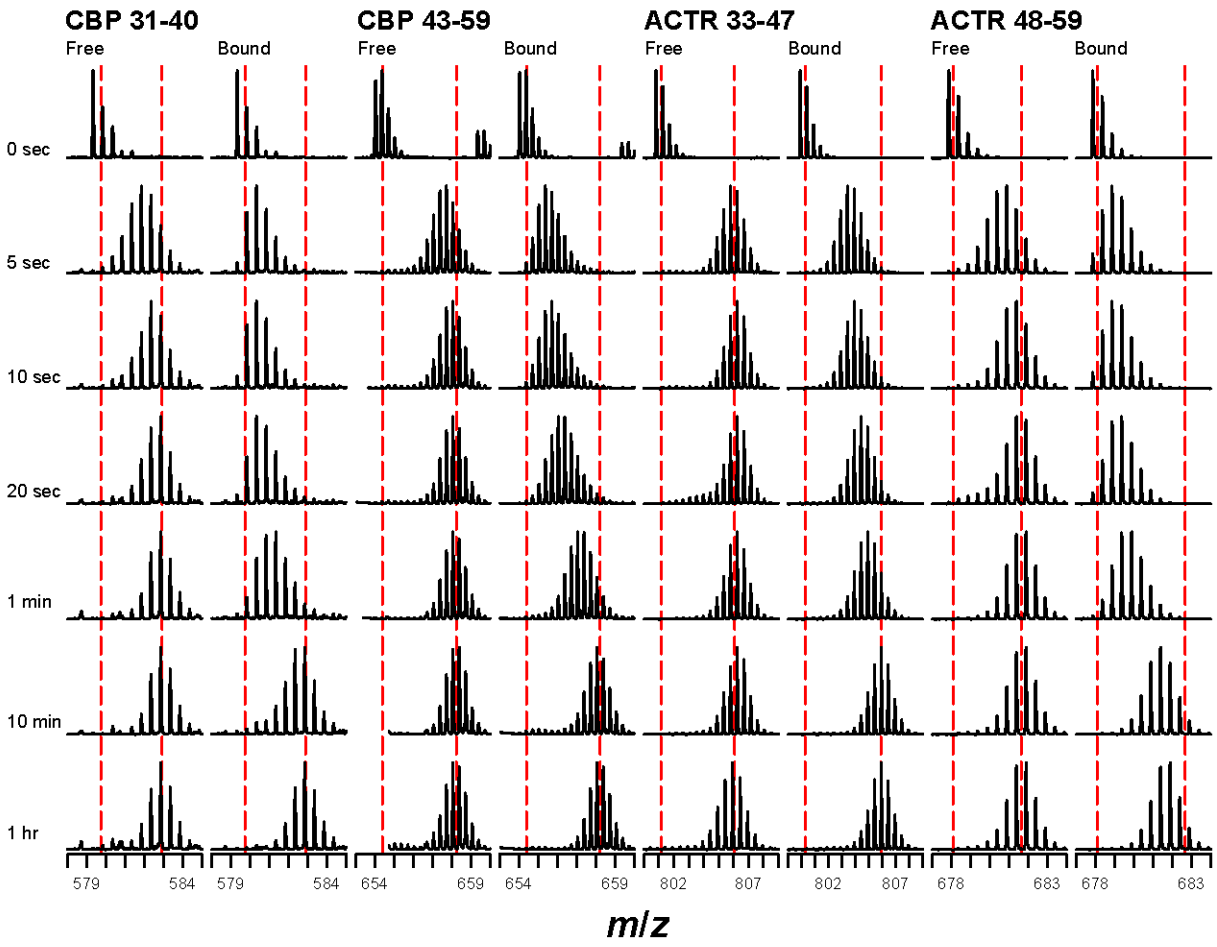


Figure 3.1. Mass spectra showing deuterium uptake in several distinctive regions of CBP and ACTR in both the free and bound states for selected D_2O exposure times. The vertical dashed lines denote the undeuterated (0 s) and fully deuterated (12 hr) limits. The mass spectra show a steady progression of deuterium uptake with no evidence of bimodal isotope profiles.

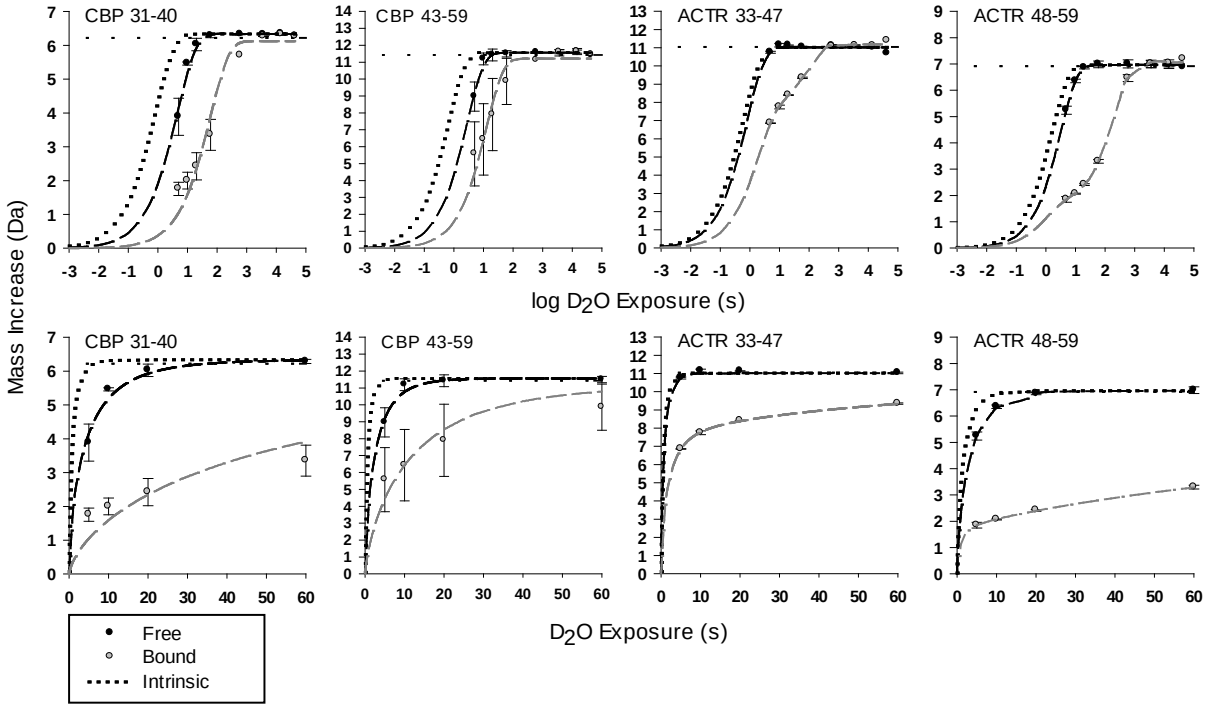


Figure 3.2. Deuterium uptake curves for the distinctive regions of CBP and ACTR shown in figure 3.1. The upper row of plots show the full time-course on a logarithmic scale while the lower row of plots are the early time points on a linear scale. The dotted curves represent the calculated intrinsic rate of exchange calculated using equation 1. Deuterium uptake data are shown with data points, with black points representing uptake by the free proteins and gray points representing uptake by the proteins in the complex. Data points represent the average values obtained from two trials with error bars extended to the specific data points from each trial. Stretched exponential (or bi-exponential for ACTR peptides 33-42, 33-47, 36-47, and 48-59 in the complex state) curves overlay each set of data, with black broken curves and gray broken curves representing the free and protein complex states, respectively. For reference, the maximum deuterium uptake, measured after 24 hr of exchange, is shown as a horizontal dashed line with the exception of ACTR 60-71, which did not produce a detectable spectral intensity. The maximum number of exchangeable amides ($n_{residues} - n_{proline} - 2$) is indicated by the limit of the vertical axis of the plot.

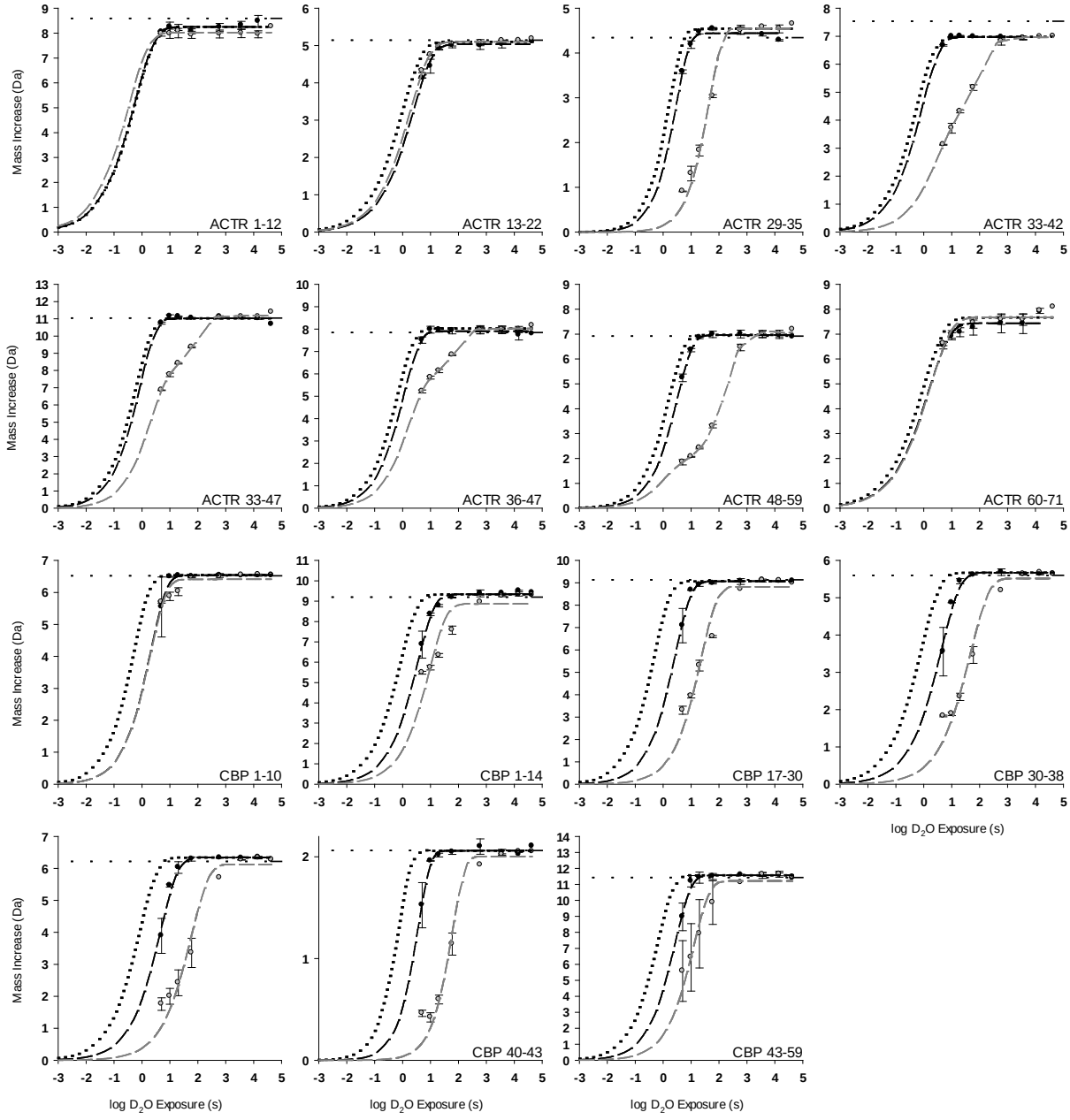


Figure 3.3. Deuterium uptake curves for all measured peptides of CBP and ACTR, plotted as shown in figure 3.2, but only using a logarithmic time scale for D_2O exposure.

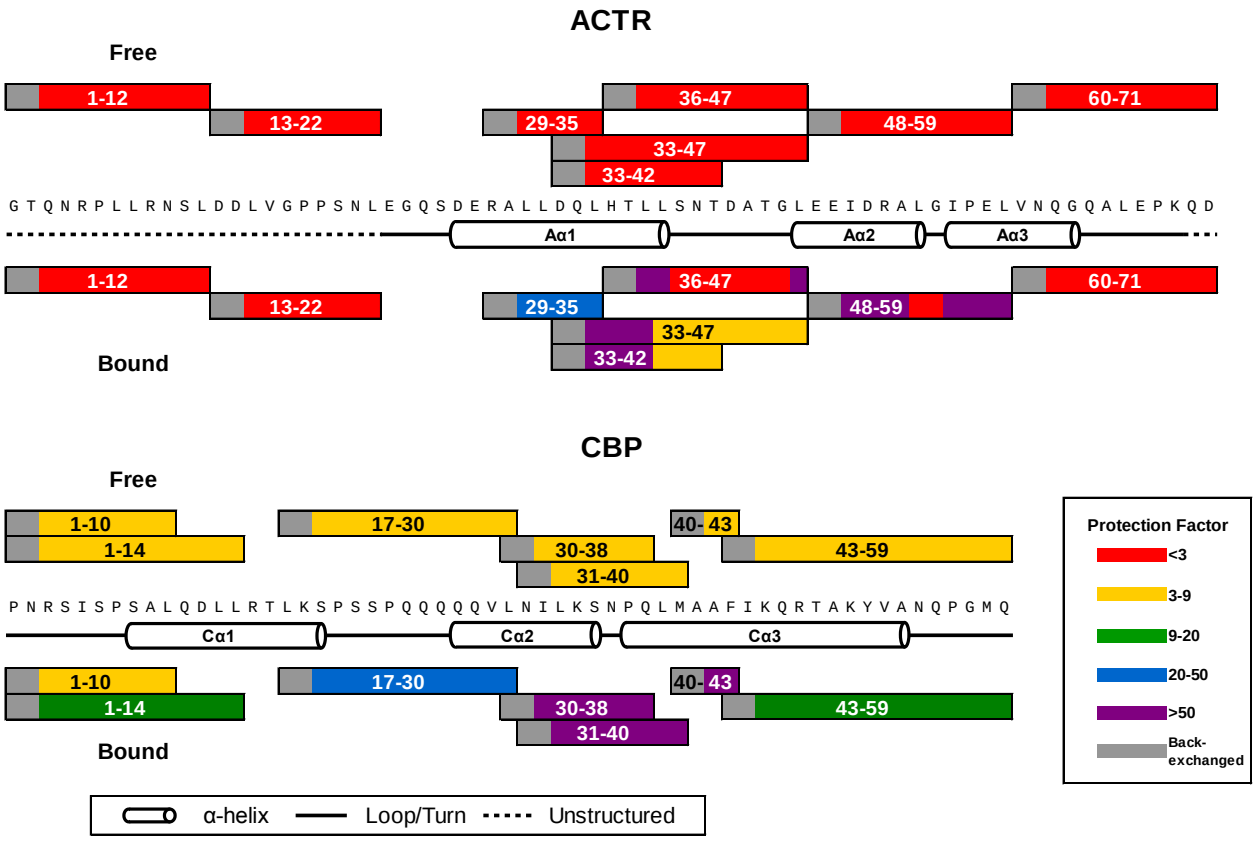


Figure 3.4. Protection factors mapped onto the sequences of ACTR and CBP for each peptide for both free and bound forms. The assigned secondary structural elements from the CBP/ACTR complex[15] are also indicated. The first two residues of each peptide are shown in gray to indicate that rapid back exchange of deuterium label at these two positions leads to no measurable deuterium uptake.

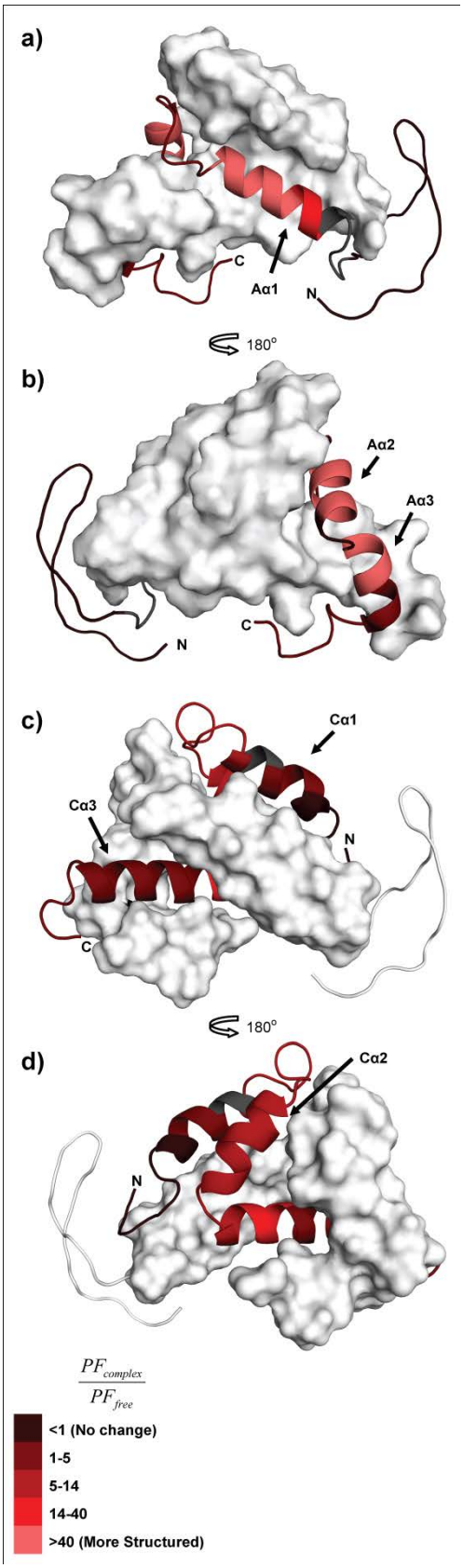

a)
C
Aα1
N
180°
b)
Aα2
Aα3
C
N
c)
Cα1
Cα3
N
C
180°
d)
Cα2
N
$\frac{PF_{complex}}{PF_{free}}$
<1 (No change)
1-5
5-14
14-40
>40 (More Structured)

Figure 3.5. The structure of the ACTR/CBP complex (PDB: 1KBH)[15] showing relative increases in protection of (a, b) ACTR and (c, d) CBP that accompany complex formation. The cognate binding partner is shown in surface representation. In a and b, a modeled unstructured N-terminal region was appended to the NMR structure of ACTR. This figure was rendered using Pymol (version 0.99, DeLano Scientific, South San Francisco, CA).

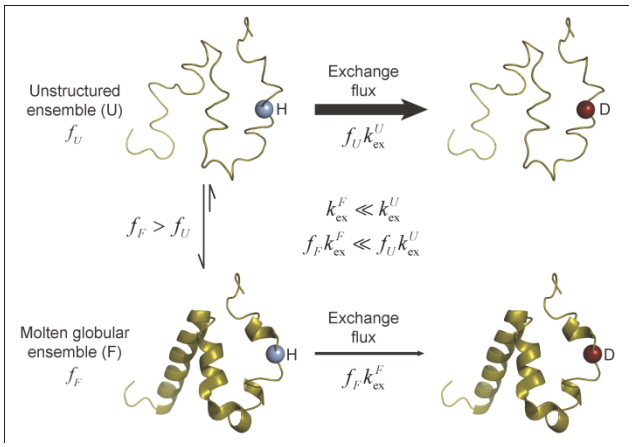


Figure 3.6. This conceptual representation of the H/D exchange flux model illustrates how the low, uniform protection observed for free CBP can be explained by inter-conversion between an ensemble of folded conformers and an ensemble of unstructured conformers. Since the unstructured conformers are much less protected from exchange, most of the total flux of exchange occurs through the unstructured ensemble, as depicted by the heavy arrow. See text for additional details. The protein structures presented in this figure are conceptual models only and do not imply any structural details about the conformations. This figure was rendered with Pymol (version 0.99, DeLano Scientific, South San Francisco, CA) using the structure of CBP from the CBP/ACTR complex (PDB: 1KBH)[15] as a starting point.

Chapter 4: Rapid H/D Exchange Analysis of IDPs Using a Quench-Flow Apparatus

4.1 Introduction

4.1.1 Complete Exchange of IDPs with Short Labeling Time

Deuterium exchange mass spectrometry (H/D-MS) is a versatile method for studying conformational dynamics of intrinsically disordered proteins (IDPs).[1-5] The previous chapter, chapter 3, discusses the H/D-MS analysis of the nuclear co-activator binding domain of the CREB binding protein (CBP) and the activation domain of the activator of thyroid and retinoid receptors (ACTR). In that study, we were able to confirm that CBP is partially folded in its native state, proposing that it exists as a molten globule. We observed that ACTR has little to no protection across its entire sequence in its native state. We also observed marked increases in protection from exchange in both of these proteins as a result of their binding together. However, the deuterium exchange labeling technique was limited by our ability to consistently and reproducibly pipette deuterium label and quench, imposing a lower reaction time limit of about 5 seconds. Because of the near-intrinsic exchange rate by native ACTR and fast exchange by native CBP, peptic peptide resolved analysis revealed that most amide hydrogens exchanged before 5 seconds of reaction.

In describing the data in chapter 3, we used stretched exponentials to fit experimental H/D-MS data. In fitting curves to mostly or completely exchanged peptides, much of the curve shape prior to the 5 second time point was estimated using only the theoretical intrinsic exchange curve shape. This revealed the need for the ability to prepare deuterium exchange samples with on-exchange times on the millisecond time scale in order for us to form a more complete deuterium exchange curve. In this chapter, we will discuss the construction and validation of a simple quench-flow apparatus, and its use in studying the rapid deuterium exchange kinetics of CBP and ACTR. A simple design allows for easy construction of the apparatus. Loop injections from a 6-port valve aid in sample conservation. The quench-flow

apparatus is capable of preparing deuterium exchange samples with reaction times of 42 milliseconds or longer.

4.1.2 Rapid H/D

Many instruments and devices have been designed for the study of sub-second timescale reactions.[6-8] The rapid progression of such reactions, and the possibility for direct connection of these devices into detectors, such as a mass spectrometer, allows for the temporal resolution of reactions such as protein denaturation or folding.[9-12] For a deuterium exchange experiment, one of the simplest designs used is the continuous-flow system, which uses multiple syringes to drive analyte buffer and labeling solution together in a mixer.[13-14] Continuous-flow systems are used to guide eluent directly into a detector for real time analysis. H/D exchange has been carried out using a continuous-flow apparatus with direct MS analysis.[12] Conformational information about the intermediate states in the folding of apomyoglobin upon binding to free heme groups was captured with a short (7 ms) D_2O pulse.[12] An adjustable reaction capillary allowed refolding progress to be monitored after varying times of the folding reaction.[12] The design of the continuous-flow apparatus provided a basis by which the quench-flow system discussed in this chapter would be designed.

Using a quench-flow system provides an alternative approach which allows eluted sample to be collected prior to analysis. A quench-flow apparatus may be designed with a third syringe, which will drive a quench solution flow into a second mixer to quench the reaction online.[15-18] Controlling the amount of time that a protein or peptide is exposed to D_2O buffer in a quench-flow device will depend on capillary volume and flow rate. Quench-flow devices may be designed to have an adjustable capillary length in order to preserve consistent flow rates.[12, 19] Using a variable flow rate and a fixed length of reaction capillary will also produce adjustable reaction times.[16, 20] The quench-flow apparatus used in this chapter has been designed to use an adjustable flow rate, controlled by a syringe pump (see below, section *4.2.2*, for construction details).

While the eluent from our quench-flow apparatus could possibly be directly connected to the HPLC for chromatography,[15] we decided to collect quenched samples into test tubes set in liquid nitrogen. The purpose for this collection prior to analysis is to reduce the variability in the amount of time quenched eluent spends in the solution phase. Other quench-flow designs, such as one used to analyze horse heart cytochrome c,[13] also integrated a proteolytic digestion stage in the quench-flow apparatus. Quench-flow systems may also be designed on microfluidic devices or chips.[17-18] One micro-reactor chip design produced H/D exchange times between 20 and 148 ms to study ribosomal proteins.[17] Another microfluidic quench-flow chip integrated a proteolytic reactor for online digestion the study of cytochrome c, ubiquitin, and DAHP synthase properties.[18] Utilizing an adjustable mixer position, H/D exchange times ranged from 42 ms to 10 s on this quench-flow chip.[18] With reference to our previous work in chapter 3, we demonstrate that comparable datasets can be used to cover a wide range of D_2O labeling times (42 ms to 3500 ms by quench-flow, 5 s to 12 hours by pipetting) and more complete deuterium uptake curves for IDPs can be measured.

4.2 Materials and Methods

4.2.1 Materials

Deuterium oxide (99.9% D) and fluorescein isothiocyanate (FITC) conjugated albumin were obtained from Sigma (St. Louis, MO). 99+% formic acid, used as a mobile phase additive, was obtained from Thermo Scientific (West Palm Beach, FL). Sodium phosphate monobasic, sodium phosphate dibasic, sodium hydroxide, hydrochloric acid, and sodium chloride were obtained from Fisher Scientific (Hanover Park, IL). Optima LC/MS grade acetonitrile and water used in chromatography were also obtained from Fisher Scientific. Individual components of the 9-peptide test mix were each obtained from AnaSpec, Inc. (Fremont, CA). The sequences for each of 9 peptides are listed here and represented in figure 4.3: MQMKKVLDS (anti-inflammatory peptide 1), SFLLRNP (thrombin receptor (42-48) agonist, human), PLARTLSVAGLPGKK (syntide – 2), ASHLGLAR (C3a (70-77)), DRVYIHPFHL (angiotensin 1,

human), YRQSMNNFQGLR (adrenomedullin (1-12), human), EQKLISEEDL (c-Myc peptide epitope), IKNLQSLDPSH (cholecystokinin (10-20)), and RRLIEDNEYTARG (pp60(v-SRC) autophosphorylation site). CBP and ACTR stocks were obtained from previous in-house preparations, as described in chapter 3, with initial concentrations of 13.0 and 14.4 µM, respectively.

4.2.2 Quench-Flow Construction

The quench-flow apparatus was constructed using entirely off-the-shelf parts, requiring no specialized machining or nano-flow components or chips. P-662 luer-to-MicroTight adapters (IDEX, Oak Harbor, WA) allow users to quickly refill and replace luer-lock syringes. Sample loop size, reaction line volume, and sample collection may also be customized. The configuration of the quench-flow apparatus is shown in figure 4.1. Solution flow and mixing was driven by a single Fusion 400 syringe pump (Chemyx, Stafford, TX). Three syringes were selected by relative inner diameter and volume to deliver specific flow rates into the mixers. In our studies, a 1 mL syringe delivered appropriate storage buffer into a 6-port valve (Valco Instruments Co. Inc.). While the valve was in the “load” position, flow was directed through 40 µm inner diameter (ID) fused silica capillary into the first mixer, a PEEK MicroTee (IDEX). The valve was switched to the “inject” position, as shown in figure 4.1, where the syringe flow pushed a plug of sample into the first mixer. The second syringe, a 5 mL syringe, delivered labeling buffer through 150 µm ID fused silica capillary into the first mixer. The resulting reaction mixture eluted from the first mixer through a 4.0 cm length of 150 µm ID fused silica capillary into a second MicroTee. A third syringe, also a 5 mL syringe, delivered the quench buffer through 150 µm ID fused silica capillary into the second mixer. The labeling was quenched and the quenched solution eluted out of the quench-flow apparatus.

To perform a qualitative assessment of mixing efficiency, magnified images of a 22 cm long, 150 µm ID fused silica capillary connected at the outlet of the first mixer were captured. An Axio Vert. A1 microscope (Carl Zeiss Microscopy, Thornwood, NY) was used to monitor the

capillary at about 6 cm downstream of the first mixer. The fused silica capillary shown in the microscope images was stripped of its coating around the length being monitored using a brief exposure to open flame. The microscope images were taken at 2500 ms exposure every 10 seconds after sample injection. A 20X magnification was used to capture the images. A single 10 µL injection of 1.2 mg mL^{-1} FITC conjugated albumin (prepared in 10 mM Tris-HCl) was pushed into the mixer by 10 mM Tris-HCl. Also flowing into the mixer was a mock labeling buffer, which was also 10 mM Tris-HCl. In the 1:20 flow ratio configuration (see section *4.3.1*), the flow rates of sample and mock labeling buffer were 5 µL min^{-1} and 100 µL min^{-1}, respectively. In the 1:5 flow ratio, the flow rates of sample and mock labeling buffer were 20 µL min^{-1} and 100 µL min^{-1}, respectively. Fluorescence of the FITC conjugated albumin is observed upon excitation with ~470 nm light, produced by a DC4100 4-channel LED driver (Thorlabs, Newton, NJ).

4.2.3 Test Peptide Mixture Labeling by H/D Exchange

The rapid H/D exchange method was initially tested using a 9-peptide test mix. The test mix was prepared by diluting the peptides to a concentration of 1 ng μL^{-1} in 10 mM phosphate, 50 mM sodium chloride, pH 6.9. An undeuterated control was prepared by diluting 50 µL of the test mix with 250 µL 10 mM phosphate, 50 mM NaCl, pH 6.9 and 250 µL 200 mM phosphate, pH 2.6 to simulate the addition of labeling buffer and quench. A totally deuterated control was prepared by diluting 50 µL of test mix with 250 µL phosphate buffer in D_2O, pD 6.9, followed by the addition of 250 µL of pH 2.6 quench after 1 hour of reaction. Rapid H/D exchange prepared samples were prepared using 10 µL injections of test mix, which was reacted with the same D_2O labeling buffer as the totally deuterated control and quenched with 200 mM phosphate, pH 2.6, as controlled by the quench-flow apparatus.

4.2.4 Protein Labeling by H/D Exchange

Each protein was prepared separately and in a 5-fold molar excess of its binding partner based on the reported dissociation constant of 34 nM for the CBP/ACTR complex.[21] This molar

ratio will bind over 97.5% of the limited protein in their complex state samples during the D_2O labeling stage. Thus, in this study, four stock solutions were prepared: CBP only (2.3 µM CBP), CBP complex (2.3 µM CBP + 11.8 µM ACTR), ACTR only (2.2 µM ACTR), ACTR complex (2.2 µM ACTR + 11.1 µM CBP). The same labeling and quench buffers were used for CBP/ACTR samples as the 9-peptide test mix. Undeuterated and totally deuterated controls were also prepared similarly to the test mix, with the exception that the totally deuterated samples were allowed to undergo H/D exchange for over 14 hours prior to quench.

Adjustments to the configured syringe pump flow rates allow a wide range of H/D exchange labeling times to be covered. In our protein studies, flow rates ranged from 10.0 to 833.3 µL min^{-1} from the 5 mL syringes, and 2.0 to 166.7 µL min^{-1} from the 1 mL syringe. Higher flow rates would have produced excessive back pressure in the syringes. With a reaction capillary volume of 0.7 µL, the applied flow rates produced rapid H/D labeling times between 42 and 3500 ms. Reaction times ranged from 42 to 5250 ms for the 9-peptide mix samples. The 5250 ms time point was included with the 9-peptide mix samples because there is no previous H/D-MS dataset with a 5 second time point to be used for comparison. 10 µL sample loop injections, enabled by the inclusion of the 6-port valve, ensured low sample consumption per sample collected. Because of the dilution caused by mixing sample with label and quench, a volume of at least 150 µL was collected for the subsequent digestion, chromatography, and MS analysis steps (see below).

Quenched samples were eluted directly into a micro-centrifuge tube set in liquid nitrogen (see figure 4.1). The tube was held a few centimeters below the end of the outlet capillary to prevent blockage by rapidly frozen eluent. The length of time required to collect sufficient volume of eluted sample varied from a few seconds to several minutes in our study. The immediate freezing by liquid nitrogen ensured that each labeled droplet within a single sample run spent the same amount of time in the liquid state following quench (within a few seconds, due to slower flow rates taking longer to produce droplets). An alternative method may include

collecting a full sample volume in a test tube first, capping the tube, and then placing the tube in liquid nitrogen or analyzing the sample immediately. For the longest time points (slowest flow rates), this alternative method would result in the first eluted droplet spending several minutes at room temperature, while the final droplet will have only spent a few seconds at room temperature. This would result in uneven back exchange within the same sample.

4.2.5 Chromatography and MS Analysis

The chromatographic method for the digestion and separation of protein was designed to replicate our previous CBP and ACTR study closely. For the 9-peptide mix samples, the immobilized pepsin column was replaced by a narrow-bore union. An Agilent Technologies 1200 series binary pump was used to carry out a high performance liquid chromatography separation of peptides. Solvents A and B were 0.1% formic acid in water and 90% acetonitrile / 10% water / 0.1% formic acid, respectively. An A to B gradient (5-65% B over 5 minutes for 9-peptide mix, 5-60% B over 5 minutes for digested protein) was applied to separate peptides on a C_{12} column (Jupiter 4 µm, 90 Å Phenomenex). Intact protein samples were initially injected onto a self-packed C_4 trap for desalting, and were separated using a 5-60% B over 2 minutes gradient with no HPLC column in place. A thermoelectrically refrigerated system kept chromatography components at a stable temperature of about 0 ^{o}C (see chapter 2).

The MS analysis method was also designed to closely replicate previous work. Due to improvements in protein preparation, sample dilutions, and instrument limits of detection and sensitivity, more peptic peptides were identified by mass spectrometry for both CBP and ACTR than were previously discovered. The intact CBP and ACTR samples served as both a global rapid H/D exchange test and as an initial mass check for the CBP and ACTR protein stocks, which confirmed the absence of any extensive post-translational modifications. Minor oxidation products were observed, but these did not produce an intense enough signal to interfere with deuterium uptake measurements.

Separate peptide databases were prepared for the 9-peptide mix, intact CBP, intact ACTR, digested CBP, and digested ACTR listing the appropriate retention time, charge state, and sequence information. The HDExaminer software package (version 1.1.0, Sierra Analytics, Inc., Modesto, CA) was used to extract peptide mass spectra, measure average mass, and measure deuterium uptake using raw mass spectral data and the peptide databases.

4.2.6 Kinetic Analysis

An intrinsic exchange curve was calculated for each peptide analyzed in this study to predict the deuterium uptake kinetics for each peptide as if it was a random-coil, as previously described in chapter 3. For CBP and ACTR samples, both intact and peptic digest, the intrinsic exchange curve was normalized by the ratio of average mass measured for totally deuterated control samples and the maximum number of exchangeable amide hydrogens. This accounted for the 83.3% maximum deuteration (1:5 sample buffer to D_2O buffer ratio) and peptide specific back-exchange. For the 9-peptide test mix, the intrinsic exchange curve was normalized to match the measured uptake at the last measured time point (see figure 4.3), as totally deuterated control samples indicated deuterium uptake beyond theoretical limits (data not shown).

The kinetic analysis was replicated from the previous work in chapter 3, which includes a stretched exponential curve-fitting to measure deuterium uptake protection factors, $\langle PF \rangle$. The curve fits were calculated using data only from 42 ms to 3500 ms of rapid H/D labeling and the totally deuterated controls. Though shown for illustrative purposes in figure 4.6, the manually labeled H/D dataset encompassing labeling times from 5 seconds to 12 hours, originally presented in chapter 3, was not included in curve-fitting calculations. Exceptions to the curve-fit method were made for peptides when the mass increase did not reach half of the number of exchangeable amides within 3500 ms of exchange. The measured protection factors for ACTR

and CBP peptic peptides, both free and bound in complex, are mapped in figure 4.7 and listed in table 4.1.

To correlate the dataset from this experiment with results presented in chapter 3, we normalized previously acquired deuterium uptake curves, $\Delta m_{\text{manual}}(t)$, to fit newly acquired curves. We adjusted the mass increase for each measurement by the ratio between the quench-flow experiment's totally deuterated control, $\Delta m^{\text{TD}}_{\text{QF}}$, and the previous experiment's totally deuterated control, $\Delta m^{\text{TD}}_{\text{manual}}$, as shown in equation 1:

$$\Delta m_{\text{manual}}(t)_{\text{adj.}} = \Delta m_{\text{manual}}(t) \times \frac{\Delta m^{\text{TD}}_{\text{QF}}}{\Delta m^{\text{TD}}_{\text{manual}}} \quad (1)$$

An exception was made for ACTR 1077-1088, which did not have a measurable totally deuterated MS feature in the previous dataset. Instead, the uptake curve was normalized by the theoretical D_{max} difference between the two datasets (*i.e.* 83.3%:95%). The normalized, previous datasets, $\Delta m_{\text{manual}}(t)_{\text{adj.}}$, are shown by black points (free protein) points and white points (bound protein) in figure 4.6. Intrinsic exchange curves (dotted lines on figure 4.6) were also normalized to totally deuterated samples from the rapid H/D dataset.

4.3 Results and Discussion

4.3.1 Optimizing Mixer Design under Laminar Flow Conditions

In designing a quench-flow mixing apparatus, several challenges needed to be overcome. One challenge was accurately measuring the amount of time analyte spent undergoing a reaction in a short capillary. Merging two streams of buffer together in a mixer to produce a uniform mixture at the outlet was another challenge. The effects of laminar and plug flow were also considered, as these fluid behaviors affect the age distribution of reacting analyte.[22] We adjusted the quench-flow apparatus design to address problems that arose with sample consumption, mixing efficiency, and sample collection and storage.

The homogenous velocity of a plug of solution traveling via turbulent flow makes the age of that plug easier to predict.[22] Under ideal turbulent flow conditions, the reaction time, t_{rxn}, for analyte undergoing deuterium exchange in the reaction line following the mixing in of D_2O is a function of the geometric reaction capillary volume, V_{geo} (determined by reaction capillary length, L, and reaction capillary radius, r_d), and total flow rate, F_1+F_2, as shown in equation 2:

$$t_{rxn} = \frac{V_{geo}}{F_1 + F_2} = \frac{\pi r_d^{\,2} L}{F_1 + F_2} \tag{2}$$

Under laminar flow, the velocity of solution will be distributed across the capillary, depending on its radial position.[19] The fastest stream will lie in the center of the capillary and the slowest will lie along the capillary walls. Turbulent flow is desirable because it will maintain constant mixing between slow and fast streams of a fluid moving through a capillary.[22] Whether a fluid is moving via turbulent flow or laminar flow through a capillary can be predicted using the Reynold's number. The Reynold's number, Re, is calculated as in equation 3, where v represents linear velocity of the flow, d represents capillary diameter, ρ represents the fluid density, and μ represents the absolute viscosity of the fluid:

$$\mathrm{Re} = v\frac{\rho d}{\mu} \tag{3}$$

At the highest flow rate used in this rapid H/D study, v = 0.94 m s^{-1} through the reaction capillary of d = 1.5 × 10^{-4} m. Assuming pure deuterium oxide and using ρ = 1105 kg m^{-3} and μ = 1.25 × 10^{-3} N s m^{-2} at 20 °C,[23] Re was about 125. Since the Reynold's number for all flow paths at the proposed flow rates was estimated to be below 2000, the approximate value necessary for turbulent flow,[24] it would seem the possible effects of laminar flow on H/D exchange reaction time should be considered.

The reaction age of molecules flowing through a capillary will be highly distributed due to laminar flow effects, but these effects will be partially corrected when molecular diffusion is

considered in a narrow capillary.[19, 22] In simulations, molecules without molecular diffusion produced a wide, trailing distribution of reaction age in a capillary. For larger molecules, such as protein, diffusion coefficients will be relatively small. When higher diffusion coefficients and longer capillaries are applied, the highest frequency age (the peak of the distribution) moves closer to what it would be under turbulent plug flow.[19, 22] In other words, radial diffusion counters the laminar flow effect of parabolic radial velocity distribution. However, even in laminar flow reaction kinetics simulations,[22] a laminar flow simulation without molecular diffusion still provides a fair first-approximation of fractional reaction progression for rate constants expected from intrinsic exchange calculations. Even a small diffusion coefficient expected for proteins will improve the disparity between homogenous flow and laminar flow in reaction kinetics calculations. In this chapter, we have calculated D_2O exposure times as though flow in the reaction capillary was plug flow.

Volumetric flow rates, *F*, are determined by the syringe pump with a single syringe pump piston producing flow rates based on syringe diameter. The respective radii of each capillary, *r*, will determine the linear velocity of the fluid, v, which is proportional to the square of these radii as given by equation 4:

$$\frac{v_2}{v_1} = \frac{F_2}{F_1}\left(\frac{r_1}{r_2}\right)^2 \tag{4}$$

Therefore, if 150 µm ID (r = 75 µm) fused silica capillary was used to deliver both sample and labeling buffer into the first mixer, the linear velocities of each stream would be proportional to the flow rates. Using 150 µm ID for all capillary lines and a 1:20 sample to labeling buffer flow rate ratio would have produced a ratio of 1:20 in their linear velocities. This 1:20 sample to labeling buffer flow rate ratio would have been desirable, as it would have allowed up to ~95% deuteration maxima before back exchange.

Simple tests using commercial food dye and water revealed problems with mixing two solution streams together at a flow rate ratio of 1:20. When the outlet capillary of the first mixer was viewed under a microscope, both the food coloring stream and the water stream could be observed, still separate. This observation revealed that the two streams were not mixing properly. The dye stream appeared to be "pinned" to the edge of the capillary by the overwhelming water stream. Initial trials may have been unsatisfactory due to a mismatch in the flow streams' linear velocities.

The sample buffer capillary was replaced with 40 µm ID (r = 20 µm) capillary, to better match the linear velocity of the sample stream and the label stream. When using a 500 µL sample buffer syringe and a 10.0 mL labeling buffer syringe, this new capillary configuration produced a linear velocity ratio of about 14:20 between the sample and labeling streams, instead of the full 1:20 ratio as it would be with equal diameter capillaries. Unfortunately, the two streams still failed to mix properly. Figure 4.2A shows a sequence of magnified fluorescence images following injection of FITC conjugated albumin and mock labeling buffer into the first mixer at a flow rate ratio of 1:20 and a linear velocity ratio of 14:20. These images show the coating-stripped, transparent portion of the capillary at a fixed 6 cm distance downstream of the first mixer. Food dye molecules diffuse much faster than more bulky protein molecules, so food dye may not provide an accurate test model. Therefore, the use of albumin, a 66 kDa protein, as an analyte better simulated mixing under the conditions of an H/D exchange labeling experiment. As can be seen over the time course of figure 4.2A, the two flow streams remained separate following the first mixer. The separate streams indicate incomplete mixing. These images also reveal that the sample stream was delayed, taking the fluorescence over 300 seconds to fully disappear after its first appearance. The expected clearance time of analyte, t_{clear}, may be calculated based on loop volume, V_{loop}, and sample stream flow rate, F_1, as in equation 5:

$$t_{\text{clear}} = \frac{V_{\text{loop}}}{F_1} \tag{5}$$

In the images captured in figure 4.2A, a 5 µL min^{-1} sample buffer flow rate pushed a 10 µL loop injection and mixed with a 100 µL min^{-1} labeling buffer flow rate out in the first mixer. Under ideal plug flow, fluorescence disappearance should reflect the clearance time and only require 120 seconds under these conditions. The longer disappearance time could indicate that sample dwells within the reaction capillary for a long time. This could mean that D_2O exposure times are skewed to be much longer than what is expected. This problem may be caused by the overwhelming 20-fold greater flow rate of labeling buffer, which pinned the smaller sample flow against the walls of the capillary under laminar flow.

To eliminate these problems, a 1:5 mixing ratio was chosen, instead. This syringe configuration actually made the linear velocity of the sample stream faster than that of the labeling buffer, with about a 2.8:1 ratio. When we repeated the food dye and fluorescent protein conjugate tests in this configuration, only a single stream of fluorescence or dye was observed in the outlet line from the first mixer. Figure 4.2B shows time-resolved sequence of capillary images for the 1:5 configuration following the injection of FITC-albumin. The fluorescence mostly disappeared about 30 seconds after its first appearance with some residual fluorescence persisting longer, which was close to the disappearance time of 30 seconds expected for ideal plug flow (*i.e.* 10 µL injection is pushed by 20 µL min^{-1} sample buffer flow). The observations of a single stream and a narrow clearance time indicate a better mixing efficiency and plug-like flow in the 1:5 flow ratio configuration.

4.3.2 Validation of Quench-Flow Labeling

We conducted a practical test to ensure deuterium exchange initiated by using the quench-flow apparatus produced measurable deuterium uptake at reaction times approximated by equation 2. To perform this test, a mixture of 9 different peptides was deuterated then

analyzed by HPLC and mass spectrometry. The deuterium uptake curves measured from these peptides, which have been labeled with D_2O exposure times ranging from 42 ms to 5250 ms as calculated according to equation 2, are presented in figure 4.3. The calculated intrinsic exchange curves (see chapter 3), $\Delta m_{int}(t)$, have been normalized to the ratio between the measured mass increase, Δm_{QF}, of the latest measured time point for each peptide, t_{max} and the predicted mass increase for intrinsic exchange, Δm_{int}, at the same t_{max}.

$$\Delta m_{\text{int}}(t)_{\text{norm.}} = \Delta m_{\text{int}}(t) \times \frac{\Delta m_{\text{QF}}(t_{\max})}{\Delta m_{\text{int}}(t_{\max})} \tag{6}$$

Intrinsic exchange was normalized this way since some totally deuterated controls were either not detected or showed C-terminal amide deuteration that was slower than the measured timescale. The 5250 ms time point for angiotensin 1 was also not detected. We were able observe early exchange points with relatively small mass increase values and the progression of exchange as D_2O exposure time lengthened. The closeness between calculated intrinsic exchange points and experimentally measured exchange points indicates that the calculated reaction times are accurate. This agreement applies to both faster exchanging peptides, such as YRQSMNNFQGLR, and slower exchanging peptides, such as EQKLISEEDL.

4.3.3. Rapid H/D Exchange by Intact CBP/ACTR

We also conducted an initial test using the quench-flow apparatus to rapidly label the ACTR and CBP proteins and analyze these samples without a peptic digest. The deuterium uptake curves from this experiment are shown in figure 4.4. The intrinsic exchange curves, $\Delta m_{int}(t)$, in this figure were normalized by the ratio between the measured mass increase of the totally deuterated control for the quench-flow experiment, Δm_{QF}^{TD}, and the totally deuterated mass increase for calculated intrinsic exchange, Δm_{int}^{TD}, as shown in equation 7:

$$\Delta m_{\text{int}}(t)_{\text{norm.}} = \Delta m_{\text{int}}(t) \times \frac{\Delta m_{\text{QF}}^{\text{TD}}}{\Delta m_{\text{int}}^{\text{TD}}} \tag{7}$$

The exchange for free ACTR and CBP is slower than intrinsic exchange, which is indicative of the residual structure discussed in chapter 3 and in previous work.[21] We also observe that protection from exchange increases when either protein is in the presence of a 5-fold molar excess of its binding partner. This effect validates the binding of ACTR to CBP and vice versa, the effects of which we will be able to analyze using rapid H/D-MS with peptic digestion.

4.3.4 Quench-Flow Precision

The quench-flow apparatus demonstrates remarkable precision in the amount of deuteration observed in H/D-MS measurements. For the peptide-level analyses of CBP and ACTR outlined in the following sections, 653 standard deviation values for measured mass increase were calculated. Most of these standard deviations were calculated for triplicate H/D-MS runs. 666 averaged data points were expected (37 peptides × 9 time points × 2 protein states). However, standard deviations were not calculated for 13 points, since a feature was either detected in only one or none of the triplicate trials. Standard deviations from runs with peptide detection in 2 of the 3 trial runs were included. Figure 4.5 is a histogram that shows the frequency of specific standard deviation values. 534 of these 653 standard deviations fell within ±0.1 Da. The mean of the 653 standard deviation values was ±0.07 Da. The median of these 653 values is ±0.04 Da. The precision of the quench-flow system in H/D mass increase measurements should be considered excellent. Replicate sampling with three or more trials was conducted in an H/D-MS biopharmaceutical comparability analysis of interferon-β-1a.[25] A standard deviation of ±0.14 Da was reported for these measurements.

4.3.5 Rapid H/D Exchange by Free ACTR

Our previous H/D-MS analysis of free ACTR revealed rapid exchange by all of the analyzed segments. The goal of this rapid H/D-MS study was to extend the labeling timescale beyond the five second minimum to the millisecond range in order to reveal any residual structural elements in the protein. The deuterium uptake curves shown in figure 4.6 reveal that exchange occurs at near intrinsic rates for unbound ACTR segments. Calculated stretched

exponential fit parameters and protection factors are shown in table 4.1 and figure 4.7. For the C- and N-terminal located regions, particularly ACTR(1018-1029) and ACTR(1077-1088), the stretched exponential function overlaps with the intrinsic exchange function very closely, giving measured protection factors of 1.5 or less. Some segments located within the sequence, such as ACTR(1040-1049), (1052-1055), (1060-1064), (1065-1076), and (1067-1076), showed slight protection from exchange with measured protection factors greater than 2.0. Deuterium uptake curves for these mid-sequence segments were shifted up the time axis, representing slightly slower exchange than terminally located segments. In the NMR study by Ebert, *et al*., slightly positive Cα and CO secondary chemical shifts were observed around ACTR 1031.[26] A similar study also observed positive ^{13}C methyl group secondary chemical shifts around ACTR 1031.[27] A slight increase in protection was also observed in rapid H/D measurements for ACTR segments (1030-1039), (1031-1039), and (1032-1039). While this region was determined to be unstructured in earlier NMR studies,[21] both results presented by Ebert *et al*. and these rapid H/D results imply a slight reduction in dynamics in this region. The region of the Aα1 helix also showed a reduction in dynamics in NMR studies of free ACTR.[26-27] This result is supported by our rapid H/D-MS data, as segments located within this region, particularly ACTR(1040-1049) and (1052-1055), exhibited larger protection factors.

4.3.6 Rapid H/D Exchange by ACTR in Complex with CBP

When ACTR was bound to CBP, we observed a marked increase in protection from exchange in regions measured to be α-helical by NMR. As shown in figure 4.7, several segments within each α-helix showed a deuterium uptake too slow to characterize by curve-fitting. Exchange that was too slow was defined as having a mass increase of less than half of the totally deuterated controls within the 3500 millisecond exposure limit. Segments that best exhibited this slowing of exchange include ACTR(1046-1049), ACTR(1046-1052), and ACTR(1052-1055), which are all located within Aα1. Likewise, the ACTR(1065-1076) and ACTR(1067-1076) segments are both located across the Aα2 region and into the Aα3 region

and also exhibited protection from exchange. The combined information for the unbound and bound ACTR supports the analysis from previous studies.[21, 26] Free ACTR is thought to exist as a near random-coil protein in its native state across most regions, but gains a more stable conformation upon binding with CBP, specifically in the regions of α-helices. The regions near the N-terminus, up to about residue 1040, remain highly dynamic, even when ACTR is bound in complex. This supports the findings of the NMR study, which showed similar chemical shift profiles for this region in both free and bound ACTR.[26]

4.3.7 Rapid H/D Exchange by Free CBP

Free CBP showed rapid exchange for most regions analyzed, with a few exceptions. Like ACTR, the C- and N-terminal regions from CBP showed near intrinsic deuterium exchange rates, with protection factors of 1.5 or less for CBP(2059-2068), CBP(2059-2071), CBP(2101-2117), and CBP(2110-2117) (see figure 4.7). Protection factors for CBP segments (2068-2071), (2089-2098), (2091-2096), (2092-2100), (2098-2101), and (2099-2101) were measured to be larger than 3.0 for free CBP. While the magnitudes of protection factors measured for these segments are small relative to studies of stable, structured proteins, they exhibit the usefulness of the application of rapid H/D to IDPs. Previously, as shown in chapter 3, all analyzed segments from free CBP showed fairly uniform protection, with $\langle PF \rangle$ ranging from 4.1 to 5.7. More of the deuterium uptake curve was covered with the rapid H/D exchange times (ranging from 42 ms to 3500 ms). This improved the accuracy of this chapter's stretched exponential fits over those from chapter 3. While the $\langle PF \rangle$ values were smaller in the quench-flow data, there was a greater contrast between individual segments. Deuterium exchange in the millisecond time scale revealed subtle differences in conformations of mostly unstructured proteins. Our data supports the model of CBP being a loosely structured protein that undergoes transient unfolding in its native, free state.[21, 26, 28] Protection was observed to be greater in

regions predicted to have transient conformation, shown as the α-helical regions in figure 4.7. These results suggest that not all conformers in the ensemble are equally unfolded.

4.3.8 Rapid H/D Exchange by CBP in Complex with ACTR

All analyzed regions show that CBP gains protection as a result of binding to ACTR. As shown in figure 4.6, many non-terminal located regions of bound CBP exhibited slowed deuterium uptake. CBP segments (2071-2089), (2088-2096), (2089-2098), (2091-2096), (2092-2100), and (2100-2109) did not show at least half the mass increase of totally deuterated controls within 3500 ms of labeling. In the short segments, CBP (2068-2071), (2098-2101), and (2099-2101), no measurable increase in deuterium uptake was observed within 3500 milliseconds of exchange, as listed on table 4.1. A stretched exponential fit would be inappropriate to describe segments that exchanged too slowly or not at all. C- and N-terminal regions of CBP also exhibited an increase in protection upon forming the CBP/ACTR complex. The overall gain in exchange protection supports the model for CBP gaining conformational stability as a result of binding, becoming a globular protein.[26, 28] This gain in conformational stability represents a very different protein dynamic change in CBP than it does in ACTR. Because native CBP is predicted to be a loosely structured molten globule, the protein ensemble contains many conformers that are stable and share similar secondary structure characteristics. NMR results support the prediction that the ensemble is undergoing rapid inter-conversion between conformers, which suggests a conversion between tertiary structures.[28] This differs from the transitions observed in the NMR study of ACTR, which showed that ACTR gains both secondary structure and tertiary structure as a result of binding to CBP.[26] When CBP interacts with other binding partners, a different secondary structure configuration might be observed. An example of another CBP conformation was observed when the CBP domain interacts with IRF-3 (PDB: 1ZOQ).[29] The secondary structure map of CBP in complex with IRF-3 differs from CBP bound to ACTR primarily in that the first helix (2066-2072) and third helix (2094-2103) are shorter. When bound to ACTR, CBP residues 2070-2071 and 2100-2101 were

found to have the greatest protection from exchange. Interestingly, both of these sets of residues are located in α-helical regions of CBP in the CBP/ACTR complex and CBP/IRF-3 complex models.

Rapid H/D-MS results captured the dynamic inter-conversion of CBP conformers in the free state. These conversions are likely to expose amide hydrogens to exchange, as hydrogen bonding is disrupted during the transitions. The binding of CBP to ACTR reduces the probability of these inter-conversion events, favoring the structure bound conformers. This reduction in inter-conversion events was reflected by the gain in exchange protection in rapid H/D-MS measurements.

4.4 Conclusions

In this work, we have described the construction and validation of a quench-flow apparatus and its application in the rapid H/D-MS analysis of the IDPs, ACTR and CBP. The extension of deuterium exchange labeling times to a lower limit of 42 milliseconds allows for more accurate deuterium uptake curves to be measured for rapidly exchanging segments. While the technology of constructing a quench-flow system was already established, our apparatus was constructed from low-cost, available materials and required no specialized machining. An advantage of using our quench-flow apparatus is the adjustable sample loop size, which helps conserve sample. A smaller loop could also be implemented in order to reduce the clearance time of sample, allowing more rapid analyses. The reaction line volume is also variable, which can allow alteration in the range of labeling times available. Modifications of the quench-flow design could allow the system to connect directly to an LC or ESI-MS for online sample labeling and analysis. The elution line may be directed through a pepsin column, and peptic peptides can then be loaded directly onto a refrigerated trap.

Our rapid H/D-MS analysis of ACTR and CBP represents an initial application of the quench-flow apparatus to study conformational dynamics of intrinsically disordered proteins and binding induced folding. Combining the rapid H/D technique with a traditional manually labeled

H/D technique allows an extended labeling time regime to be measured. Lower limits in the tens of milliseconds range to upper limits of days or more can be measured by the experimenter. Our study further supports the native form models of ACTR as a mostly random-coil protein and CBP as a molten globule protein. Even in the D_2O exposure time range of 42 to 3500 milliseconds, we can monitor the increase in conformational stability that results from ACTR binding to CBP and vice versa.

References

1. Zhang, Z.; Smith, D. L., Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* **1993,** *2* (4), 522-531.
2. Croy, C. H.; Bergqvist, S.; Huxford, T.; Ghosh, G.; Komives, E. A., Biophysical characterization of the free IκBα ankyrin repeat domain in solution. *Protein Sci.* **2004,** *13* (7), 1767-1777.
3. Mitchell, J. L.; Trible, R. P.; Emert-Sedlak, L. A.; Weis, D. D.; Lerner, E. C.; Applen, J. J.; Sefton, B. M.; Smithgall, T. E.; Engen, J. R., Functional characterization and conformational analysis of the *Herpesvirus saimiri* Tip-C484 protein. *J. Mol. Biol.* **2007,** *366* (4), 1282-1293.
4. Pan, J.; Han, J.; Borchers, C. H.; Konermann, L., Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* **2009,** *131* (35), 12801-12808.
5. Hansen, J. C.; Wexler, B. B.; Rogers, D. J.; Hite, K. C.; Panchenko, T.; Ajith, S.; Black, B. E., DNA Binding Restricts the Intrinsic Conformational Flexibility of Methyl CpG Binding Protein 2 (MeCP2). *J. Biol. Chem.* **2011,** *286* (21), 18938-18948.
6. Dickson, P. N.; Margerum, D. W., Extension of accessible first-order rate constants and accurate dead-time determinations for stopped-flow spectroscopy. *Anal. Chem.* **1986,** *58* (14), 3153-3158.
7. Shastry, M. C. R.; Luck, S. D.; Roder, H., A Continuous-Flow Capillary Mixing Method to Monitor Reactions on the Microsecond Time Scale. *Biophys. J.* **1998,** *74* (5), 2714-2721.
8. Kolakowski, B. M.; Simmons, D. A.; Konermann, L., Stopped-flow electrospray ionization mass spectrometry: a new method for studying chemical reaction kinetics in solution. *Rapid Commun. Mass Spectrom.* **2000,** *14* (9), 772-776.
9. Konermann, L.; Rosell, F. I.; Mauk, A. G.; Douglas, D. J., Acid-Induced Denaturation of Myoglobin Studied by Time-Resolved Electrospray Ionization Mass Spectrometry *Biochemistry* **1997,** *36* (21), 6448-6454.
10. Yang, H.; Smith, D. L., Kinetics of cytochrome *c* folding examined by hydrogen exchange and mass spectrometry. *Biochemistry* **1997,** *36* (48), 14992-14999.
11. Konermann, L.; Collings, B. A.; Douglas, D. J., Cytochrome c Folding Kinetics Studied by Time-Resolved Electrospray Ionization Mass Spectrometry *Biochemistry* **1997,** *36* (18), 5554-5559.
12. Simmons, D. A.; Konermann, L., Characterization of Transient Protein Folding Intermediates during Myoglobin Reconstitution by Time-Resolved Electrospray Mass Spectrometry with On-Line Isotopic Pulse Labeling. *Biochemistry* **2002,** *41* (6), 1906-1914.
13. Dharmasiri, K.; Smith, D. L., Mass Spectrometric Determination of Isotopic Exchange Rates of Amide Hydrogens Located on the Surfaces of Proteins. *Anal. Chem.* **1996,** *68* (14), 2340-2344.
14. Truhlar, S. M. E.; Croy, C. H.; Torpey, J. W.; Koeppe, J. R.; Komives, E. A., Solvent Accessibility of Protein Surfaces by Amide H/2H Exchange MALDI-TOF Mass Spectrometry. *J. Am. Soc. Mass. Spectrom.* **2006,** *17* (11), 1490-1497.
15. Rist, W.; Rodriguez, F.; Jorgensen, T. J. D.; Mayer, M. P., Analysis of subsecond protein dynamics by amide hydrogen exchange and mass spectrometry using a quenched-flow setup. *Protein Sci* **2005,** *14* (3), 626-632.

16. Clarke, D. J.; Stokes, A. A.; Langridge-Smith, P.; Mackay, C. L., Online Quench-Flow Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Elucidating Kinetic and Chemical Enzymatic Reaction Mechanisms. *Anal. Chem.* **2010,** *82* (5), 1897-1904.
17. Yamamoto, T.; Shimizu, Y.; Ueda, T.; Shiro, Y.; Suematsu, M., Application of micro-reactor chip technique for millisecond quenching of deuterium incorporation into 70S ribosomal protein complex. *Int. J. Mass spectrom.* **2011,** *302* (1-3), 132-138.
18. Rob, T.; Liuni, P.; Gill, P. K.; Zhu, S.; Balachandran, N.; Berti, P. J.; Wilson, D. J., Measuring Dynamics in Weakly Structured Regions of Proteins Using Microfluidics-Enabled Subsecond H/D Exchange Mass Spectrometry. *Anal. Chem.* **2012,** *84* (8), 3771-3779.
19. Wilson, D. J.; Konermann, L., A Capillary Mixer with Adjustable Reaction Chamber Volume for Millisecond Time-Resolved Studies by Electrospray Mass Spectrometry. *Anal. Chem.* **2003,** *75* (23), 6408 -6414.
20. Attwood, P. V.; Geeves, M. A., Kinetics of an enzyme-catalyzed reaction measured by electrospray ionization mass spectrometry using a simple rapid mixing attachment. *Anal. Biochem.* **2004,** *334* (2), 382-389.
21. Demarest, S. J.; Martinez-Yamout, M.; Chung, J.; Chen, H.; Xu, W.; Dyson, H. J.; Evans, R. M.; Wright, P. E., Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **2002,** *415* (6871), 549-553.
22. Konermann, L., Monitoring Reaction Kinetics in Solution by Continuous-Flow Methods: The Effects of Convection and Molecular Diffusion under Laminar Flow Conditions. *The Journal of Physical Chemistry A* **1999,** *103* (36), 7210-7216.
23. Haynes, W. M., *CRC Handbook of Chemistry and Physics*. 93rd Edition, Internet Version 2013 ed.; Taylor Francis Group: Boulder, Colorado, **2013**.
24. Johnson, K. A., [2] Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. In *Methods Enzymol.* Academic Press: **1995**; Vol. Volume 249, pp 38-61.
25. Houde, D.; Berkowitz, S. A.; Engen, J. R., The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J. Pharm. Sci.* **2011,** *100* (6), 2071-2086.
26. Ebert, M.-O.; Bae, S.-H.; Dyson, H. J.; Wright, P. E., NMR Relaxation Study of the Complex Formed Between CBP and the Activation Domain of the Nuclear Hormone Receptor Coactivator ACTR *Biochemistry* **2008,** *47* (5), 1299-1308.
27. Kjaergaard, M.; Iešmantavičius, V.; Poulsen, F. M., The interplay between transient α-helix formation and side chain rotamer distributions in disordered proteins probed by methyl chemical shifts. *Protein Sci.* **2011,** *20* (12), 2023-2034.
28. Kjaergaard, M.; Teilum, K.; Poulsen, F. M., Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (28), 12535-12540.
29. Hiscott, J.; Lin, R., IRF-3 Releases Its Inhibitions. *Structure* **2005,** *13* (9), 1235-1236.

Figures

Table 4.1. Peptide deuterium uptake kinetics.

Peptide	Sequence	D_{max} [a]	N (Da)[b]	β_{int}	k_{int} (ms^{-1})	β_{free}	$k_{HX,free}$ (ms^{-1})	$<PF>_{free}$	β_{bound}	$k_{HX,bound}$ (ms^{-1})	$<PF>_{bound}$
ACTR(1018-1025)	GTQNRPLL	4.17	2.99	0.52	1.48E-03	0.52	2.43E-03	0.61	0.52	3.27E-03	0.45
ACTR(1018-1029)	GTQNRPLLRNSL	7.50	5.73	0.56	2.15E-03	0.56	2.41E-03	0.89	0.56	2.83E-03	0.76
ACTR(1020-1029)	QNRPLLRNSL	5.83	5.17	0.56	1.57E-03	0.56	2.09E-03	0.75	0.56	2.54E-03	0.62
ACTR(1030-1039)	DDLVGPPSNL	5.00	3.90	0.60	1.22E-03	0.60	7.28E-04	1.68	0.60	1.09E-03	1.12
ACTR(1031-1039)	DLVGPPSNL	4.17	3.21	0.61	1.70E-03	0.61	1.04E-03	1.64	0.61	1.55E-03	1.09
ACTR(1032-1039)	LVGPPSNL	3.33	2.92	0.76	3.05E-03	0.76	1.68E-03	1.81	0.76	2.61E-03	1.17
ACTR(1033-1045)	VGPPSNLEGQSDE	7.50	5.69	0.67	2.00E-03	0.67	1.73E-03	1.16	0.67	2.35E-03	0.85
ACTR(1040-1049)	EGQSDERALL	6.67	4.35	0.76	1.91E-03	0.76	5.39E-04	3.55	0.76	3.24E-04	5.91
ACTR(1046-1049)	RALL	1.67	1.53	0.98	4.48E-04	0.86	2.36E-04	1.90 [c]		too slow	
ACTR(1046-1052)	RALLDQL	4.17	3.72	0.90	7.07E-04	0.80	3.82E-04	1.85 [c]		too slow	
ACTR(1046-1064)	RALLDQLHTLLSNTDATGL	14.17	11.84	0.69	1.49E-03	0.69	8.88E-04	1.68	0.69	2.32E-04	6.42
ACTR(1050-1064)	DQLHTLLSNTDATGL	10.83	9.19	0.72	2.17E-03	0.72	1.20E-03	1.82	0.72	3.62E-04	6.00
ACTR(1050-1059)	DQLHTLLSNT	6.67	5.85	0.63	2.23E-03	0.63	1.43E-03	1.56	0.63	1.61E-04	13.79
ACTR(1052-1055)	LHTL	1.67	1.22	0.64	2.32E-03	0.64	7.25E-04	3.19		too slow	
ACTR(1053-1059)	HTLLSNT	4.17	3.12	0.61	1.82E-03	0.61	1.32E-03	1.38	0.61	2.09E-04	8.68
ACTR(1053-1063)	HTLLSNTDATG	7.50	5.37	0.72	2.01E-03	0.72	1.42E-03	1.41	0.72	6.66E-04	3.02
ACTR(1053-1065)	HTLLSNTDATGLE	9.17	5.19	0.70	1.53E-03	0.70	1.49E-03	1.03	0.70	6.54E-04	2.34
ACTR(1060-1064)	DATGL	2.50	2.01	0.77	1.95E-03	0.77	7.70E-04	2.53	0.77	1.14E-03	1.71
ACTR(1065-1076)	EEIDRALGIPEL	7.50	5.61	0.79	7.52E-04	0.79	3.41E-04	2.20		too slow	
ACTR(1067-1076)	IDRALGIPEL	5.83	4.96	0.77	8.79E-04	0.77	3.61E-04	2.44		too slow	
ACTR(1077-1088)	VNQGQALEPKQD	7.50	6.38	0.57	1.18E-03	0.57	8.77E-04	1.35	0.57	8.97E-04	1.32
CBP(2059-2068)	PNRSISPSAL	5.83	4.92	0.74	2.39E-03	0.74	3.68E-03	0.65	0.74	1.14E-03	2.10
CBP(2059-2071)	PNRSISPSALQDL	8.33	6.89	0.73	1.76E-03	0.73	1.41E-03	1.25	0.73	3.62E-04	4.86
CBP(2059-2072)	PNRSISPSALQDLL	9.17	7.66	0.70	1.46E-03	0.43	8.29E-04	1.76 [c]	0.39	1.52E-04	9.61 [c]
CBP(2068-2071)	LQDL	1.67	1.46	0.80	7.59E-04	0.53	4.84E-05	15.68 [c]		no exchange	
CBP(2071-2089)	LLRTLKSPSSPQQQQQVLN	12.50	10.93	0.77	2.20E-03	0.77	7.73E-04	2.85		too slow	
CBP(2075-2088)	LKSPSSPQQQQQVL	8.33	7.53	0.74	2.49E-03	0.74	9.36E-04	2.66			
CBP(2077-2094)	SPSSPQQQQQVLNILKSN	12.50	8.33	0.67	2.06E-03	0.67	2.66E-03	0.77	0.67	2.20E-04	9.35
CBP(2088-2096)	LNILKSNPQ	5.00	4.24	0.62	1.50E-03	0.62	6.22E-04	2.42		too slow	
CBP(2089-2098)	NILKSNPQLM	5.83	5.23	0.67	1.48E-03	0.67	4.73E-04	3.13		too slow	
CBP(2091-2096)	LKSNPQ	2.50	2.00	0.69	4.61E-03	0.69	6.80E-04	6.78		too slow	
CBP(2092-2100)	KSNPQLMAA	5.00	4.24	0.78	1.99E-03	0.78	6.02E-04	3.31		too slow	
CBP(2098-2101)	MAAF	1.67	1.64	0.97	1.63E-03	0.76	4.76E-04	3.43 [c]		no exchange	
CBP(2099-2101)	AAF	0.83	0.84	1.00	1.26E-03	0.75	3.67E-04	3.42 [c]		no exchange	
CBP(2100-2109)	AFIKQRTAKY	6.67	5.16	0.85	1.91E-03	0.54	1.47E-03	1.30 [c]		too slow	
CBP(2101-2117)	FIKQRTAKYVANQPGMQ	11.67	8.97	0.74	1.84E-03	0.74	2.32E-03	0.79	0.74	2.30E-04	7.98
CBP(2110-2117)	VANQPGMQ	4.17	3.22	0.49	1.52E-03	0.49	2.41E-03	0.63	0.49	1.04E-03	1.47

[a] D_{max} is calculated as the number of exchangeable amide hydrogens adjusted for the 83.3% D_2O composition of samples undergoing deuterium exchange.

[b] *N* represents the average mass increase measured for totally deuterated controls. *N* was used in stretched exponential curve calculations as discussed in chapter 3. However, *N* was used as a constant and not an adjustable variable.

[c] Stretched exponential rate constants were calculated using stretched exponential functions that had stretching factors, β, that differed from the exponential fit for intrinsic exchange.

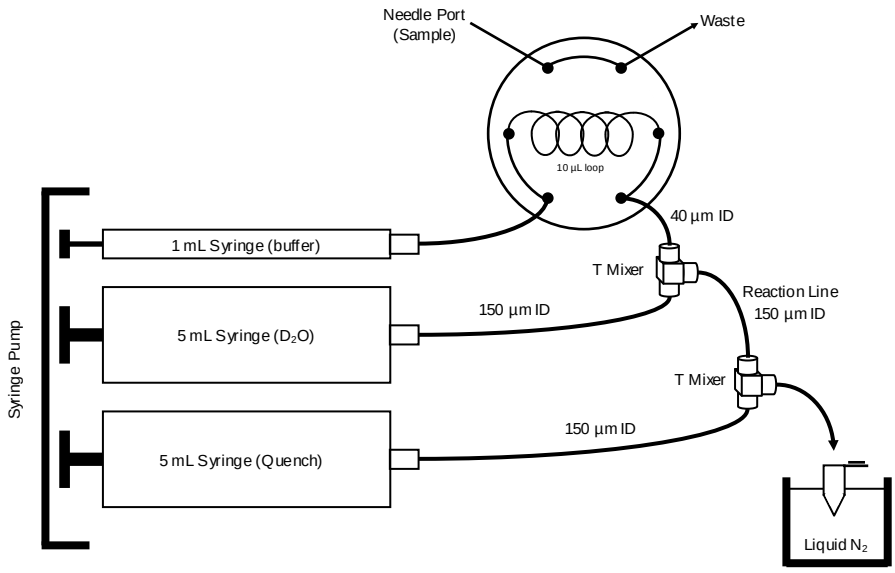


Figure 4.1. The quench-flow apparatus. The system is configured for a rapid H/D experiment, with quenched sample collected in a micro-centrifuge tube set in liquid nitrogen for rapid freezing and storage.

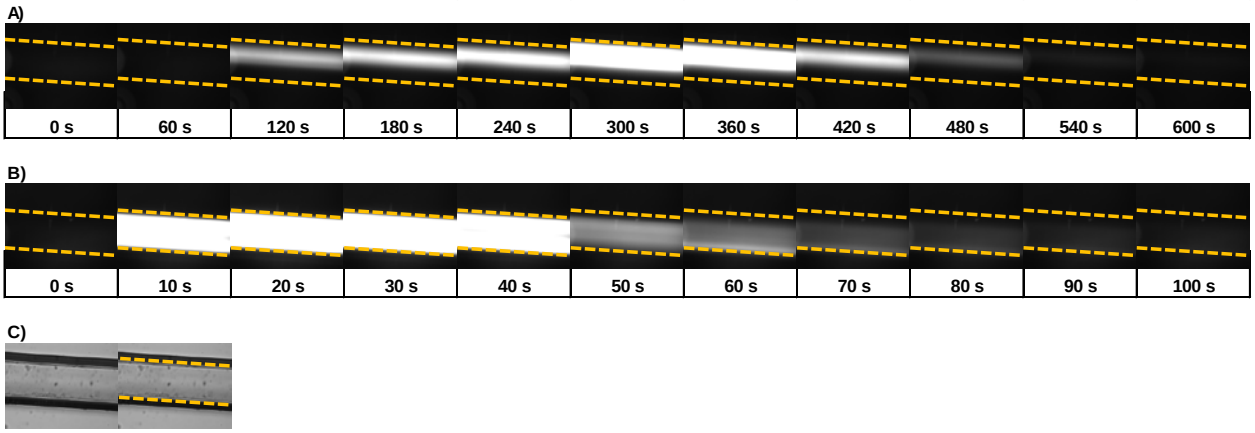


Figure 4.2. Time lapse photographs of a coating-stripped reaction capillary following injection of 10 µL of FITC-conjugated albumin with ~470 nm excitation. The capillary was photographed every 10 seconds, with 2500 ms exposures. Capillary walls are shown in yellow dashed lines. A) The capillary when sample to label flow ratio is 1:20, with 105 µL min^{-1} overall flow rate through the capillary. Due to the length of observed fluorescence, only photographs from every 60 s are shown. B) The capillary when sample to label flow ratio is 1:5, with 120 µL min^{-1} overall flow rate through the capillary. C) An ambient light photograph of the reaction capillary, which was used to locate capillary walls in fluorescence images.

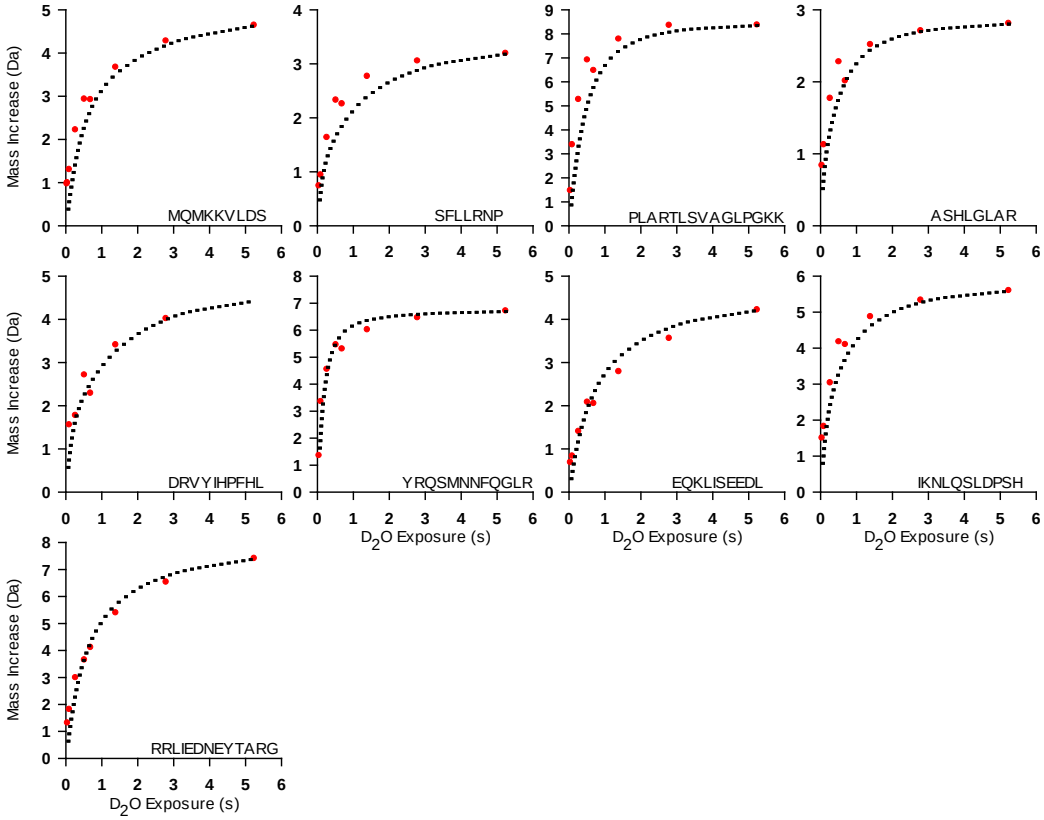


Figure 4.3. 9-peptide mix deuterium uptake curves. Red data points represent experimental average mass increase measurements. The intrinsic exchange curves, represented as the black dotted curve, were normalized to match the latest measured time point of each respective curve.

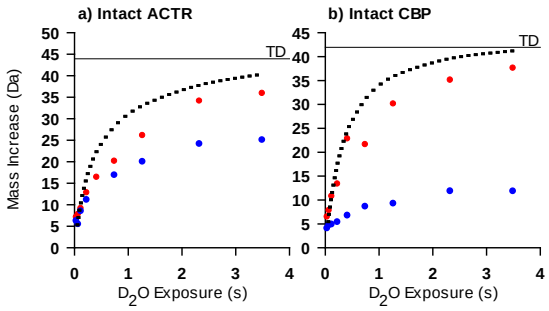


Figure 4.4. Intact protein deuterium uptake curves of (a) ACTR and (b) CBP. Totally deuterated control average mass increase results plotted as a horizontal line marked “TD”. Intrinsic exchange curves (dotted black line) have been normalized to totally deuterated samples. Deuterium uptake data shown as data points, with red points representing free protein and blue points representing bound protein.

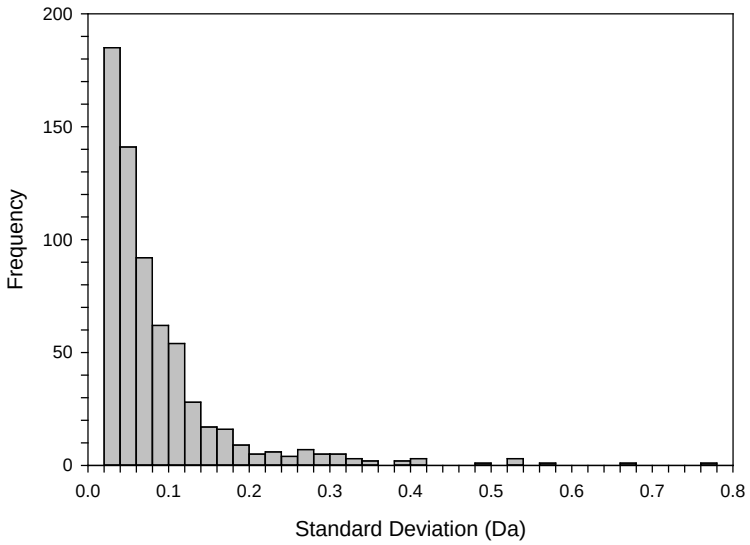


Figure 4.5. Histogram displaying the frequency of standard deviation values in CBP and ACTR peptide quench-flow H/D-MS measurements.

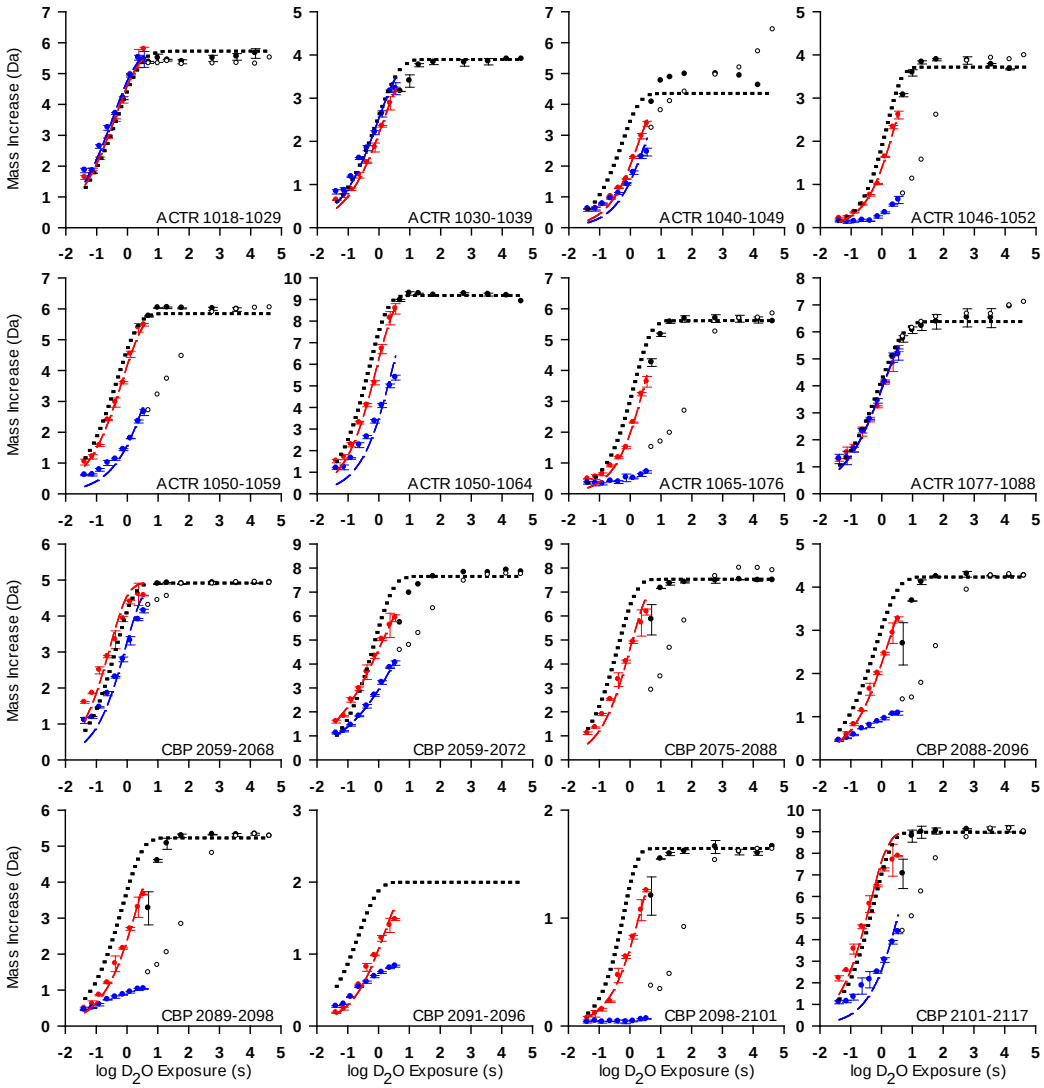


Figure 4.6. Deuterium uptake curves for peptides analyzed both in this study and in the previous work, where deuterium labeling was performed by hand (chapter 3). Previously acquired data (black filled points for free state, white filled points for bound state) were normalized to account for experimental condition changes. The average rapid H/D mass increase data are represented by red (free protein) and blue (bound in complex) colored points. The red and blue dashed lines represent the respective stretched exponential fits.

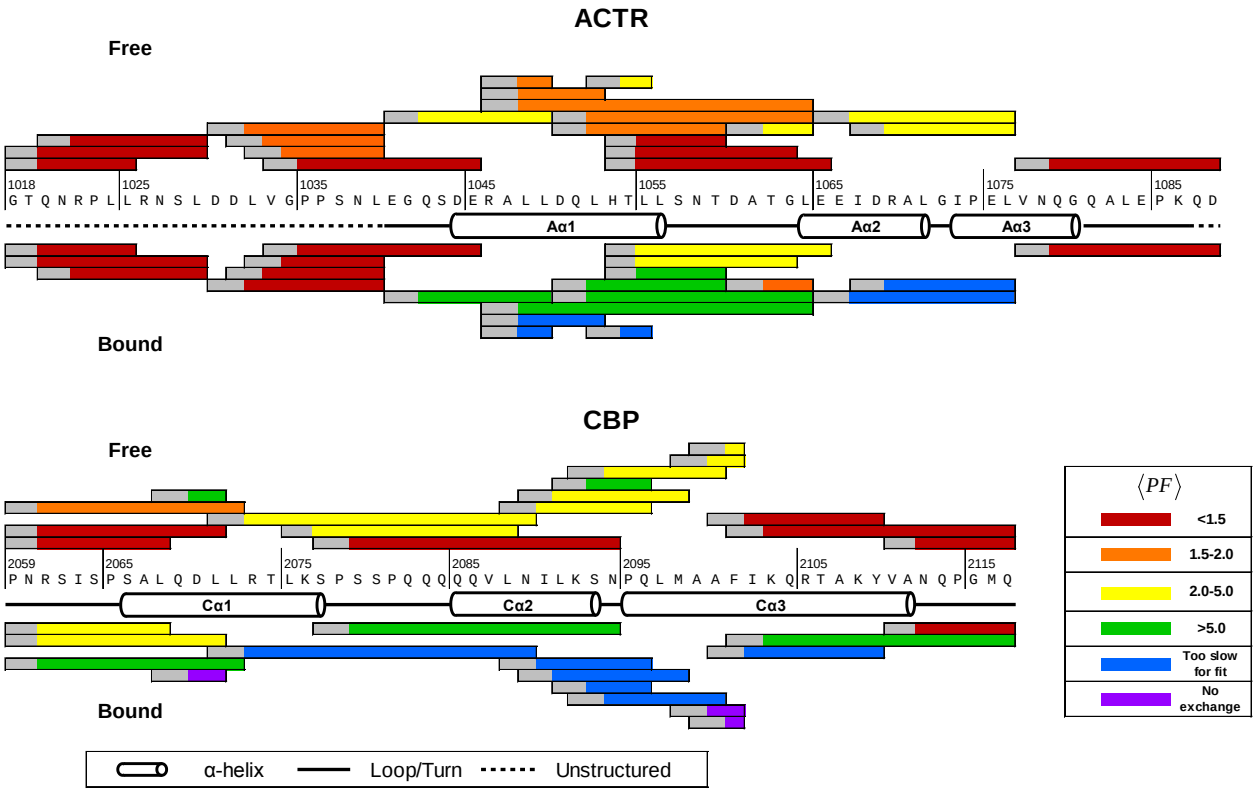


Figure 4.7. Peptide protection factors mapped onto ACTR and CBP sequences for both free and bound states. Secondary structure assignments for the CBP/ACTR complex are transposed from previous NMR studies (PDB: 1KBH).[21] Gray coloring indicates the rapid loss of deuterium label for the first two residues in each peptide due to back exchange. Peptides were determined to be too slow for stretched exponential fits when the largest rapid H/D mass increase did not reach about half of the mass increase of the totally deuterated control samples within 3500 ms exchange. Peptides with no discernible exchange within 3500 ms are also indicated.

Chapter 5: Improvements to H/D-MS Data Analysis for Multiple Protein States

5.1 Analyzing H/D-MS Data

5.1.1 Amino Acid Residue Resolution

Protein conformational changes are observed using hydrogen/deuterium (H/D) exchange experiments.[1-4] Structural perturbations, both local fluctuation events and global unfolding transitions, represent events necessary for H/D exchange in a structured protein.[5-6] The dynamic motions between such states, however, are difficult to predict and quantify accurately. The conformational stability is described by the equilibrium, K_u, between protected and unprotected states within a conformational ensemble. If single-residue resolution H/D exchange rate constants can be determined, then calculated protection factors, *PF*, can be used to calculate conformational stability, $\Delta G_u^{o'}$:[2]

$$K_u = \frac{k_{\text{op}}}{k_{\text{cl}}} \approx \frac{k_{\text{ex,EX2}}}{k_{\text{int}}} \tag{1}$$

$$\Delta G_{\text{u}}^{\ o'} = -RT \ln K_{\text{u}} = RT \ln(PF) \tag{2}$$

Determining single residue resolved exchange rate constants through H/D exchange mass spectrometry (H/D-MS) is more difficult than it is through NMR. Depending on the protein being analyzed, NMR may be an obvious choice in determining localized H/D exchange protection. The technique allows single residue resolution for solution phase, intact protein.[5-6] However, as discussed in chapter 1, NMR applications are limited by several factors. These limitations include the requirement of isolating proteins in high enough concentrations for NMR, the difficulty of studying large proteins, and the unresolvable, overlapping signals that are characteristic for disordered protein regions.[7]

Residue level localization of deuterium exchange information is a long sought after goal in H/D-MS experiments. H/D-MS offers localized information through the use of tandem MS

fragmentation, which often produces overlapping fragments. Using a residue specific analysis of overlapping fragment data, deuterium uptake localization would be improved. Electron capture dissociation (ECD) and electron transfer dissociation (ETD) have demonstrated the greatest potential for near single residue resolution.[8-11] The application of ETD-MS to study H/D exchange in a set of test peptides has been demonstrated.[9] Single residue resolution of deuterium uptake is accomplished by comparing the average mass increase, Δ*m*, of incremented c- or z-ions.[9] For example, the Δ*m* for the c6´ ion minus the Δ*m* for a c5´ ion will yield the incremental increase in Δ*m* for the sixth residue.) Another study demonstrated the analysis of horse myoglobin folding intermediates via a continuous-flow rapid mixing device and ECD-MS.[8] The fragmentation of proteins with ECD allowed short sequence increments to be analyzed, giving an average resolution of about 2.2 residues.[8] A possible advantage of top-down MS techniques, such as these, is the reduction of back exchange effects when chromatography steps are bypassed. However, bypassing chromatography requires the use of volatile buffer components. The cost and expertise required for ECD and ETD instrumentation may place such techniques out of reach of laboratories that wish to perform routine H/D-MS analyses. Large proteins, such as β-glucuronidase or monoclonal antibodies, will also prove impractically difficult to study in intact protein ECD and ETD experiments.[12] New approaches, often referred to as middle-down, combine bottom-up digests and top-down analyses to H/D exchange experiments.[13]

Computational techniques may provide an alternative for localized protection analyses. Overlapping peptide H/D-MS information was used to calculate exchange rate constants with near amino acid resolution in signal transduction kinase ERK2.[14] Peptides were partitioned into short segments with estimated rate constants of exchange. Overlapping peptide data was used to identify the residues that exhibited a specific rate constant.[14] The use of qualitatively estimated rate constants and overlapping peptide average will limit quantitative analysis of deuterium uptake.

A second computational technique integrates controls to correct for back exchange and derive protection factors at an amino acid residue resolution.[15] An automated platform to process experimental H/D-MS data was developed for analyzing proteins. The use of a peptide mixture for back exchange correction and a tetrapeptide (PPPI) for intrinsic exchange correction helped inter- and intra-lab comparability in H/D exchange experiments.[15] A modeling algorithm used experimentally measured uptake data and the above corrections to derive local protection factors. The algorithms did not derive only a single set of protection factors for a specific sequence. Instead, multiple probable protection factor solutions were calculated. While protection factors for some specific regions were consistent across many of the solutions, most regions showed large variance in derived protection factors from these same solutions. While overlapping peptide averaging may improve on the localization of estimated rate constants, amino-acid specific rate constant determination remains limited to top-down and middle-down studies. Despite limitations listed above, H/D-MS could reveal information about relative protection from exchange and conformational stability. The purpose of this chapter is to address the limitations in data analysis techniques when the above approaches cannot be applied for conventional peptide-level H/D-MS.

5.1.2 Reference States and Differential Studies

To quantify protein dynamics, measured deuterium uptake can be compared to theoretical uptake. For example, protection factors can be calculated by comparing experimental exchange to calculated intrinsic exchange.[16-18] However, protein conformation and intrinsic exchange rates are sensitive to individual experimental conditions. The minor variability in intrinsic exchange rates is often overlooked as these rates serve as a comparable reference. Many labs continue to use the same set of published rate constants to calculate intrinsic exchange, adjusting the calculation for specific pH values and temperatures.[16, 19]

Uneven sequence coverage can lead to multiple peptides covering the same region of the protein. As mentioned in section *5.1.1*, overlapping peptide information can better localize

H/D exchange information. The issue of how to average overlapping peptide information must be addressed in H/D-MS studies. While data should not be thrown out, the issue of how to describe a single residue covered by multiple peptides remains. Data from each individual peptide only represents a segmented average. This chapter explains how overlapping segments can be averaged to better resolve H/D exchange kinetics.

Deuterium uptake kinetics are often analyzed as relative differences between two protein states. Directly comparing two or more states of a single protein without using a theoretical baseline of exchange can provide useful information. These comparisons can be used to identify protein-protein contact regions and the proximal and distal effects of ligand binding.[20-21] However, comparisons made in differential analyses are often qualitative and subjective. It is difficult to classify major changes or minor fluctuations in a protein. Even when quantified, differential H/D exchange experiments often overlook factors such as peptide length. Laboratories lack a widely accepted method of analyzing H/D exchange data. Such a method would allow quantified deuterium uptake to be compared between laboratories, even though the reproducibility problems discussed above would remain. A more universally applicable procedure for assigning and quantifying changes in conformation using H/D-MS should be established.[22] The new approach outlined in this chapter delivers a possible normalization method for differential H/D-MS experiments. This approach was developed using experimental H/D-MS measurements of a model protein, β-glucuronidase. This differential analysis illustrates the localized effects of a single point mutation. The conformational rescue caused by a small molecule additive is also illustrated.

5.1.3 Restoring β-Glucuronidase Mutant Function with Allosteric Control

The β-glucuronidase protein is responsible for β-D-glucuronic acid hydrolysis. Its activity is lost when the tryptophan residue 492 is mutated to glycine.[23] The loss of this structurally important residue causes a conformational disruption that extends to the active site. When indole is added into solution with the mutant, enzyme activity is partially restored.[23] This chapter

describes the application of H/D-MS to identify the conformational changes in β-glucuronidase that result from the W492G mutation, the introduction of indole to the W492G mutant, and any effects indole may have on the wild-type protein.[24] The goal of this chapter is to present how a new data analysis method can resolve limitations in exchange localization in current H/D-MS data analysis strategies. Mutated β-glucuronidase protein provides a model that will aid in demonstrating a new strategy to localize modification effects.[24] By localizing where indole restores structural characteristics lost by mutation, it may be possible to form a link between region and function.

5.2 Experimental

5.2.1 β-Glucuronidase Experiment Overview

Wild-type and W492G mutant β-glucuronidase stocks were provided by Professor John Karanicolas (University of Kansas, Lawrence, KS) at a concentration of 5 μg μL^{-1} in 50 mM sodium phosphate, 100 mM sodium chloride, pH 7.4. Protein stocks of both wild-type and W492G mutant were diluted to 0.5 μg μL^{-1} aliquots in 90% water / 10% ethanol, 50 mM sodium phosphate, 100 mM sodium chloride. Since indole addition was to be tested, one aliquot of wild-type and one of W492G mutant contained 5 mM indole. Ethanol was used to initially dissolve solid indole.

All deuterium exchange experiments were carried out by a LEAP HDX-PAL robot (LEAP Technologies, Carrboro, NC). The robot was used to mix 2 μL of protein solution with 40 μL deuterium exchange buffer, followed by the addition of 40 μL quench buffer after the appropriate labeling time. The deuterium exchange buffer was 50 mM sodium phosphate, 100 mM sodium chloride, with and without 5 mM indole in 90% D_2O, 10% ethanol, pD 7.4. Undeuterated control samples were prepared using similar buffers which contained 90% H_2O instead of D_2O, pH 7.4. All deuterium exchange buffers were kept at ambient room temperature. The exchange quench buffer was 0.75 M hydrochloric acid, kept at 1°C. Deuterium exchange labeling times ranged from 10 seconds to 12 hours. Following the labeling, quench buffer was added to sample at a

1:1 volume ratio. Quenched samples were immediately injected onto a 100 µL loop prior to pepsin digestion and peptide separation (see chapter 2).

An isocratic pump flowed 0.1% formic acid at 0.2 mL min^{-1} through the sample loop to inject samples through an immobilized pepsin column, prepared in-house, for online proteolysis of labeled samples[3] (see chapter 2). Peptides were captured on a C12 trap, packed in-house. A 4.5-40.5% acetonitrile gradient over 10 minutes was used to separate peptides on a 50 mm × 1 mm Ascentis Express ES-C18 column (2.7 µm particle size, 160 Å pore size, Supelco Analytical, Bellefonte, PA). All HPLC components were kept at 0°C to reduce back-exchange. Eluted peptides were analyzed by TOF-MS (Agilent Technologies, Santa Clara, CA). HDExaminer (Sierra Analytics, Modesto, CA) was used to obtain average mass values for peptide spectra at each labeling time and for the undeuterated controls.

5.3 Analysis of Differential H/D Exchange Data

5.3.1 Butterfly Plots

A general method to interpret H/D-MS data should be developed. When comparing two protein states using H/D-MS, one of the simplest qualitative assessments may come from comparing the two deuterium uptake curves. Even a survey of deuteration at each time point can provide an idea of whether a specific peptide gains or loses protection, or stays about the same between two conditions. Figure 5.1 shows four example plots with four deuterium uptake curves each. These curves represent typical results from H/D-MS analyses, where the average deuterium uptake is plotted as a function of D_2O exposure time. The scaling of each y-axis to the maximum number of exchangeable amide hydrogens for a given peptide allows relative exchange to be observed when comparing plots. There is little difference between wild-type and mutant deuterium uptake in peptide #61 as shown in figure 5.1. The W492G point mutation in β-glucuronidase causes a loss of exchange protection in regions distant from the mutation site, such as in peptides #75 and #76. The mutant protein's peptide #140 is also less protected from exchange when compared to the wild-type protein. The comparison of deuterium uptake

curves from one peptide to another may be more difficult due to differences in sequence. The sequence influences intrinsic exchange behavior, as intrinsic exchange rate constants will depend on amino acid residue identity and the identity of neighboring residues. Peptide length is also an important factor to consider when comparing different peptides, as the complete peptide sequence will determine the number of exchangeable amide hydrogens. Simplifying the comparison of two curves with an overall difference measurement provides a comparable quantity that may be plotted along a list of peptides.

Differential exchange data can be presented graphically in a “butterfly” plot. Such a plot facilitates the comparison between two experimental conditions. In butterfly plots, deuterium uptake is quantified using a relative fractional exchange; the average deuterium uptake is divided by the maximum number of exchangeable amide hydrogens. This quantity represents what can be inferred from figure 5.1. The relative fractional exchange for a reference experiment is plotted on the positive y-axis. Exchange for an experimental condition change is mirrored and plotted along the negative y-axis. Peptide position, ordered by its sequence location midpoint, is plotted along the x-axis. The relative exchange at each time point is represented by a differently colored plot. This representation can provide a comparison of where exchange occurs slower or faster along the protein, so a degree of protection can be inferred. A quick survey of these mirrored plots can also reveal where deuterium uptake behavior differs between two states. Multiple preparations of interferon-β-1a were compared using an H/D-MS method and butterfly plots.[25]

A graph, such as figure 5.3, illustrates deuterium uptake differences between two conditions in measured peptides. For each peptide, the difference in average mass increase values, $\Delta\Delta m$, is calculated at a single time point, t, between the mass increase of a conditionally altered state, Δm_i, and that of a reference state, Δm_i^{ref}. Figure 5.2 illustrates how $\Delta\Delta m$ is determined between two curves at a single time point. On a graph such as figure 5.3, the sum

of the differences in deuterium uptake, $\Sigma\Delta\Delta m$, for all time points, n_t, is plotted for each peptide to compare two protein states.

$$\sum \Delta\Delta m = \sum_{i=1}^{n_t} \left(\Delta m_i - \Delta m_i^{\text{ref}} \right) \tag{3}$$

This summed difference can be positive or negative in value. Protection factor may increase (a negative difference) or decrease (a positive difference) in transitioning from one form to the second. The difference in deuterium uptake plots are summed for each peptide ordered along the x-axis. Using β-glucuronidase data, figure 5.3 shows a graph where data is prepared in this manner. Figure 5.3 illustrates the deuterium uptake differences between mutant β-glucuronidase and wild-type protein (reference) as red bars. Most of these uptake differences show an increase in deuterium uptake, and thus lower protection, across most of the mutant protein segments. Figure 5.3 also illustrates deuterium uptake differences in mutant β-glucuronidase with added indole as the conditional change and absence of indole as the reference state. For most segments, the addition of indole causes a decrease in deuterium exchange.

Presenting the data in this manner does not represent each peptide's length. A peptide may have a small summed deuterium uptake difference, such as ±1 Da. If that difference is the result of one measurable amide gaining complete protection from previously having no protection, then it should still be considered a significant change. On the other hand, a ±1 Da summed difference may also be the resulting sum of a minor structural fluctuation spread across many residues. That is, a +1 Da difference on a 4 residue peptide shows a greater overall effect than a +1 Da difference on an 11 residue peptide. For example, the summed deuterium uptake difference between mutant and wild-type protein in peptide #80, which is 4 residues long, and peptide #129, which is 11 residues long, are very similar at around +4.6 Da. Thus, the relative difference between mutant and wild-type protein in peptide #80 is much larger

than that of peptide #129. The important observation of a large relative difference in a small peptide may be missed as it may not produce a large value on the difference plot.

Referring back to the difference plot in figure 5.3 and the individual uptake plots in figure 5.1, the large difference between mutant and wild-type at peptide #76 (residues 331-357) is +18.4 Da. This is a result of this particular peptide containing 24 exchangeable amide hydrogens. The magnitude of 18.4 Da almost reaches the difference magnitude of +22.2 Da for peptide #140 (residues 584-609), which contains 23 exchangeable amide hydrogens. Peptide #76 and peptide #140, therefore, share similar summed difference magnitudes, but the uptake curves display differing uptake kinetics. However, looking at peptide #75 (residues 331-339), which contains 7 exchangeable hydrogens located within peptide #76's sequence, a summed difference of +3.8 Da is observed. The uptake curves of peptides #75 and #76 share a similar profile, implying similar uptake kinetics. The huge difference in the summed difference values between peptide #75 (+3.8 Da) and #76 (+18.4 Da) makes it difficult to infer this similarity between the peptides as plotted in figure 5.3 because peptide length is not represented. A mathematical normalization to account for peptide length would allow direct comparison between these peptides.

A similar plot in an earlier study of interferon-β-1a addressed this issue of comparing peptide lengths by plotting the relative length of peptides alongside fractional exchange data.[26] Peptides were ordered by their numerical sequence midpoint values to form a "quasi-sequence."[25] Figure 5.4(A) shows a mock list of peptides mapped onto a brief sequence. Figure 5.4(B) shows how the mock list is translated to a quasi-sequence. A normalized exchange difference plot would better aid in the visualization of differential exchange data as opposed to an absolute difference plot. Absent from the plot in figure 5.4(B) is a depiction of which regions of the protein are represented and which regions are not covered. The peptides used may have high overlap in one particular region, such as between residues 9 and 19. These peptides, which are colored in teal, dominate the quasi-sequence despite covering less

than half of the actual sequence. Other regions may have no peptide coverage at all, such as between residues 20 and 24. All peptides are presented as evenly spaced integers in figure 5.4(B). Regions with no peptide coverage are not represented in a plot such as figure 5.4(B). Mock deuterium uptake values are plotted along the quasi-sequence in figure 5.4(B). While deuterium uptake for the teal colored peptides appears greater, it is difficult to infer relative exchange from figure 5.4(B) alone because peptide length is not established.

5.3.2 Relative Deuterium Uptake Differences

Relative deuterium uptake measurements allow two different peptides' deuterium uptake behavior to be compared. Differential deuterium exchange analyses can average the uptake difference between two curves at every measured time point to calculate and overall fractional difference between two conditions. An average relative deuteration difference, *ARDD*, was used to quantify exchange protection in free and complex-bound recombinant cashew 11S globulin allergen, rAna o 2.[21] ARDD values were calculated as the relative difference between the deuterium uptake of one conditional state and a reference state, averaged across all exchange time points:

$$ARDD = \frac{\sum_{i}^{n_t} \left(\frac{\Delta m_i^{\text{ref}} - \Delta m_i}{\Delta m_i^{\text{ref}}} \right)}{n_t} \tag{4}$$

Figure 5.2 illustrates two mock deuterium uptake curves being compared. The strategy of averaging across all time points places equal statistical weight on each time point. This may skew the average toward a range of closely spaced times points. That is, if many points are taken for early time points, as shown in the left-side blue box on figure 5.2, and fewer at later times, the average skews toward early exchange behavior rather than exchange across the entire range. Another possible problem with such an interpretation arises if two curves have many time points where the difference was near zero, such as in the two blue boxes on figure

5.2. These points would be measured before any significant exchange occurs or after complete exchange occurs in both conditions. The abundance of no-difference points may dilute the average difference in a way that overall difference information, shown by the large ΔΔm region, may be lost. Curve differences are better described when time points are spaced evenly along a logarithmic scale and entire uptake curve profiles are covered by the D_2O exposure time range.

5.3.3 Stretched Exponential Fits

To eliminate the problem of uneven time point spacing, deuterium uptake curves could be described as a function of a peptide's deuterium exchange rate constant. Rather than compare two curves point by point, a comparison of segment-averaged rate constants for a particular peptide would account for an overall difference in apparent deuterium uptake curves. This method of comparison was discussed in the study of CBP and ACTR free and bound states (see chapter 3).[18] However, rate constants may be difficult to determine for a number of exchange behaviors. Rapidly exchanging peptides may become fully deuterated before the smallest measurable D_2O exposure time (a limitation addressed in chapter 4). This results in an almost horizontal uptake curve near the value of the totally deuterated control. Very slowly exchanging peptides may not show measurable deuterium uptake within an experiment's longest time point (a limitation encountered in chapter 4). This slow exchange results an almost horizontal uptake curve near 0 Da of mass increase. A peptide may also exhibit multi-modal and/or slow deuterium uptake, which cannot described by any single, simple curve fit function. Previous work has demonstrated rate constant determination for bi-modal deuterium uptake curves,[17] but such calculations introduce greater uncertainty as more parameters are incorporated into the interpretation. Subjective alteration to fit parameters may also reduce quantitative validity. The approach outlined in the next section introduces an averaging technique to better localize differential amide H/D exchange data between multiple conditions.

The availability of overlapping peptides also raises another issue in data interpretation. To assign specific values to specific protein segments, extra peptide information can be discarded or merged mathematically or subjectively together. An averaging of overlapping peptide information allows more useful experimental results to be included in the analysis and improves information localization, even if it is not truly single amino acid resolution. The next section also explains how overlapping peptide data is averaged in order to improve H/D exchange localization.

5.4 Normalized Deuterium Difference Approach

5.4.1 Normalization of Deuteration Differences

An alternative method to compare deuterium uptake kinetics between a condition and a reference should include a normalized measurement of their differences in exchange. The first step in the data processing for this normalization is to calculate the summed mass difference between two deuterium uptake curves, $\Sigma\Delta\Delta m$ using equation 3. This is done at each individual time point, t, for a specific peptide. Individual time point mass increase difference values, $\Delta m_i - \Delta m_i^{\text{ref}}$ in equation 3, are calculated and then summed from all measured time points. The resulting sum value is assigned to that peptide's location and sequence.

Peptides will vary in length and time points may be lost due to limited detection or experimental conditions. It is because of these reasons that these uptake curve difference sums alone will not be comparable from peptide to peptide. Instead, normalized deuterium difference, *NDD*, values should be calculated to eliminate the factors of peptide length and number of measured time points from comparisons of uptake behavior.

$$NDD = \frac{\sum_{i=1}^{n_t} \left(\Delta m_i - \Delta m_i^{\text{ref}} \right)}{D_{\max} \times n_t} \tag{5}$$

By normalizing the sum of the differences between uptake curves by the maximum number of exchangeable amide hydrogens for a specific peptide, D_{max}, the problem caused by varied peptide length is addressed (see chapters 3 and 4 for determination of D_{max} values based on peptide sequence). The sum is also normalized by the number of usable time points that are present in both uptake curves, n_t. Therefore, if one curve is missing one or more time points, those time points will not be used in subsequent analyses of either curve. This also ensures that uptake difference results are not exaggerated by a particular peptide having more usable time points than another peptide. These *NDD* values can range from -1 to 1. A value at either extreme represents the difference between no deuteration and complete deuteration at every time point, with no back-exchange. While the extremes represent an unrealistic possibility, they provide an easy to interpret boundary by which deuterium uptake curves for a single peptide or group of peptides can be compared.

5.4.2 Averaging Overlapping Peptides

Once a particular peptide's deuterium uptake data has been normalized, it may be compared to the data from another peptide. Table 5.1 shows an abbreviated example spreadsheet calculator to average NDD values on a per-residue basis. The amino acid sequence is listed in individual cells within a single row on the top of spreadsheet. Residue start and end numbers are ordered along the two columns on the left. The NDD value is assigned to all residues in the peptide, except for the first two residues and proline residues. The first two residues of each peptide are listed as cells with no data, "ND", as these residues will rapidly back-exchange for H_2O in the HPLC solvent.[16, 19] Similarly, cells list "ND" for proline residues, as proline residues do not have amide hydrogens. Once every peptide's NDD values are listed, the NDD data for each residue is then averaged. This NDD average will describe whether the deuterium uptake data showed an increase or decrease in protection relative to a reference state. Negative values represent an increase in exchange protection in relation to the reference state. Positive values represent a decrease in exchange protection in relation to the reference

state. By including all measured peptides, sub-peptide resolved information is obtained for overlapped regions. The region shown in Table 5.1 is covered by several overlapping peptides. NDD values for most peptides in this region were measured within a range of 0.011 to 0.030. The NDD value of 0.000 for peptide (23-40) is anomalously small compared to similar peptides covering the same region. By using the experimental data averaging strategy described above, peptide (23-40)'s small NDD value is effectively overruled by the availability of other peptides covering the same region. While this strategy can aid in difference localization, the measured average NDD values should not be considered truly residue-resolved.

5.4.3 Mapping the Regions of Maximum Differences

While a quantitative measurement of a condition's effects is important, a graphical presentation may be more comprehensible. In a spreadsheet bar chart, a protein's sequence numbering is plotted along the x-axis and the per-residue NDD is plotted along the y-axis. By using partial transparency and different coloring schemes, two sets of conditional changes can be easily compared on a single graph. Examples of this type of chart are presented in figure 5.5, which are representations of the data for the β-glucuronidase experiment. These numerical, per-residue averages can also be used as the basis for a color scale representation onto a model structure.

5.4.4 Estimated Significance Threshold

H/D-MS is highly sensitive to changes in protein structure and results can be highly reproducible. However, determining what constitutes a significant difference between two states of a single peptide requires uncertainty calculations provided by replicate sample analyses. A single standard deviation of ±0.14 Da for a replicated time point (*i.e.* a Δm measurement) with a 98% confidence interval of about ±0.5 Da has been reported in previous work.[25-26] The confidence interval appears to be unaffected by labeling time, peptide length, or mass difference caused by deuterium uptake.[25] A difference between two peptide states, $\Delta\Delta m$, that exceeds this confidence interval at any time point could then be considered statistically

significant. The ±0.14 Da standard deviation may be used to calculate a statistical significance threshold between multiple datasets which contain all time points, $\Sigma\Delta\Delta m$, using the Student's t-distribution. (In the work noted above, 5 time points were sampled to give a 98% confidence interval of ±1.1 Da).[25] Because of single trial sampling conducted for this chapter, uncertainty calculations are omitted. Data herein is interpreted in terms of simple relative observations, instead. If uncertainty and artificial triplicate sampling are applied to this work and assumed to be similar to that of the previous work, the significance threshold would be calculated to be about ±1.7 Da for the 9 sampled time points. Peptides with only 1 exchangeable amide hydrogen could not meet this threshold, despite showing obvious effects on deuterium uptake relative to a reference state.

5.5 Mutated β-Glucuronidase Rescue

5.5.1 Effects of Indole on Wild-Type β-Glucuronidase

β-glucuronidase protein activity can be inhibited by a single point mutation at residue 492. Residue 492 is tryptophan in the wild-type protein. The mutation of residue 492 to glycine represents the loss of an indole for that residue's side chain.[23] Activity may be partially restored in the W492G mutant protein by the addition of indole to the solution. This work outlines how H/D-MS is used to identify changes that result from the mutation as well as from the addition of indole. As a control experiment, the effect of indole on the deuteration of wild-type β-glucuronidase protein was measured. The indole is expected to have no significant effects on the structure of wild-type protein. Indeed, the addition of indole molecule to wild-type protein does not alter its enzymatic activity.[23-24] The effect of indole on wild-type protein was measured using NDD values relative to wild-type protein without indole. The results of this control experiment are graphed in figure 5.5(A). NDD values calculated from equation 5 range from -0.033 to 0.032 for 220 analyzed peptides. Within these 220 analyzed peptides, only 50 have difference values beyond a -0.010 to 0.010 range. As will be shown below, these 50 largest NDD magnitudes for this control experiment are relatively small compared to the magnitudes

given by the effects of mutation and the effects indole has on the mutant. A significance threshold of |NDD|>0.050 would represent an average deuteration difference of 5% of D_{max} at all measured time points between two deuterium uptake curves. Because the subsequent experiments show much greater ranges of normalized deuteration difference values, this control experiment provides insight into the inherent variability of the deuterium exchange method and any slight effects the indole may have on exchange.

5.5.2 Effects of Mutation on Wild-Type β-Glucuronidase

The mutation of tryptophan to glycine at residue 492 of β-glucuronidase removes an indole side chain from the protein. The effect of the W492G mutation was measured using NDD values relative to wild-type protein. The averaged normalized deuteration differences for the mutation effects are mapped in figure 5.5(B) as red bars, with proline residues and regions with no peptide coverage occupying the gaps in the plot. The W492G mutation causes pronounced changes across much of the protein, shown in figure 5.5(B). Most peptides show an increase in deuterium uptake when compared to the wild-type protein. One noticeable effect occurs between residues 65 and 89, where these difference values range from +0.021 to +0.120. These positive values indicate a loss of protection from exchange. The region between residues 303 and 356 also show an increase in deuterium uptake, with difference values between +0.046 and +0.089. Some shorter regions around residue 377 and between residues 401 and 406 have difference values greater than +0.120, as well. Starting at about residue 433, the difference values become even greater in magnitude. Beyond residue 433 are the regions of the protein with closer structural proximity to the mutation site. Many portions of these regions, until about residue 603, have measured difference values greater than +0.100, with the maximum difference value of +0.295 at residues 547-549.

Measured mutation effect NDD values are mapped onto a three dimensional crystal structure of β-glucuronidase in figure 5.6(A). Figure 5.6(A) shows a scaled color scheme of the measured NDD values with green representing positive differences and magenta representing

negative differences. Gaps in coverage and proline residues are colored yellow. The largest effects of mutation reside in the α-helical lobe located around residue 492 in the crystal structure. The loss of protection from exchange across much of the protein may indicate a partial destabilization process that results from the replacement of the tryptophan residue with glycine. The measured difference values provide information on the destabilization of amide hydrogen bonding in the protein structures, but they do not provide information on secondary structural elements themselves. Further analysis of the kinetics of deuterium uptake at the peptide level can also provide insight toward specific conformational elements.

5.5.3 Effects of Indole on Mutated β-Glucuronidase

Indole was added to the sample buffer and deuterium labeling buffer to understand how indole rescued the W492G mutated β-glucuronidase.[23] W492G mutant β-glucuronidase acts as the reference state in this experiment. The replacement of tryptophan with glycine at residue 492 destabilized many regions of the protein and increased deuterium exchange in those regions, as discussed in section *5.5.2* and seen in figure 5.5(B) as red bars. This destabilization correlates to the loss in enzymatic activity in the mutant protein.[23] Figure 5.5(B) shows the averaged normalized deuteration difference values for the indole effects on W492G mutant β-glucuronidase as blue bars. In contrast to the mutation effects, many regions possess negative difference values, indicating that indole increased protection in these regions. Indeed, in some regions, such as around residues 76, 276, and 403, the difference values indicate that indole has the opposite effect of mutation. This implies that indole returns some protection to regions destabilized by the mutation. However, the smaller magnitudes of the NDD values indicate that wild-type levels of protection are not reached.

Some regions gain protection from exchange even beyond that measured for the wild-type protein. Residues 170-185 have a measured difference value of -0.053, a magnitude that exceeds the +0.012 difference value measured for the mutation effect alone. Another region, between residues 380 and 406, shows increase in protection for the mutant β-glucuronidase,

though these changes do not have a directly opposite effect from the mutation effects. Some regions show a destabilization effect when indole is included, as indicated by positive average NDD values. Between residues 410 and 432, the difference values range from +0.023 to +0.048. Similarly, difference values range from +0.022 to +0.075 between residues 445 and 466. The largest negative difference values for the indole effects lie in the same regions of the largest positive values for the mutation effects. Between residues 547 and 549, where the mutation effects have the largest difference values, the difference values measured for the effect of indole on mutated protein are -0.095.

Measured indole effect on mutant protein NDD values are mapped onto the crystal structure of β-glucuronidase in figure 5.6(B). Figure 5.6(B) shows a scaled color scheme of the measured NDD values for these indole effects, with green, again, representing positive differences and magenta representing negative differences. The largest protection recovery effects reside in the α-helical lobe around residue 492. Residue 492, shown as red spheres in figure 5.6(B), is where the largest effects of mutation were also located. The pairing of the opposing maxima between the mutation effects on wild-type protein and indole on the mutant protein provides some important insights toward the dynamics of the protein structure. To reiterate, the W492G mutation destabilizes much of the β-glucuronidase protein, with the most pronounced increases in deuterium uptake located along the sequence close to the mutation site. These peptides which show the greatest deuterium uptake increase as a result of mutation also show a decrease in deuterium uptake once indole is added. While the indole fails to match the magnitude of difference that the mutation causes, it still has a stabilizing effect on the protein. This implies that indole partially counteracts the mutation effects around the mutation site. The shortfall may result from indole concentration limits and a difference between a structurally bound indole group, as with tryptophan, and indole simply being non-covalently bound in the indole pocket.

An assumption inherent in the calculations of NDD values is that every deuterium uptake time point carries the same statistical weight. The spacing of time points is often arbitrarily chosen by investigators to cover a broad range on a logarithmic scale. Labeling will often range from several seconds to days.[18, 27-28] The midpoint of exchange may change when altering the conditions under which exchange is taking place. If this midpoint lies out of the range of time points under one or both of conditions being compared, then the full magnitude of these difference values is not entirely realized. Again, figure 5.2 compares two uptake plots. If time points are concentrated in either of the regions indicated by the blue boxes, then any comparison between the two curves will be skewed by the incompleteness and/or unevenness of sampled time points.

5.5.4 Normalized Deuterium Uptake

The information provided by ΔΔ*m* values alone excludes a quantitation of conformational dynamics. A limitation to interpreting deuterium uptake data solely as NDD values is that it provides no indication of native flexibility for each peptide studied. A measurement of absolute deuterium uptake can aid in localizing protection or flexibility. In the β-glucuronidase experiment, using a normalized deuterium uptake value, *NDU*, calculated as in equation 6, can measure absolute deuterium uptake in any one condition:

$$NDU = \frac{\sum_{i=1}^{n_t} \left(\Delta m_i \right)}{D_{\max} \times n_t} \tag{6}$$

An NDU value differs from an NDD value in that mass increase is not related to any reference state in an NDU calculation. An NDU value provides a normalized value of deuterium uptake that can be compared between peptides. Figure 5.7 shows an example representation of NDU values calculated for the four β-glucuronidase conditions discussed in this chapter (wild-type, wild-type with indole, W496G mutant, W492G mutant with indole). The NDU values for wild-

type protein without indole are mapped onto a three-dimensional model of β-glucuronidase as a white to red heat map in figure 5.8 (Gaps in peptide coverage and proline residues are, again, colored yellow). This mapping of NDU values can provide a survey of backbone dynamics and flexibility. A redder region in figure 5.8 corresponds to a larger NDU value and a faster deuterium uptake. Regions more protected from exchange will have smaller NDU values and will be colored whiter in figure 5.8.

Surveying deuterium uptake plots (figure 5.2) can indicate whether a peptide is exchanging rapidly (less protection) or more slowly (more protection) based on expected maximum exchange values. Figure 5.7 shows NDU values measured for all four conditions of the β-glucuronidase protein. Exchange is relatively slow for all conditions between residues 204 and 365, with NDU values measured below 0.35 throughout. Some regions, such as those near residues 161, 378, 489, and 576, exhibit relatively fast exchange with NDU values above 0.45. Measuring the difference between two plots alone, such as with NDD values, will not include this protection information. As a result, folding or protection changing events cannot be readily quantified. A major structural conversion is indistinguishable from a less severe stabilization event with deuterium uptake difference information alone. The inclusion of a theoretical dataset representing intrinsic exchange, Δm_i^{int}, could provide another degree by which the multiple experimental datasets could be compared. Calculations of Δm_i^{int} can be conducted as discussed in chapter 3. By calculating a normalized deuterium difference from intrinsic exchange, *NDDI*, this comparison may give some indication of initial protection present for each peptide:

$$NDDI = \frac{\sum_{i=1}^{n_t} \left(\Delta m_i - \Delta m_i^{\text{int}} \right)}{D_{\max} \times n_t} \tag{7}$$

NDDI values are calculated using equation 7 in the same manner as NDD values are calculated using equation 5, with the exception being that intrinsic exchange represents the reference state for the analyzed peptide. A negative NDDI value, like a negative NDD value, will indicate an increase in protection. NDDI values will always be expected to be negative, since intrinsic exchange is the expected deuterium exchange for a protein as if it were a random coil.

References

1. Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L., Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100* (12), 7057-7062.
2. Konermann, L.; Tong, X.; Pan, Y., Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *J. Mass Spectrom.* **2008,** *43* (8), 1021-1036.
3. Wang, L.; Pan, H.; Smith, D. L., Hydrogen Exchange-Mass Spectrometry: Optimization of Digestion Conditions. *Mol. Cell. Proteomics* **2002,** *1* (2), 132-138.
4. Wales, T. E.; Engen, J. R., Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* **2006,** *25* (1), 158-170.
5. 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.
6. Skinner, J. J.; Lim, W. K.; Bédard, S.; Black, B. E.; Englander, S. W., Protein hydrogen exchange: Testing current models. *Protein Sci.* **2012,** *21* (7), 987-995.
7. Sharma, S.; Zheng, H.; Huang, Y. J.; Ertekin, A.; Hamuro, Y.; Rossi, P.; Tejero, R.; Acton, T. B.; Xiao, R.; Jiang, M.; Zhao, L.; Ma, L.-C.; Swapna, G. V. T.; Aramini, J. M.; Montelione, G. T., Construct optimization for protein NMR structure analysis using amide hydrogen/deuterium exchange mass spectrometry. *Proteins* **2009,** *76* (4), 882-894.
8. Pan, J.; Han, J.; Borchers, C. H.; Konermann, L., Characterizing Short-Lived Protein Folding Intermediates by Top-Down Hydrogen Exchange Mass Spectrometry. *Anal. Chem.* **2010,** *82* (20), 8591-8597.
9. Zehl, M.; Rand, K. D.; Jensen, O. N.; Jørgensen, T. J. D., Electron Transfer Dissociation Facilitates the Measurement of Deuterium Incorporation into Selectively Labeled Peptides with Single Residue Resolution. *J. Am. Chem. Soc.* **2008,** *130* (51), 17453-17459.
10. Kweon, H. K.; Hakansson, K., Site-specific amide hydrogen exchange in melittin probed by electron capture dissociation Fourier transform ion cyclotron resonance mass spectrometry. *Analyst* **2006,** *131* (2), 275-280.
11. Charlebois, J. P.; Patrie, S. M.; Kelleher, N. L., Electron Capture Dissociation and 13C,15N Depletion for Deuterium Localization in Intact Proteins after Solution-Phase Exchange. *Anal. Chem.* **2003,** *75* (13), 3263-3266.
12. Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L., On the Scalability and Requirements of Whole Protein Mass Spectrometry. *Anal. Chem.* **2011,** *83* (17), 6868-6874.
13. Rand, K. D., Pinpointing changes in higher-order protein structure by hydrogen/deuterium exchange coupled to electron transfer dissociation mass spectrometry. *Int. J. Mass spectrom.* **2012**, Article in press.
14. Resing, K.; Hoofnagle, A.; Ahn, N., Modeling deuterium exchange behavior of ERK2 using pepsin mapping to probe secondary structure. *J. Am. Soc. Mass. Spectrom.* **1999,** *10* (8), 685-702.

15. Zhang, Z.; Zhang, A.; Xiao, G., Improved Protein Hydrogen/Deuterium Exchange Mass Spectrometry Platform with Fully Automated Data Processing. *Anal. Chem.* **2012,** *84* (11), 4942-4949.

16. Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W., Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 75-86.

17. Chetty, P. S.; Mayne, L.; Lund-Katz, S.; Stranz, D.; Englander, S. W.; Phillips, M. C., Helical structure and stability in human apolipoprotein A-I by hydrogen exchange and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (45), 19005-19010.

18. 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.

19. Connelly, G. P.; Bai, Y.; Jeng, M.-F.; Englander, S. W., Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins* **1993,** *17*, 87-92.

20. Houde, D.; Berkowitz, S. A., Conformational comparability of factor IX–Fc fusion protein, factor IX, and purified Fc fragment in the absence and presence of calcium. *J. Pharm. Sci.* **2012,** *101* (5), 1688-1700.

21. Zhang, Q.; Willison, L. N.; Tripathi, P.; Sathe, S. K.; Roux, K. H.; Emmett, M. R.; Blakney, G. T.; Zhang, H.-M.; Marshall, A. G., Epitope Mapping of a 95 kDa Antigen in Complex with Antibody by Solution-Phase Amide Backbone Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2011,** *83* (18), 7129-7136.

22. Chalmers, M. J.; Pascal, B. D.; Willis, S.; Zhang, J.; Iturria, S. J.; Dodge, J. A.; Griffin, P. R., Methods for the analysis of high precision differential hydrogen–deuterium exchange data. *Int. J. Mass spectrom.* **2011,** *302* (1–3), 59-68.

23. Deckert, K.; Budiardjo, S. J.; Brunner, L. C.; Lovell, S.; Karanicolas, J., Designing Allosteric Control into Enzymes by Chemical Rescue of Structure. *J. Am. Chem. Soc.* **2012,** *134* (24), 10055-10060.

24. Karanicolas, J. Weis Lab Collaboration: β-Glucuronidase Project. Personal Communication, Lawrence, KS, **2012**.

25. Houde, D.; Berkowitz, S. A.; Engen, J. R., The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J. Pharm. Sci.* **2011,** *100* (6), 2071-2086.

26. Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D., Conformation and Dynamics of Biopharmaceuticals: Transition of Mass Spectrometry-Based Tools from Academe to Industry. *J. Am. Soc. Mass. Spectrom.* **2010,** *21* (3), 323-337.

27. Hsu, Y.-H.; Burke, J. E.; Li, S.; Woods, V. L.; Dennis, E. A., Localizing the Membrane Binding Region of Group VIA Ca2+-independent Phospholipase A2 Using Peptide Amide Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Biol. Chem.* **2009,** *284* (35), 23652-23661.

28. Houde, D.; Arndt, J.; Domeier, W.; Berkowitz, S.; Engen, J. R., Characterization of IgG1 Conformation and Conformational Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2009,** *81*, 2644-2651.

29. Wallace, B. D.; Wang, H.; Lane, K. T.; Scott, J. E.; Orans, J.; Koo, J. S.; Venkatesh, M.; Jobin, C.; Yeh, L.-A.; Mani, S.; Redinbo, M. R., Alleviating Cancer Drug Toxicity by Inhibiting a Bacterial Enzyme. *Science* **2010,** *330* (6005), 831-835.

Figures and Tables

Table 5.1. The abbreviated sub-peptide average deuterium difference spreadsheet. Blue boxes are drawn to represent which *NDD* values contribute to each residue's average calculation.

Peptide		Amino Acid ID:	M	L	R	P	V	E	T	P	T	R	E	I	K
Start	End	sequence #:	21	22	23	24	25	26	27	28	29	30	31	32	33
21	31		ND	ND	0.030	ND	0.030	0.030	0.030	ND	0.030	0.030	0.030		
21	38		ND	ND	0.017	ND	0.017	0.017	0.017	ND	0.017	0.017	0.017	0.017	0.017
21	39		ND	ND	0.019	ND	0.019	0.019	0.019	ND	0.019	0.019	0.019	0.019	0.019
21	40	*NDD*	ND	ND	0.018	ND	0.018	0.018	0.018	ND	0.018	0.018	0.018	0.018	0.018
21	41		ND	ND	0.016	ND	0.016	0.016	0.016	ND	0.016	0.016	0.016	0.016	0.016
23	40				ND	ND	0.000	0.000	0.000	ND	0.000	0.000	0.000	0.000	0.000
21	43		ND	ND	0.011	ND	0.011	0.011	0.011	ND	0.011	0.011	0.011	0.011	0.011
32	40													ND	ND
		Average *NDD*	#N/A	#N/A	0.019	#N/A	0.016	0.016	0.016	#N/A	0.016	0.016	0.016	0.014	0.014

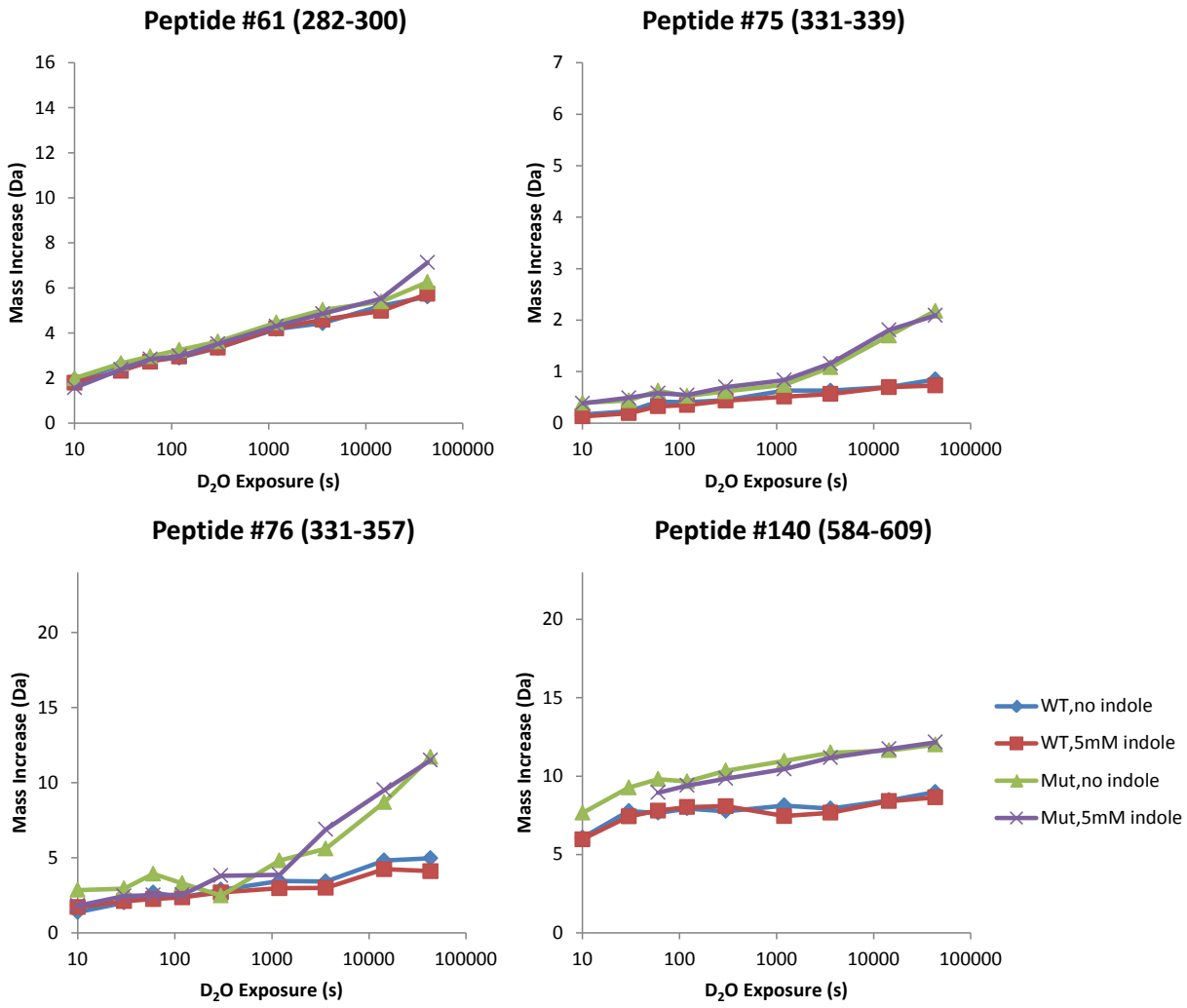


Figure 5.1. Deuterium uptake curves for four selected peptides of β-glucuronidase. Each y-axis is scaled to the maximum number of exchangeable amide hydrogens.

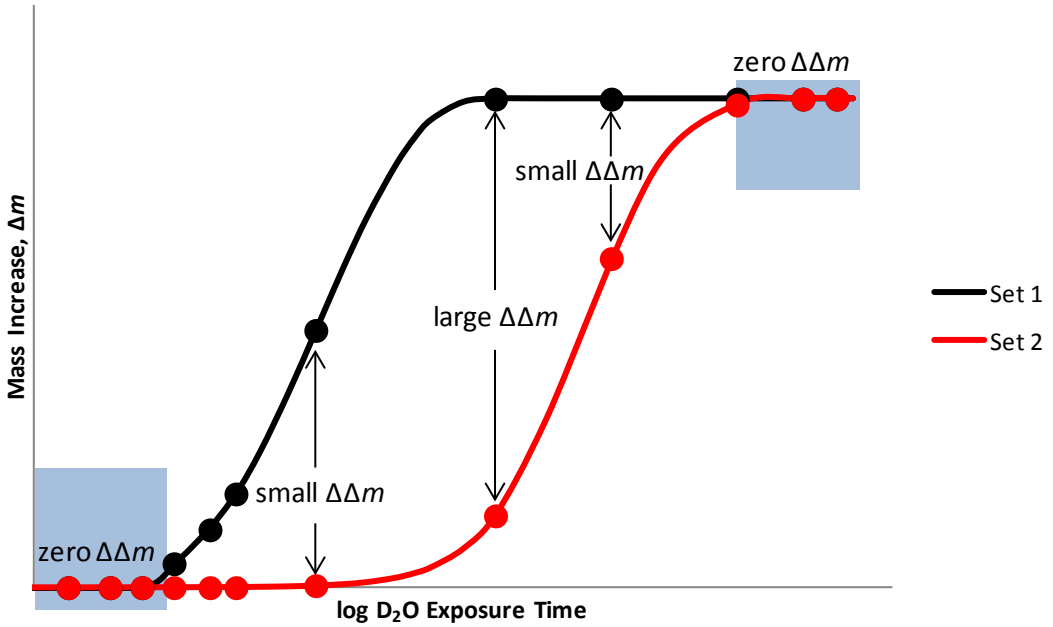


Figure 5.2. Mass increase over time is illustrated by two mock deuterium uptake curves. If ΔΔm is measured at more late or early time points (as indicated by the blue boxes) and all time point ΔΔm values contribute equivalent weight to an average, then those time points where ΔΔm is small or near-zero will skew the average to a small value.

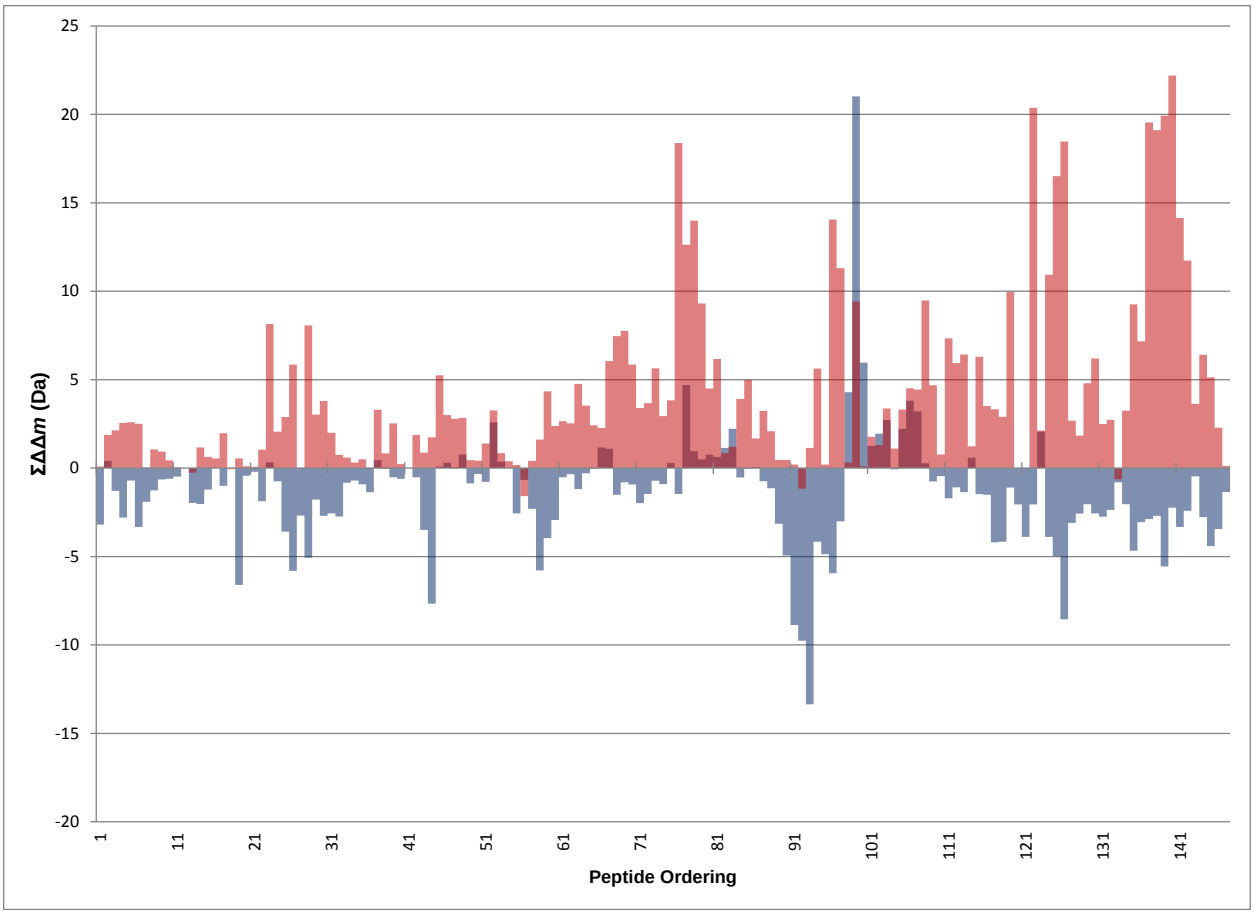


Figure 5.3. The deuterium uptake effects of mutation on wild-type β-glucuronidase (red) and of indole on mutated β-glucuronidase (blue) are presented simultaneously. The magnitude of deuteration difference is measured as the sum of the differences between deuterium curves at every time point, $\Sigma\Delta\Delta m$. Peptides are ordered by their sequence midpoint location in reference to the sequence of the whole protein.

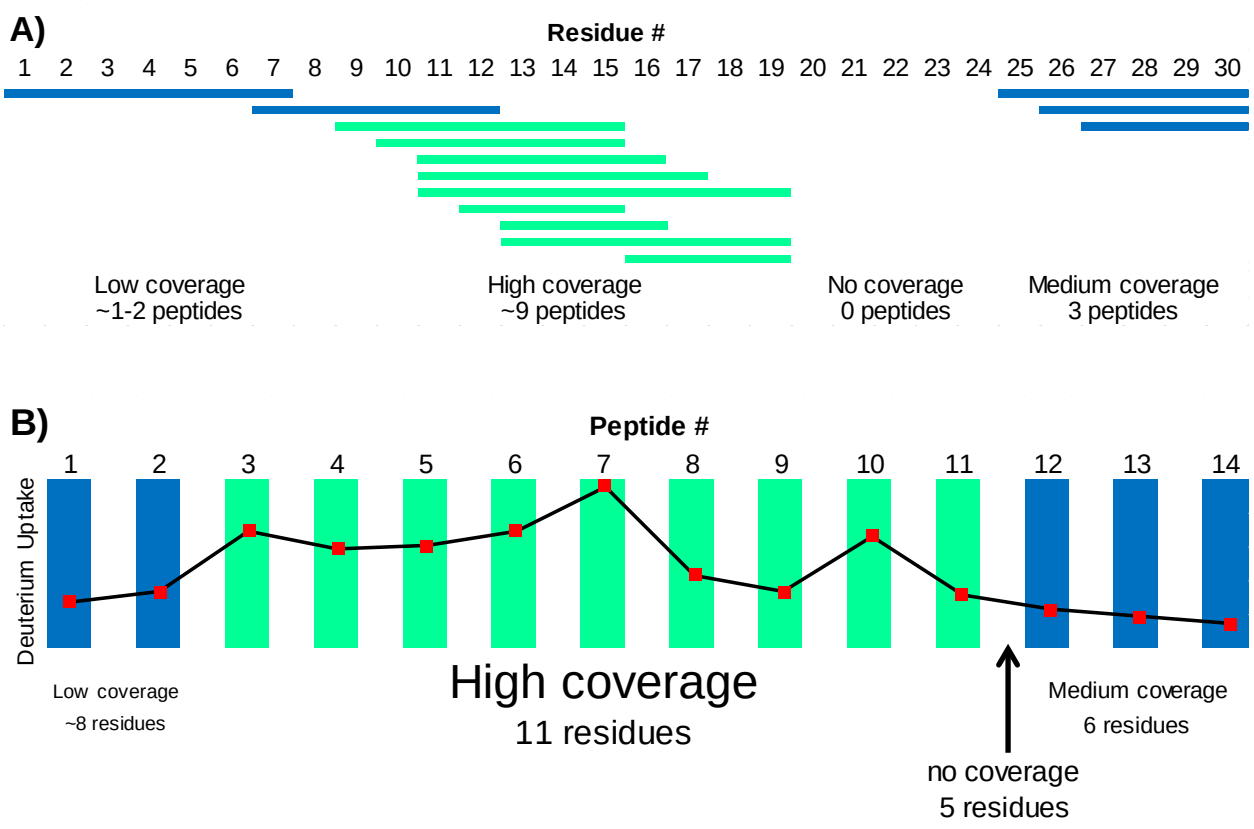
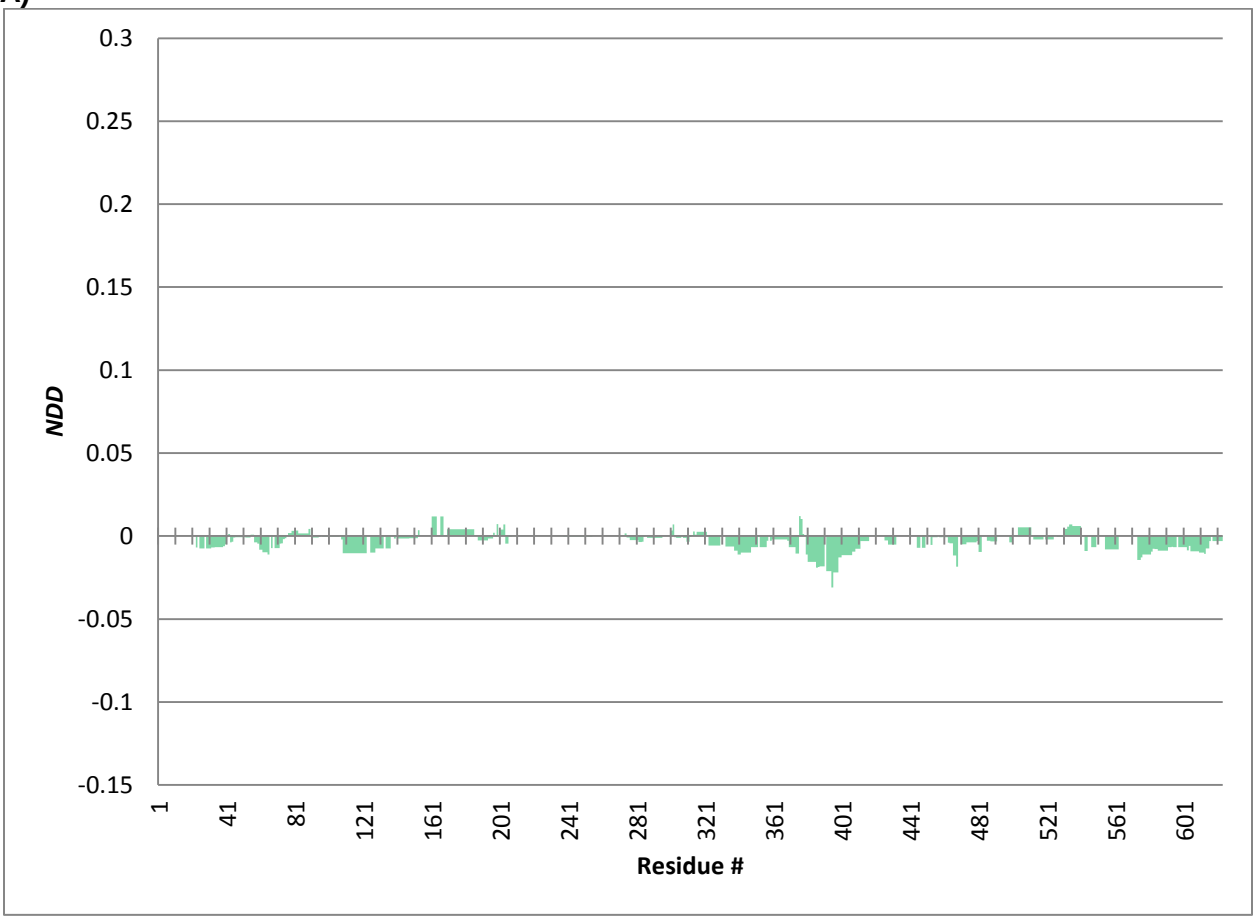


Figure 5.4. A set of mock peptides was used to illustrate a peptide data map. (A) Peptides are mapped along the sequence by their identities. (B) Peptide exchange information is translated to a plot that lists peptides in order of their "quasi-sequence".[25] Regions with no peptide coverage are not represented this plot. The red points overlaid onto (B) are mock exchange values.

A)


0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
NDD
1
41
81
121
161
201
241
281
321
361
401
441
481
521
561
601
Residue #

B)

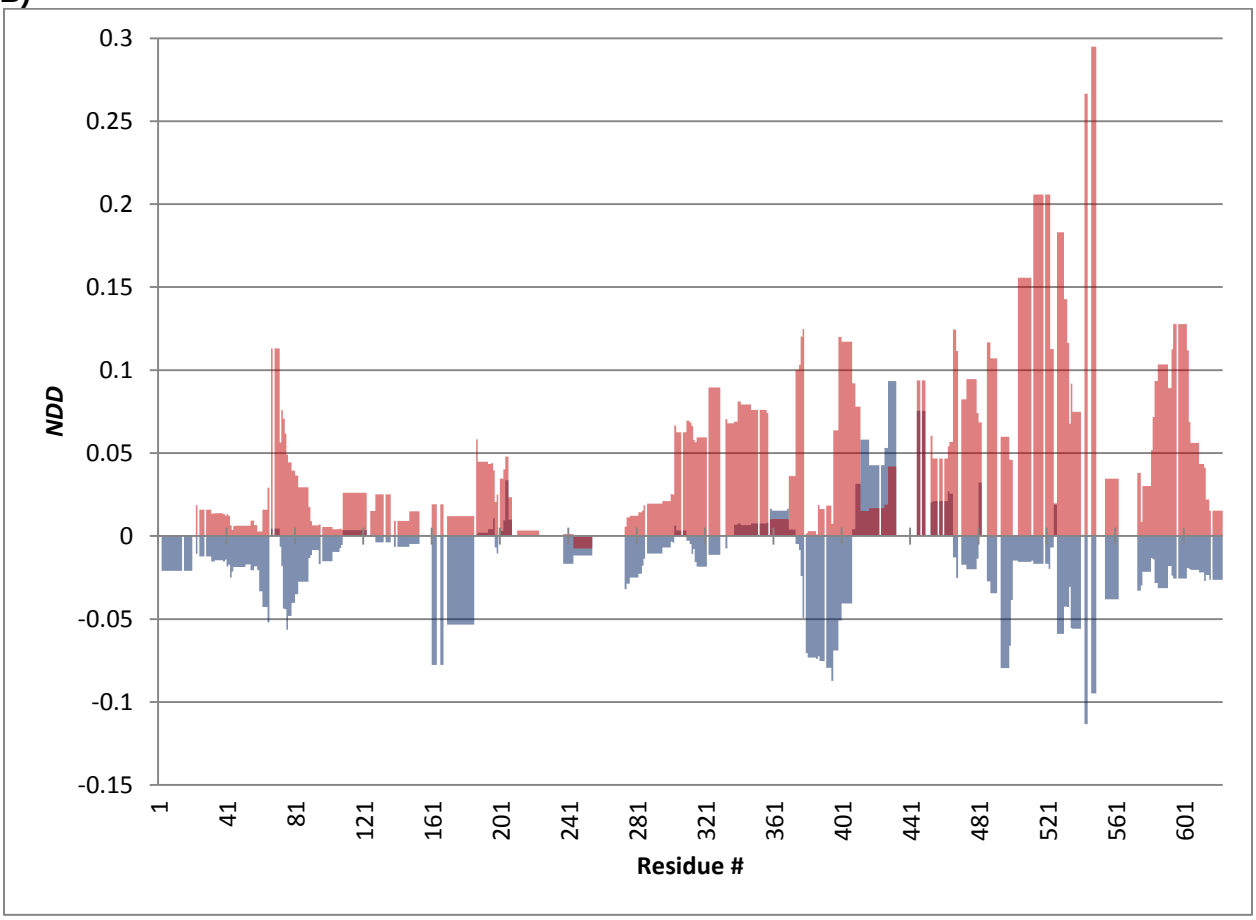


Figure 5.5. A) The effects of indole on wild-type β-glucuronidase (green) deuterium uptake. **B)** The simultaneous representation of mutation effects on wild-type β-glucuronidase (red) and indole effects on mutated β-glucuronidase (blue) on deuterium uptake. The magnitude of deuteration difference is measured as the normalized average difference between deuterium curves at each time point, *NDD*, in both plots.

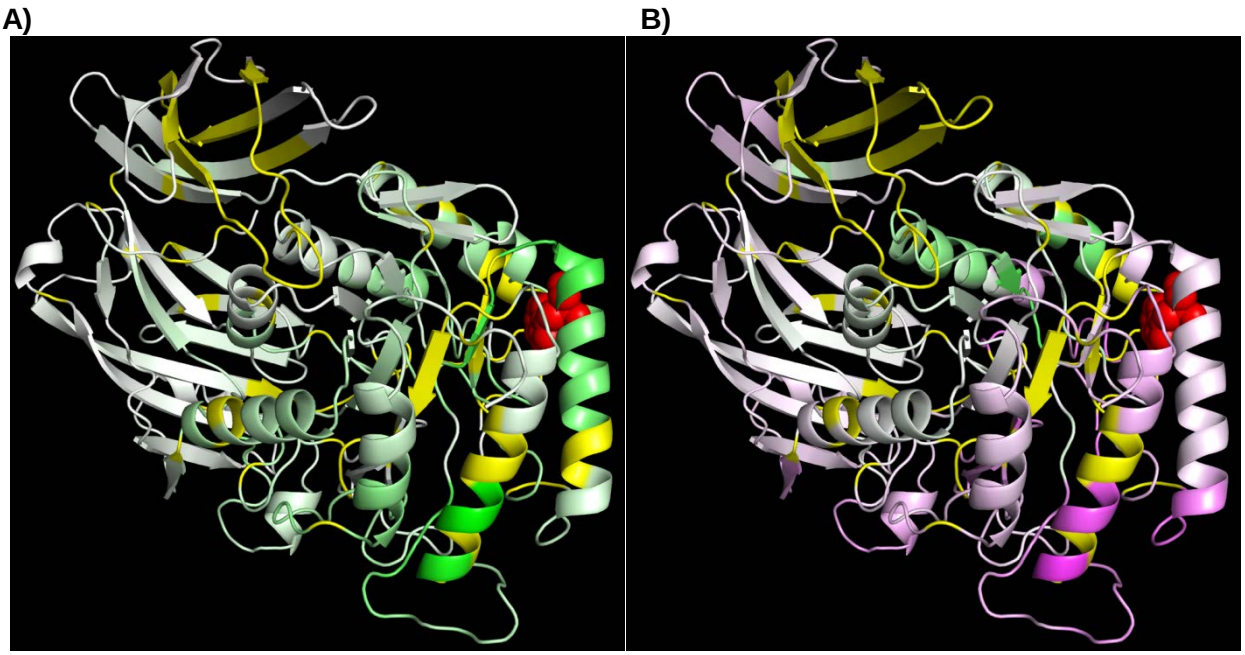


Figure 5.6. An image of β-glucuronidase protein with mapped mutation effects ±indole. The color scale transitions from negative NDD values in magenta, to zero NDD values in white, to positive NDD values in green. Residue 492 is highlighted, using atomic spheres, as a red tryptophan in these models. Residues with no peptide coverage and prolines are colored in yellow. **A)** The W492G mutation decreases protection from H/D exchange across much of the protein, relative to wild-type protein. The effects appear more localized around the mutation site. The color intensity scale covers an NDD range between −0.30 and +0.30. **B)** The addition of indole causes a recovery of protection, which is, again, more pronounced in helices in close proximity to the mutation site. The color intensity scale covers an NDD range between −0.15 and +0.15. This is a PyMOL generated model of β-glucuronidase protein, PDB: 3K4D.[29] The crystal structure used was the asymmetric unit structure.

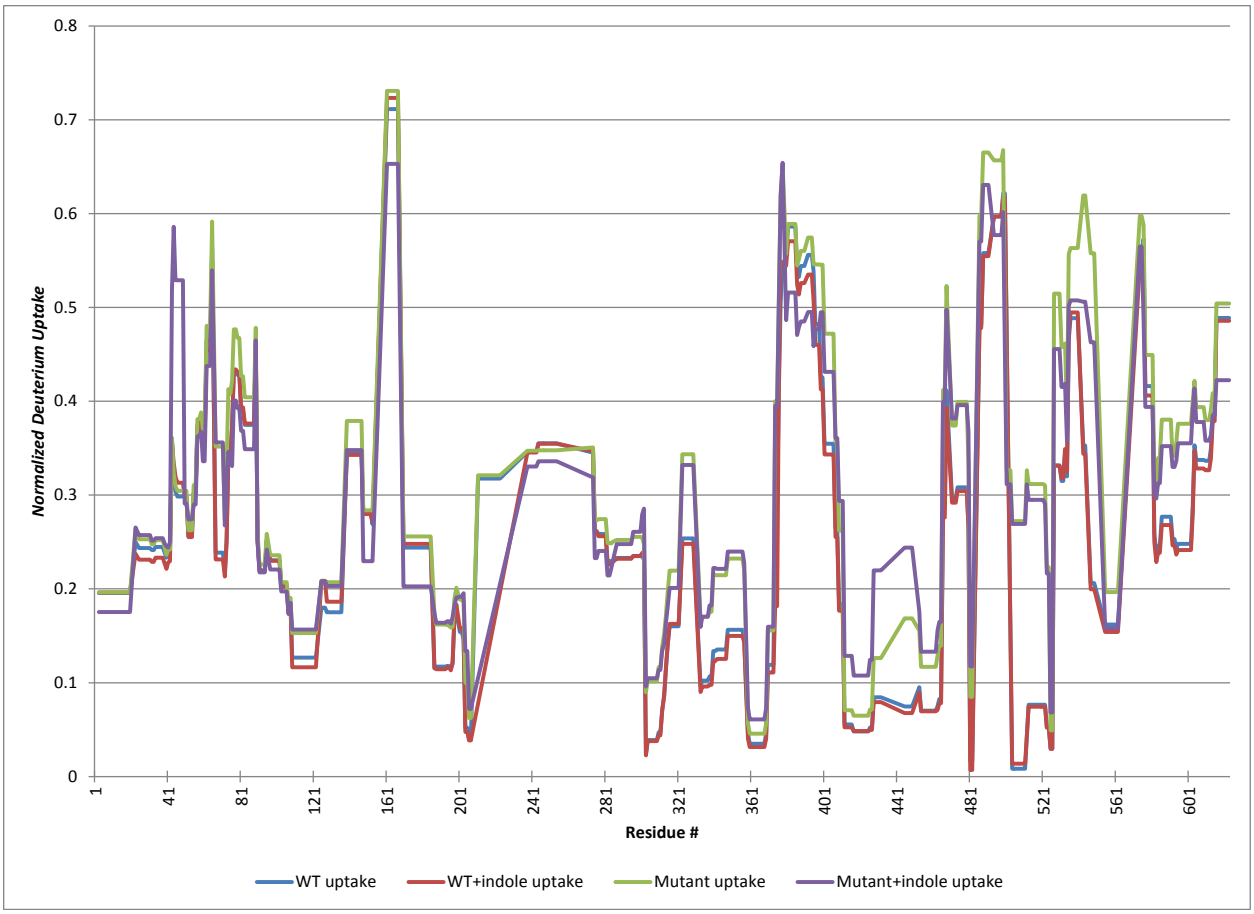


Figure 5.7. A normalized deuterium uptake plot representing protection from exchange on a residue by residue basis. Gaps in the data are not illustrated and the connecting line simply connects the points around them.

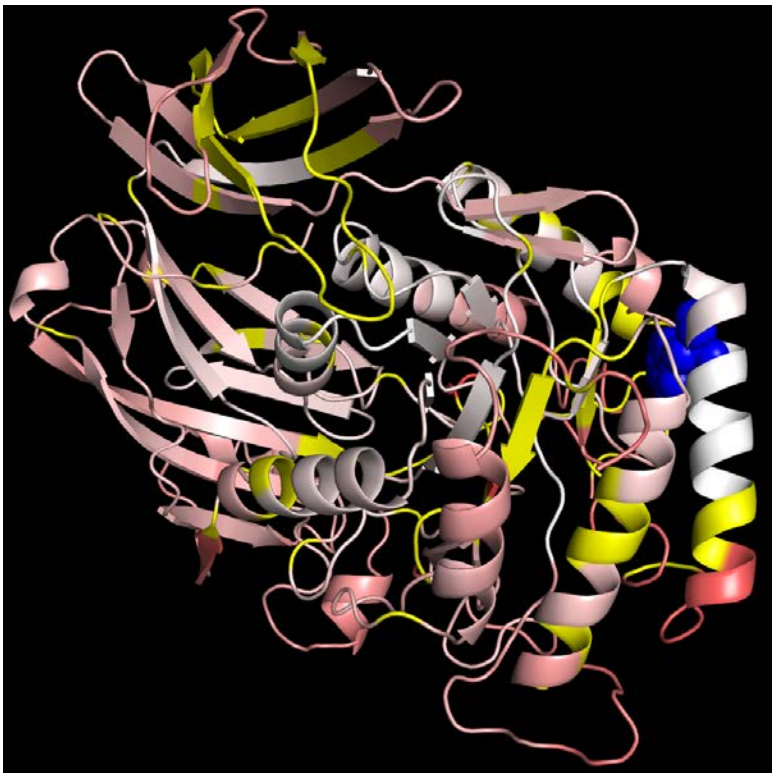


Figure 5.8. An image of β-glucuronidase protein with mapped NDU values for the wild-type protein control sample (no indole). The color scale transitions from 0.0 values in white to 0.8 values in strong red. Residue 492 is highlighted, using atomic spheres, as a blue tryptophan in this model. Residues with no peptide coverage and prolines are colored in yellow. This is a PyMOL generated model of β-glucuronidase protein, PDB: 3K4D.[29] The crystal structure used was the asymmetric unit structure.