Characterization of Structures and Dynamics of Intrinsically Disordered Domain and Proteins Using Time Resolved-Hydrogen Deuterium Exchange Mass Spectrometry

SHAOLONG ZHU

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FUFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CHEMISTRY
YORK UNIVERSITY
TORONTO, ONTARIO

February 2017

Abstract

Proteins are inherently dynamic. Virtually all of the processes that underlie biological activity, including binding partner recognition, catalysis and allostery among many others, require transient adoption of specific high energy conformations. While many proteins contain highly dynamic regions or Intrinsically Disordered Domains (IDDs), Intrinsically Disordered Proteins (IDPs) are free of well-defined structural features altogether. These domains/proteins play integral roles in a wide variety of biological processes, but are often associated with neurodegenerative disease and cancer. Hence it is crucial to study these biomacromolecules to understand the structural underpinnings of their biological functions as well as the factors that drive disease pathology. However, many of the current high resolution structural techniques (like X-ray crystallography and NMR) are not feasible for such studies because of their reliance on the presence of a well-defined 'native' conformation. An emerging structure labeling technique called Time Resolved-Hydrogen Deuterium Exchange (TR-HDX) can offer an alternative strategy. In this approach, structure-dependent hydrogen/deuterium labelling is measured on the millisecond timescale, providing an exquisitely sensitive picture of the weak hydrogen bonding networks that impart 'residual structure' in IDDs and IDPs. In this work, TR-HDX was implemented onto microfluidic chip to achieve solution phase site-specific analysis of the structures and dynamics of p53 N-terminal transactivation domain and tau protein. Gas-phase TR-HDX was also implemented to study gas-phase protein structures that become populated during Differential Mobility Spectrometry (DMS) ion mobility analyses. Critical advances include the discovery and characterization of an amyloidogenic tau intermediate that may underlie Alzheimer's disease pathology, a dynamics-based model for recruitment of co-factors to cancer protein p53 upon phosphorylation, a detailed account of how a set of anti-amyloid drug candidates affect tau structure and dynamics, and development of a method to predict solution-phase protein stability from DMS-HDX measurements.

Acknowledgements

Science is never one person's work but rather a collaborative effort and there are so many great people that I want to thank during this facet of my academic career. Without them, none of this would have been possible. First, I would like to express my heartfelt appreciation and thank my supervisor, Dr. Derek J. Wilson from the bottom of my heart for your guidance and advices. You took me in when I was a desperate undergraduate thesis student. Not only you were a great mentor, you also gave us the freedom to explore our own scientific curiosities. You made us think critically, enabled us to grow as great researchers and solidify our passion for science. I want to thank Drs. Gerald F. Audette and Dasantila Golemi Kotra for taking their time to be my committee members all these years and offering thought provoking discussions, valuable criticisms as well as life lessons. Thank you, Gerald, for providing me a space in your lab so that I could conduct most of my experiments, you certainly made me feel like a part of the Audette lab! I also want to thank Drs. Lars Konermann and Georg Zoidl for being my external committee members and sharing your valuable suggestions and enthusiasm for science!

Drs. Tamanna Rob, Yanfang Liang and Peter Liuni, you guys were the first generation in the Wilson lab and I am so fortunate to have met you all. All of you taught me everything there is to know when I first started. Tamanna, you were the mother of the lab that glued us all together like a family. Yanfang was there to cheer us up with her catch phrase "Don't worry! It's easy!" when experiments failed. Peter, you were like a big brother to me and was always there when I needed help, even if it was past midnight trying to get that microfluidic chip to work (Remember that nitrile glove and staring at the window, mumbling "conformers"?!) I also want to thank Dr. Agnesa Shala, Michelle Ferraro, Anna Petrov, Dr. Queenie Hu and Dayana D'Amora for teaching me the secrets to cloning and protein purifications. Cristina Lento for always staying on top of things, being organized and helping me with edits whether if its manuscript, thesis or administrative paper works. Kerene Brown, for being a ball full of positive energy (please stay that way forever!). Diana Resetca for teaching me great work ethics and never giving up. John Van Nostrand for maintaining our mass specs and sharing your MS knowledge with all of us (and by the way, finish your thesis!). Greg Koyanagi for your technical support when no one was around. Preet Kamal Gill for being a great friend all these year and always helping me around. Shamina Prova, for always being happy and providing insightful suggestions. Dr. Declan Williams, for teaching me the fundamentals of (phospho)proteomics.

Many thanks to my collaborators for helping me publish our work together! Drs. Larry Campbell and Yves Leblanc, for not only teaching but also letting me explore and work with their DMS system at Sciex, it was definitely an awesome experience! Drs. Adnan Sljoka, Alexandr Bezginov, and Jane Heffernan, for simulating thousands of structures for my very first paper. Karen Rethoret for teaching me the use of Transmission Electron Microscopy. Dr. Katherine Donovan, our adopted kiwi scientist, for bringing "heaps" of pyruvate kinase and popping out data like there is no tomorrow. Dr. Yi Sheng and Rahima Khatun for providing the reagents and helping me tremendously with the p53 project. Drs. Mark Reed, Marcia Taylor and especially Bin Deng for helping me with the project in collaboration with Treventis (I truly believe that we found an effective epitope site for tau!). Dr. Difei Sun for insightful discussions as well as optimising our MSMS parameters.

I would also like to thank Nicole McGregor, Magy Baket and Mary Mamais for clarifying and facilitating graduate studies issues. We all know those can be complicated sometimes.

To my "canadian" aunt, Kim Yeon Soon and her family for helping me settle in Canada and always making me feel at home.

Finally, I would like to thank my parents, sister, and brother for their unconditional support and love, especially my mother who sacrificed everything so that I could be where I am today. You fostered my interest in science since I was young and pushed me towards this direction. As promised, this work is dedicated to you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Aβ- Amyloid-β

ACTR- Activator of Thyroid and Retinoid Acid Receptor

AD- Alzheimer's Disease

ADC- Analog- to -Digital Converter

AS- Alternative Splicing

ATM- Ataxia Telangiectasia Mutated

ATP- Adenine TriPhosphate

ATR- Ataxia Telangiectasia and Rad3

CCS- Collisional Cross Section

CD- Circular Dichroism

Cdc2- Cyclin-dependent kinase 2

Cdk5- Cyclin-dependent kinase 5

CH3OD- Deuterated Methanol

Chk- Checkpoint Kinase

CID- Collisional Induced Dissociation

CV- Compensation Voltage

DC- Direct Current

DMS- Differential Mobility Spectrometry

DTIMS- Drift Time Ion Mobility Spectrometry

ELISA- Enzyme-Linked ImmunoSorbent Assay

ESI- ElectroSpray Ionization

FAIMS- Field Asymmetric Ion Mobility Spectrometry

FRET- Förster Resonance Energy Transfer

FRODAN- Framework Rigidity Optimized Dynamics Algorithm New

GSK3β- Glycogen Synthase Kinase-3β

GTFIIH- General Transcription Factor II H

HDX- Hydrogen Deuterium Exchange

HPV- Human PapillomaVirus

IC_{50}- Inhibition Concentration at 50%

IDD- Intrinsically Disordered Domain

IDP- Intrinsically Disordered Protein

IEX- Ion Exchange

IMS- Ion Mobility Spectrometry

ITC-Isothermal Titration Calorimetry

JNK- c-Jun N-terminal Kinase

LC- Liquid Chromatography

m/z- mass to charge

MALDI- Matrix-Assisted Laser Desorption Ionization

MAPK- Mitogen-Activated Protein Kinase

MS- Mass Spectrometry

ND4OD- Deuterated Ammonium Hydroxide

NFT- NeuroFibrillary Tangle

NMR- Nuclear Magnetic Resonance

NPLC- Normal Phase Liquid Chromatography

NTD- N-terminal Transactivation Domain

PH- Pleckstrin Homology

SEC- Size Exclusion Chromatography

rf- radio frequency

RPLC- Reverse Phase Liquid Chromatography

PEG- PolyEthylene Glycol

PF- Protection Factor

PHF- Paired Helical Filament

PMMA- Poly(Methyl Methacrylate)

PTM- Post-Translational Modification

SAXS- Small Angle X-ray Scattering

SV- Separation Voltage

TDC- Time- to- Digital Converter

TEM- Transmission Electron Microscopy

TR-HDX- Time Resolved Hydrogen Deuterium Exchange

TWIMS- Travelling Wave Ion Mobility Spectrometry

VADAR- Volume-Area-Dihedral Angle Reporter

List of Publications

1. Lento, C.; **Zhu, S.**; Brown, K.A.; Knox, R.; Liuni, P.; Wilson, D. J. Time-Resolved ElectroSpray Ionization Hydrogen-Deuterium Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. 2016. *In-Press*.
2. **Zhu, S.**; Campbell, J.L.; Chernuschevich, I.; Le Blanc, J.C.; Wilson, D. J. Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution. *J. Am. Soc. Mass Spectrom.* 2016, 27 (6), 991-999.
3. Donovan, K.A.; **Zhu, S.**; Liuni, P.; Peng, F.; Kessans, S.A.; Wilson, D. J.; Dobson, R.C. Conformational Dynamics and Allostery in Pyruvate Kinase. *J. Biol. Chem.* 2016, 291 (17), 9244-56.
4. **Zhu,S.**; Shala, A.; Bezginov, A.; Sljoka, A.; Audette,G.; Wilson, D. J. Hyperphosphorylation of Intrinsically Disordered Tau Protein Induces an Amyloidogenic Shift in Its Conformational Ensemble. *PLoS One.* 2015, 10(3):e0120416
5. Liuni, P.; **Zhu, S.**; Wilson, D. J. Oxidative Protein Labeling with Analysis by Mass Spectrometry for the Study of Structure, Folding, and Dynamics. *Antioxid. Redox Signal.* 2014, 21 (3), 497–510.
6. 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.

Chapter 1

Introduction

1.1 Mass Spectrometry

Mass spectrometry (MS) has existed for more than a hundred years and has proven to be a powerful analytical technique that is capable of measuring ionised atomic and molecular masses[1,2]. Even at its infancy, there was motivation to advance this nascent technology due to its high degree of selectivity. It was suggested by Wilhelm Wien's discovery that positively charged particles can be deflected using a strong magnetic field, which gave MS its start. However, it was J.J. Thompson who spearheaded this technology when he discovered two different neon isotopes using electric and magnetic fields[1,2]. Ionization is a prerequisite for transferring analytes into the gas-phase for accurate mass determination. Since the advent of soft ionization techniques, ElectroSpray Ionization (ESI) and Matrix-Assisted Laser Desorption Ionization (MALDI), the analytical field in MS has developed rapidly and paved the way to study large biomolecules such as proteins, and large macromolecular protein complexes such as viruses[3–5].

1.1.1 ElectroSpray Ionization (ESI)

ESI, one of the two soft ionization techniques, is widely adopted in the field of life sciences. Although John Fenn won and shared the Nobel Prize for his contribution in the ionization technique, it was Malcolm Dole and coworkers who first discovered the process in the 1960s. Generally, the ESI process is initiated by applying a high voltage (2-6kV) to the tip of a thin metal stainless steel capillary containing flowing analyte solution at low flow rates (Figure 1.1)[6,7]. There are two modes of operation, positive and negative. For example, in positive ion mode, an electrochemical reaction of the solution (water oxidation) occurs causing electrons to flow to the outer walls of the capillary and the accumulation of positively charged ions in solution. The positive potential generated will repel and drive positively charged analytes towards the end of the capillary forming a conical structure called Taylor cone[8]. The strong electric field will disperse the liquid solution in the Taylor cone into smaller highly charged parent droplet aerosols due to strong coulombic repulsion[6,7]. As the droplets migrate towards the mass spectrometer, this repulsion force passes the Rayleigh limit which will eventually generate progeny droplets resulting from jet fission[7,9]. This process occurs numerous times during the transit period with the analytes usually containing numerous charges. Concomitantly, dry nitrogen sheath gas is flown coaxially around the capillary to aid in nebulisation. This helps with the desolvation process and reduces the degree of cluster ion formation[7].

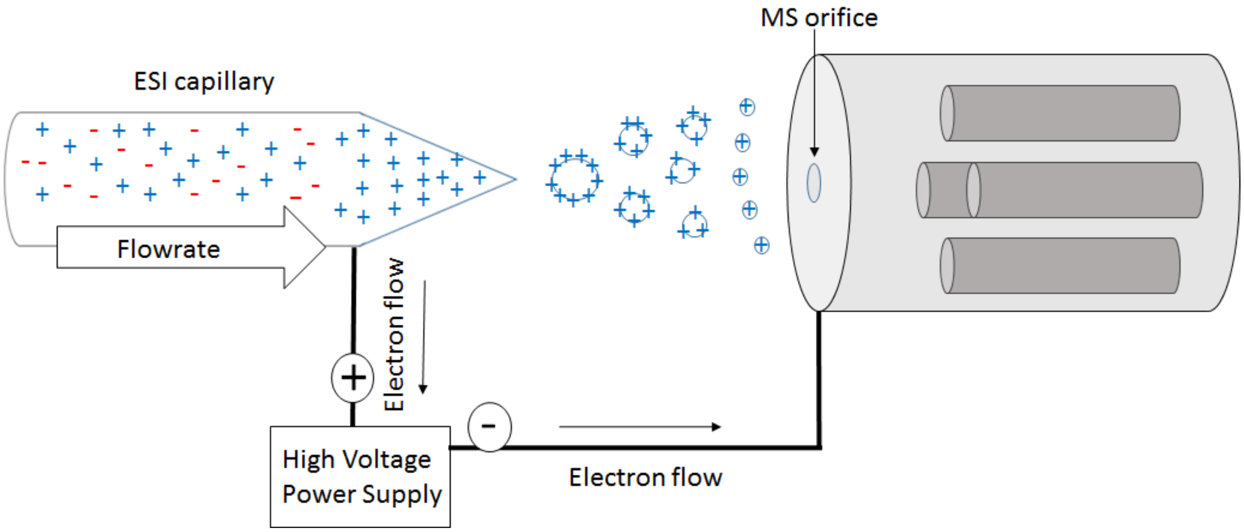


Figure 1.1. A representation of the electrospray ionization process in positive ion mode. High electric field applied to the ESI metal capillary containing analyte solution causes the formation of the Taylor cone. This results in the ejection of the positively charged parent ions due to overpowering of the columbic force while the negatively charged ions align at the walls of the capillary. Subsequently, these parent ions fly towards the negatively charged curtain plate of the mass spectrometer. During the transit, there is an interplay between solvent evaporation and the columbic repulsion force. When the columbic forces are greater than the surface tension of the droplets, parent ions undergo a jet fission process giving rise to smaller progeny droplets. This jet fission process occurs numerous times and at an increasingly faster rate due to smaller droplet size until the charged analytes are transferred into the mass spectrometer.

1.1.2 Quadrupole Time-of-Flight (QToF) Mass Spectrometers

QToF mass spectrometers are powerful robust instruments that are widely utilised in the analytical community mainly due to its high sensitivity and high mass accuracy[10]. Conventional QToF instruments usually consist of quadrupole 0 (Q0) for focusing of ions as they enter the mass spectrometer, Q1 for mass selection and transmission, Q2 for parent ion fragmentation in the collision cell, ToF for mass separation, and a detector component (Figure 1.2)[10]. Of course, the above describes one of the earliest designs as well as current conventional models and currently there many variants with additional features. For example, Waters G1 and G2S models use an ion guide and a stepwave ion guide, respectively, in place of Q0 and also contains an additional separation dimension called the ion mobility cell[11,12].

Briefly, Q0 and Q2 operate under radio frequency (rf) mode and is normally maintained in the range of 10-50 mTorr[10]. This rf field provides radial confinement and collisional cooling resulting in better focusing of the ions. Ions are focused in Q0 as a concentrated beam as it enters the mass spectrometer to the Q1 chamber[10]. Q1 acts as a mass filter and is operated under both rf and dc voltages to allow certain mass to charge (*m/z*) transmission window[10]. Q1 is normally maintained at a low pressure of 10^{-5} torr to avoid ion diffusion and fragmentation of the parent ions[10]. If MS/MS or fragmentation mode is operated following the transmission, the ions are accelerated into the pressurised Q2 chamber where collisional induced dissociation (CID) occurs with the bath gas (typically N_2), giving rise to product ions[10]. Upon exiting the final quadrupole, the ions (product ions if MS/MS is operated) are then reaccelerated and focused by ion optics into the ToF analyser (Figure 1.2)[10]. Here, the ion modulator pulses the incoming ions (several kilohertz) orthogonal to their trajectory where mass separation occurs[10]. For detection, variants of electron multipliers are used such as microchannel plates to amplify the signal[10]. Classically, mass spectra are recorded with time-to-digital converters (TDC), however, with advancing technology, the current systems now commonly use analog-to-digital converters (ADC).

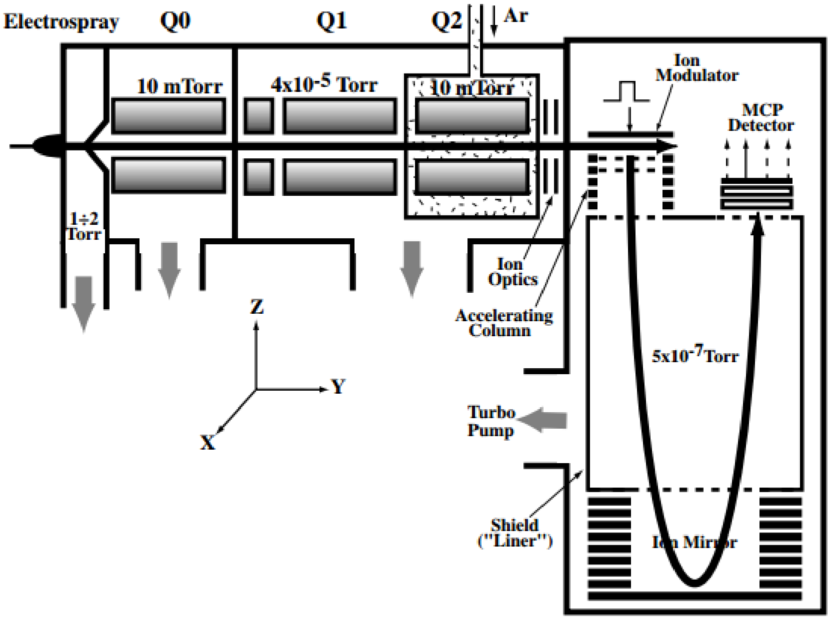


Figure 1.2. Schematic representation of a QToF mass spectrometer. Electrosprayed gas-phase ions are cooled and focused at Q0 quadrupole upon entering the orifice of the mass spectrometer and are *m/z* selected at Q1 using rf field. Q2 is operated at higher pressure containing neutral particles such as nitrogen and provides a collision chamber for CID if desired. Ions exiting Q2 are then refocused into a beam by ion optics entering the ToF analyser. The ion modulator then pulses ions orthogonal to their original trajectory into a high vacuum space where mass separation occurs. Ions are deflected at the ion mirror for enhanced separation which are then detected using ToF detector such as multichannel plate detectors as shown. Adapted from Chernushevich *et al*[10].

1.2 Ion Mobility Spectrometry- Brief Introduction

Ion Mobility Spectrometry (IMS) is a powerful analytical technique that can separate gas-phase ions based on their size and shape[13]. Thus far, three main IMS techniques exist: drift time IMS (DTIMS), travelling wave IMS (TWIMS), and field asymmetric IMS (FAIMS) or differential mobility spectrometry (DMS)[13]. IMS coupled to MS can significantly improve selectivity, speed, and limit of detection[13]. The additional separation dimension along with liquid chromatography-MS (LC-MS) or MS alone can improve the selectivity process especially in isolating isobaric species and different ion conformations[14,15]. IMS can also perform high throughput analysis by separating ions on the millisecond timescale as opposed to seconds timescale in conventional chromatography, enhancing the speed of analysis[13]. A lower limit of detection can be achieved by filtering out background noise during the separation process[16]. The next two paragraphs will briefly introduce DTIMS. TWIMS and DMS will be discussed in the following sections.

DTIMS represents the first IMS technique, where low uniform electric fields of 5-100V are applied to propel ions in a drift tube filled with drift gas, typically He (Figure 1.3)[13,17]. The time for an ion to traverse through the tube can be correlated with the ion's rotational collisional cross section (CCS) where ions with smaller CCS will travel faster than that of bigger CCS due to fewer collisions or interactions with the drift gas[13,17]. The reduced mobility coefficient, k_0 can be calculated using the following equation[18]:

$$k_0 = \left(\frac{L^2}{t_D V}\right)\left(\frac{273.15}{T}\right)\left(\frac{p}{760}\right) \qquad (1)$$

where L refers to the drift tube length, t_D is the drift time, V is the applied voltage across the drift cell, T corresponds to temperature in kelvin, and p is the gas pressure in torr. Upon finding k_0, CCS can then be determined using the Mason-Schamp equation[18]:

$$\Omega = \frac{3ze}{16N}\left(\frac{2\pi}{\mu k_B T}\right)^{1/2}\left(\frac{1}{k_0}\right) \qquad (2)$$

where z is the charge of the ion, e is the elementary charge, μ is the reduced mass of the ion-neutral pair, and k_B is the Boltzmann constant.

DTIMS offers several advantages over the other two IMS techniques such as high resolving power (R>100), however, due to its gated feature (this also applies to TWIMS), only a limited percentage of the total ions are detected[13]. Although DTIMS is the earliest version of IMS, it is still widely utilized for accurate gas-phase CCS determination of macromolecules such as peptides and proteins and correlates well compared to other structural biophysical techniques such as X-ray crystallography and Nuclear Magnetic Resonance (NMR)[13].

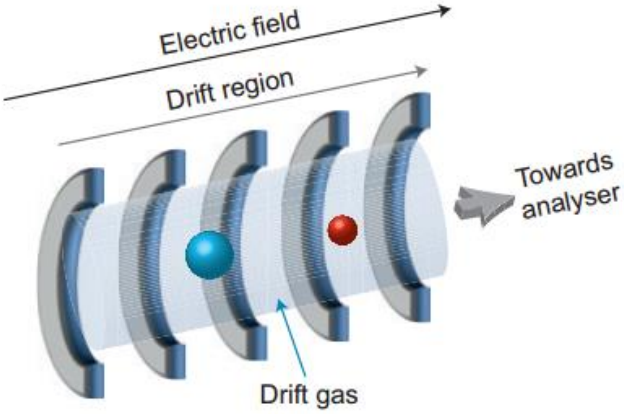


Figure 1.3. A schematic representation of DTIMS. Low electric field (5-100V) is applied along the same direction as the ion path and drift gas is flown in the opposition direction. Ions are then separated by interaction with the drift gas. Hence ions with bigger CCS are slowed down compared to smaller CCS due to more frequent interactions with the drift gas. Adapted from Lanucara *et al*[13].

1.2.1 IMS- Travelling-Wave Ion Mobility Spectrometry

Similar to DTIMS, TWIMS also operates at low fields in the presence of neutral gas such as helium or nitrogen, however there are a few differences in operation as shown in Figure 1.4. TWIMS is comprised of a series of stacked ring ion guides where alternating phases of rf voltage are applied to adjacent rings such that ions are radially confined[12]. In addition, a direct current (DC) voltage is superimposed to this rf voltage, creating a “wave” such that ions can surf axially on this wave. As a result, ions with higher mobility can surf over the wave effectively reducing their transit time whereas ions with lower mobility cannot due to “rollover” effect[12]. TWIMS has also been useful in CCS determination (requires prior calibration with known CCSs) and is being increasingly adopted in the field of structural biology. Robinson and colleagues were one of the first groups to study large protein complexes using this method[19]. For example, they transferred a large protein complex of 90kDa called trp RNA binding attenuation protein (TRAP) into the gas-phase and demonstrated that its structural topology remained intact. Furthermore, its determined CCS was in close agreement with data generated from X-ray crystallography. Later on, Heck and colleagues pushed the boundary and studied intact hepatitis B virus capsids (4MDa)[20].

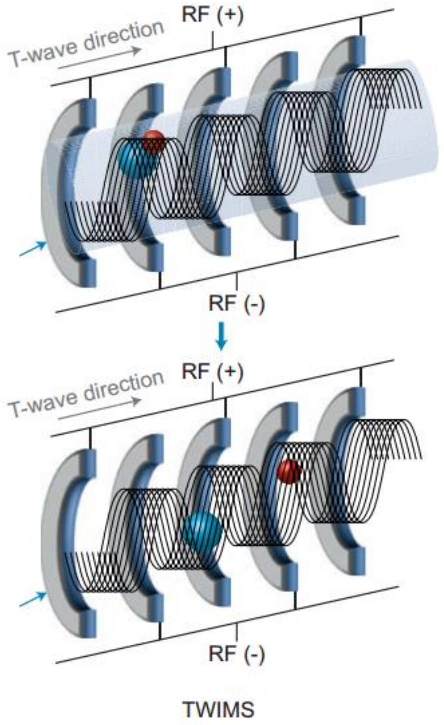


Figure 1.4. A schematic representation of TWIMS. Application of alternating rf voltage along with superimposed DC voltage to a series of stacked ring ion guides results in travelling wave voltage allowing for ions to surf and separate on these travelling waves. Ions with higher mobility are able to surf over the wave whereas ions with lower mobility tend to roll back, thus increasing the transit time in the tube. Ion separation can be optimized by adjusting the wave height and speed. Adapted from Lanucara *et al*[13].

1.2.2 IMS- Differential Mobility Spectrometry

DMS and its analogue, FAIMS, is an emerging bioanalytical technique that can act as an orthogonal separation dimension to existing LC-MS or MS for enhanced selectivity and sensitivity[21]. Originally developed to detect illegal and explosive substances, it is now becoming a common tool in structural biology for analysis of proteins' three-dimensional structure and its conformational dynamics, at least for small proteins (10kDa>)[13,22–25]. Unlike conventional IMS, where ion separation is dependent on the drift time through a neutral gas and under low electric field strength, in DMS, ions are 'pre-selected' for transfer into the mass spectrometer based on the difference between their high-field (~30,000 V cm^{-1}) and low-field (~1,000 V cm^{-1}) mobilities, ΔK [21,26]. To do this, an asymmetric rf voltage (Separation Voltage, SV) is applied perpendicular to the flow of ions, resulting in a net lateral displacement and an 'unstable' trajectory which will eventually lead to and discharge at one of the electrodes. Another sweeping DC offset potential (Compensation Voltage, CV or CoV) is applied to successfully transmit ions of different ΔK (Figure 1.5)[21,26]. The mobility of an ion in these electric fields can therefore be described by the following equation:

$$k_h(E) = k_o[1 + f(E)] \tag{3}$$

where k_h, k_o is the mobility in high and low field respectively and $f(E)$ is the ion mobility dependent on electric field strength. In DMS, there are three types of ion behaviours: type A, type B, and type C. Type A ions exhibit an increase in mobility with increasing fields (clustering effect). Type B ions exhibit an initial increase followed by a decrease in mobility as a function of electric field (polarizing effect). Type C ions exhibit a decrease in mobility as a function of increasing field strength (rigid sphere model).

A typical DMS setup consists of two parallel plates where the high/low voltages are applied with two gas inlets: a transport gas (typically N_2 and other modifier gases) at the entrance of DMS for separation enhancement, and a resolving gas at the interface between the exit of DMS and MS entrance for resolution enhancement[21,26]. One attractive characteristic of this approach is that the separation is principally dependent on the parameter, ΔK, hence interactions with different types of transport gases within the DMS instrument can have a substantial impact on ΔK[27,28]. This can be advantageous where optimization of the transport gas mix (i.e., the modifier gas) can allow for high-resolution separation of isobaric analytes and even of species with identical CCS areas[21,27]. For larger proteins (>15kDa) with significant molecular dipoles, a particular problem is field-induced alignment which results in bimodal CV profiles for unimodal structural distributions[29]. As a result, there has been little effort to apply DMS to large biomolecules and no success in linking gas-phase structural insights acquired by DMS to the solution phase conformational ensembles of proteins[29,30].

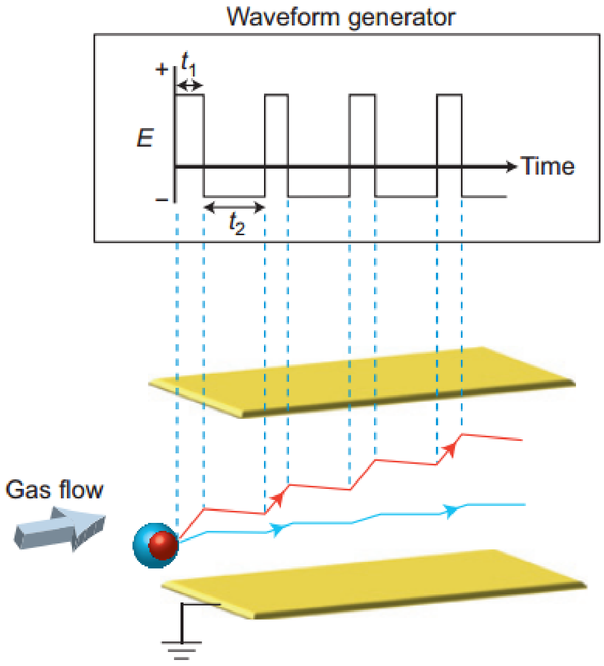


Figure 1.5. Schematic diagram of two ions travelling in a planar DMS cell. Top plot illustrates the field asymmetric wave pulse sequence. Briefly, the lower plate is held at a ground potential whereas asymmetric waveform is applied by the upper plate and net sum of the high and low field voltage is 0 as shown by the plot. This results in a short, high field pulse followed by a longer low field portion resulting in a net trajectory of ions towards the mass spectrometer. Ions are then separated by their ΔK (mobility difference in high- vs. low- field). The red ion's high field mobility increases compared to its low field mobility hence causing it to be more sensitive at high fields compared to the blue ion. Adapted from Lanucara *et al*[13].

1.2.2.1 IMS- DMS, Modifier Gas Effects

As mentioned above, adding different modifier gases in the transport gas line can have beneficial effects in separation and resolution[31]. Inert and/or polar modifiers can be used in the transport gas[31,32]. When inert gas such as helium or nitrogen is implemented, ion and modifier gas interaction is predominantly dictated by the rigid sphere model[32]. However, when polar modifiers such as acetonitrile, methanol, isopropanol, acetone etc. are used, the interaction of modifier gases with analytes in the DMS chamber can be explained using the dynamic cluster/decluster model[31,32]. The model suggests that during the transit of ions in low rf field, modifier gas molecules can form clusters with the analytes due to low energetic conditions and these complexes are then declustered during the high rf field transit[32]. Hence over a series of these high/low field cycles, the analytes and modifier gases are undergoing dynamic clustering and declustering, affecting ΔK[31,32]. All polar modifiers have the potential to induce this mechanism with different efficiencies. Hence with proper optimization, this technique can provide a powerful orthogonal separation dimension to the existing MS platform. For example, Figure 1.6 shows the DMS ionogram of 140 different small molecule compounds. SV was constant and the CV was ramped from -100V to 20V[32]. Without any modifier gas, most peaks were clustered in the narrow CV window (~20V range) and were not separated. However, upon addition of 1.5% isopropanol as a modifier gas, the separation of the compounds was improved drastically encompassing a CV range of ~100V[32].

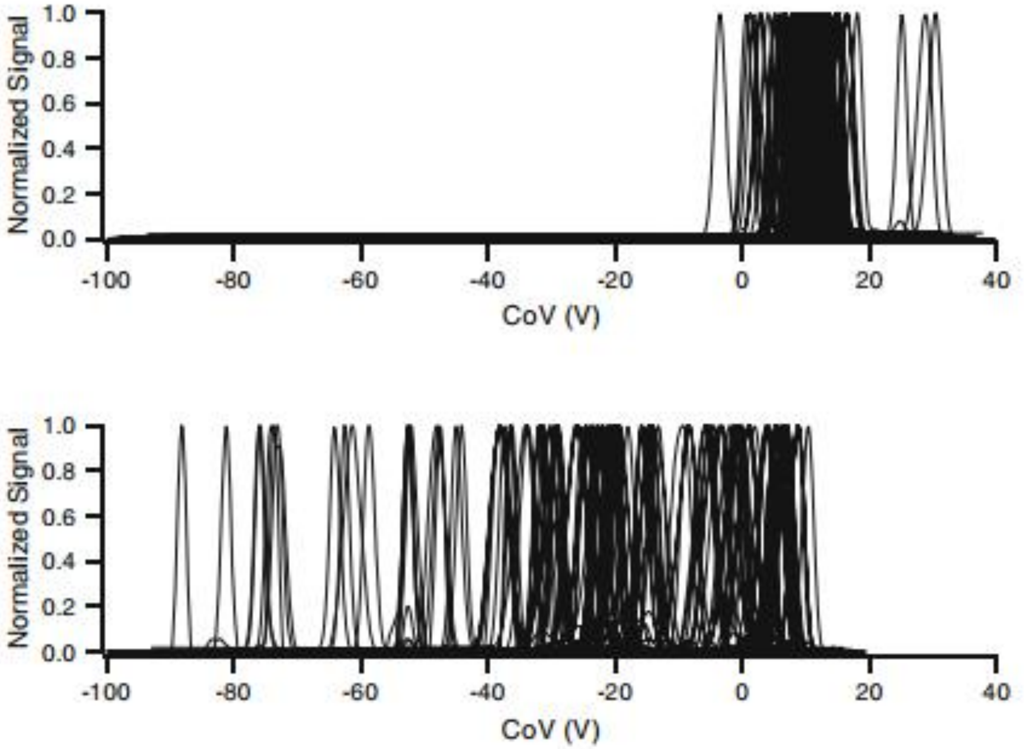


Figure 1.6. Effect of DMS separation upon addition of modifier gas. Top/bottom spectra represents DMS ionogram of 140 compounds without/with any modifier gas respectively. In the absence of any modifier gas, the separation capabilities were poor with the majority of the peaks clustered in a narrow Compensation Voltage (CV or CoV) window. Upon addition of modifier gas, 1.5% isopropanol, the separation was drastically improved spanning over a 3-fold increase in the CV region. Adapted from Schneider *et al*[32].

1.2.2.2 IMS- Differential Mobility Spectrometry Hydrogen Deuterium Exchange (DMS-HDX).

Gas-phase HDX is a structure-dependent labeling technique in which hydroxyl, thiol, amine and amide hydrogens undergo exchange with deuterons from the surrounding gas[28]. Although, the reaction mechanism in the gas-phase is not as well understood as solution phase

HDX, it is still viewed as a sensitive probe of gas-phase structure that can be reflected to solution phase structures[33]. Gas-phase HDX in conjunction with a variant of the DMS setup, FAIMS, has been performed previously to provide a simultaneous, orthogonal characterization of protein structure; however the resulting profiles were difficult to interpret in the context of solution structure[24,29]. For instance, adjacent charge states from the same solution-structure subpopulations exhibited wildly different DMS-HDX profiles[22,34]. Moreover, the profiles displayed a multitude of co-populated conformations with substantially different CVs which suggests that much of the conformational heterogeneity was generated during DMS separation[22,24,34]. Thus, DMS-HDX was demonstrated as a useful tool for investigating different gas-phase conformers, but also shown to have poor correlation to solution phase structure.

1.3 Liquid Chromatography (LC)

LC is an important analytical separation technique and still present as a major workhorse in many analytical applications such as pharmaceutical, environmental, and clinical settings[35]. It is a quantitative and qualitative technique that is commonly used to separate complex mixtures for purification, identification, and quantification[35]. Typically, LC involves a column with packing material referred to as the stationary phase, a pump to move pressurised mobile phase(s) through the column and a detector that displays the time at which analytes elute called retention time[35,36]. Retention times vary depending on the nature of the interaction of the analyte and the column (chemical or physical interaction) as well as the solvents used[35]. Normally, there are five types of LC: normal phase, reverse phase, size exclusion, ion exchange, and bioaffinity chromography.

Normal phase LC (NPLC) uses a polar stationary phase (amino functional group moieties) and non-polar solvents (hexane) as the mobile phase[37]. Hence polar analytes are retained by the polar stationary phase and the adsorption interaction increases with more polar analytes[37]. NPLC is useful in separating hydrophobic molecules such as polysaccharides[38].

Reverse-phase LC (RPLC), unlike NPLC, uses a non-polar stationary phase (C_{20}, C_{18}, C_4) and a polar mobile phase (acetonitrile, water, methanol) which operates via hydrophobic interactions between the analytes and the stationary phase[39]. A typical application of RPLC is for the separation of complex peptide mixtures from digested cell lysates[40].

Size exclusion LC (SEC) also known as gel permeation or gel filtration chromatography separates analytes based on size[41]. This is useful for the determination of tertiary or quaternary structures of proteins or protein complexes[41]. In addition, molecular weights can sometimes be determined as well[41].

Ion exchange LC (IEX) has two variants, anion and cation exchange, and both separate analytes based on attraction to charged sites on the stationary phase[42]. Ions of the same charge as the stationary phase are eluted with little or no interaction whereas oppositely charged ions are retained based on their ionic strength[42]. The mobile phase is composed mainly of increasing concentrations of counter-ions (sodium chloride) that competes with the binding of analytes[42]. Furthermore, pH can be varied in the mobile phase such that the charge of the analyte of interest is changed which subsequently affects its binding to the stationary phase[42]. Typical applications of IEX are water purification and protein purification[42,43].

Bio-affinity chromatography separates based on the specific interaction of proteins with their respective ligands[44]. These ligands are covalently linked to the packed beads in the stationary phase. Elution can be performed in two ways: biospecific elution which involves competing ligand in the mobile phase and aspecific elution which involves changes in pH and salt concentrations that can weaken the interaction of the protein with the affinity column[44]. As the name implies, this method is specific to certain analytes and the resulting products are highly pure.

For detection, typically UV spectroscopy is used where the wavelength is set to a value at which the analytes can absorb[35]. However, LC combined with MS as a detection system is much better because of higher sensitivity and the added selectivity feature provided by MS[45]. LC-MS has many applications, however, it is most renowned in the "omics" fields such as proteomics, metabolomics, lipidomics, and glycoproteomics[45–49].

1.4 Hydrogen Deuterium Exchange (HDX)

HDX is a chemical reaction whereby labile hydrogen atoms undergo exchange with deuterium. It serves as a powerful analytical technique to probe protein conformation and dynamics by monitoring the isotopic exchange of protein backbone amides[50]. HDX was first observed by Kaj Ulrik Linderstrøm-Lang using ultracentrifugation and Englander *et al.* further explored and measured HDX to single amide resolution using nuclear magnetic resonance (NMR)[51]. Despite its high resolving powers, NMR has its limitations such as difficulties in studying transient species, long acquisition times due to complex pulse sequences ultimately resulting in data that represents protein in an equilibrium ensemble, and requirement for highly concentrated samples. Nevertheless, HDX NMR is still widely used in studies of protein-ligand binding and protein folding/unfolding. HDX was first implemented on LC-MS by Zhongqi

Zhang and David L. Smith in 1993 and provided the foundation for the majority of HDX MS workflows still widely used today[52]. Figure 1.7 shows the conventional “bottom-up” HDX workflow using cytochrome c. Briefly, the protein was reacted in excess deuterium solution. The reaction was quenched and the labelled protein was subsequently digested using pepsin in cooled acidic conditions (pH 2.5, 0°C). The labelled peptides were further separated by LC and detected by mass spectrometer[52–54]. HDX MS, unlike HDX NMR, has no limitations on protein size and is highly sensitive requiring protein concentrations down to the nM range.

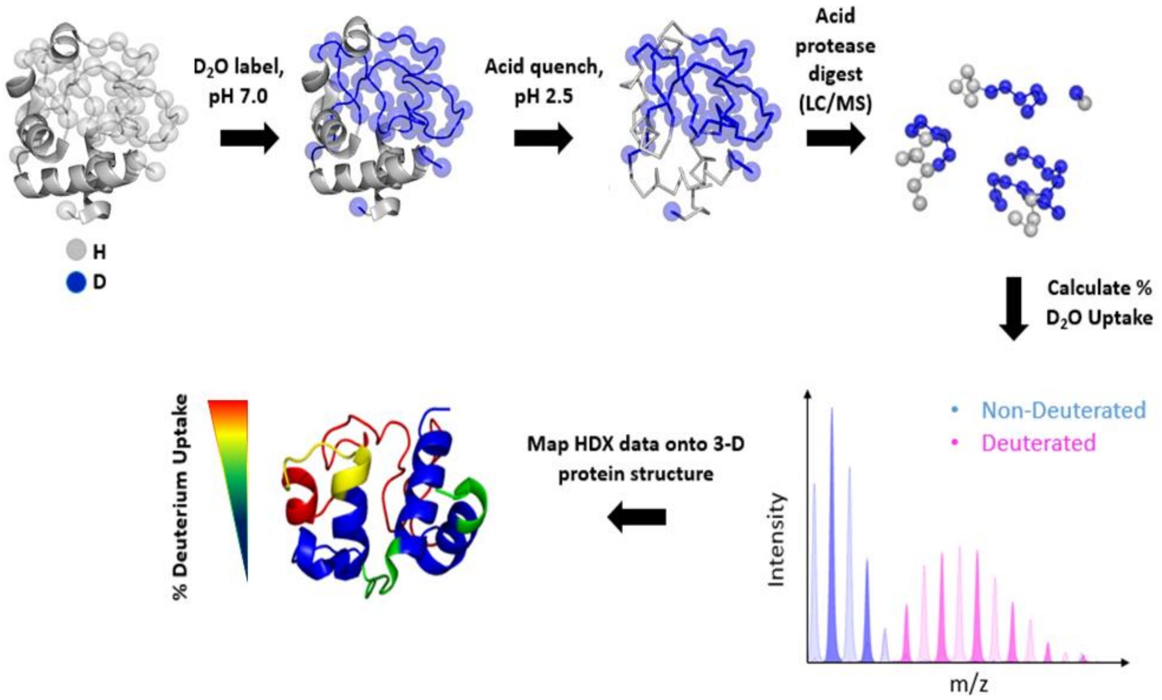


Figure 1.7. Conventional “bottom-up” HDX workflow. HDX is initiated when the protein (cytocrhrome c) is allowed to mix with deuterated buffer for a set time. During this mixing time, solvent exposed amides will exchange much faster than protected regions. Subsequently, the reaction is quenched holding the label in place and digested in acidic conditions. Peptides are then separated and resolved using LC-MS yielding spatial resolution HDX data which are then mapped onto 3-D protein structures. Global HDX experiments work in a similar fashion, however without the quenching and digesting steps.

1.4.1 HDX- Fundamentals

Labile hydrogen atoms bound to sulfur, oxygen, and nitrogen undergo rapid isotopic exchange due to the electronegative heteroatoms[55]. Exchange of backbone amide hydrogens, however, is highly dependent on pH and temperature and can be quenched under the right conditions (pH 2.5, 0°C). Hence HDX rates of backbone amide hydrogens serve as a sensitive probe for protein conformation. The HDX reaction can be acid (H_3O^+), base (OH^-), and water (H_2O) catalyzed, however, effects from water catalysis are minimal. The chemical HDX rate constant (K_{ch}) of an amide in an unstructured peptide is as follows:

$$K_{ch}=k_{int,H}[H^+]+k_{int,OH}[OH^-]+k_{int,H2O}[H_2O] \tag{4}$$

where $k_{int,H}$, $k_{int,OH}$, and $k_{int,H2O}$ are the intrinsic rate constants for the acid, base, and water catalyzed reactions respectively[56,57]. Figure 1.8 demonstrates the mechanisms of base- and acid-catalyzed backbone amide HDX. In base-catalysis, OH^- abstracts the amide proton through a nucleophilic attack forming an amidate ion. Subsequently, the amidate ion is then reprotonated by the deuterium atom from deuterated water (deuterium oxide, D_2O). Acid catalysis involves two possible mechanisms: N-protonation or O-protonation. N-protonation is similar to that of base catalysis except the protonation order is reversed; the initial step is the protonation of the backbone amide hydrogen by D_3O^+ followed by deprotonation of the N-protonated intermediate. O-protonation involves protonation of amide oxygen by D_3O^+ leading to amide proton acidification. The amide proton is then abstracted by water producing imidic acid which is subsequently protonated by D_3O^+ to form the labeled amide[58].

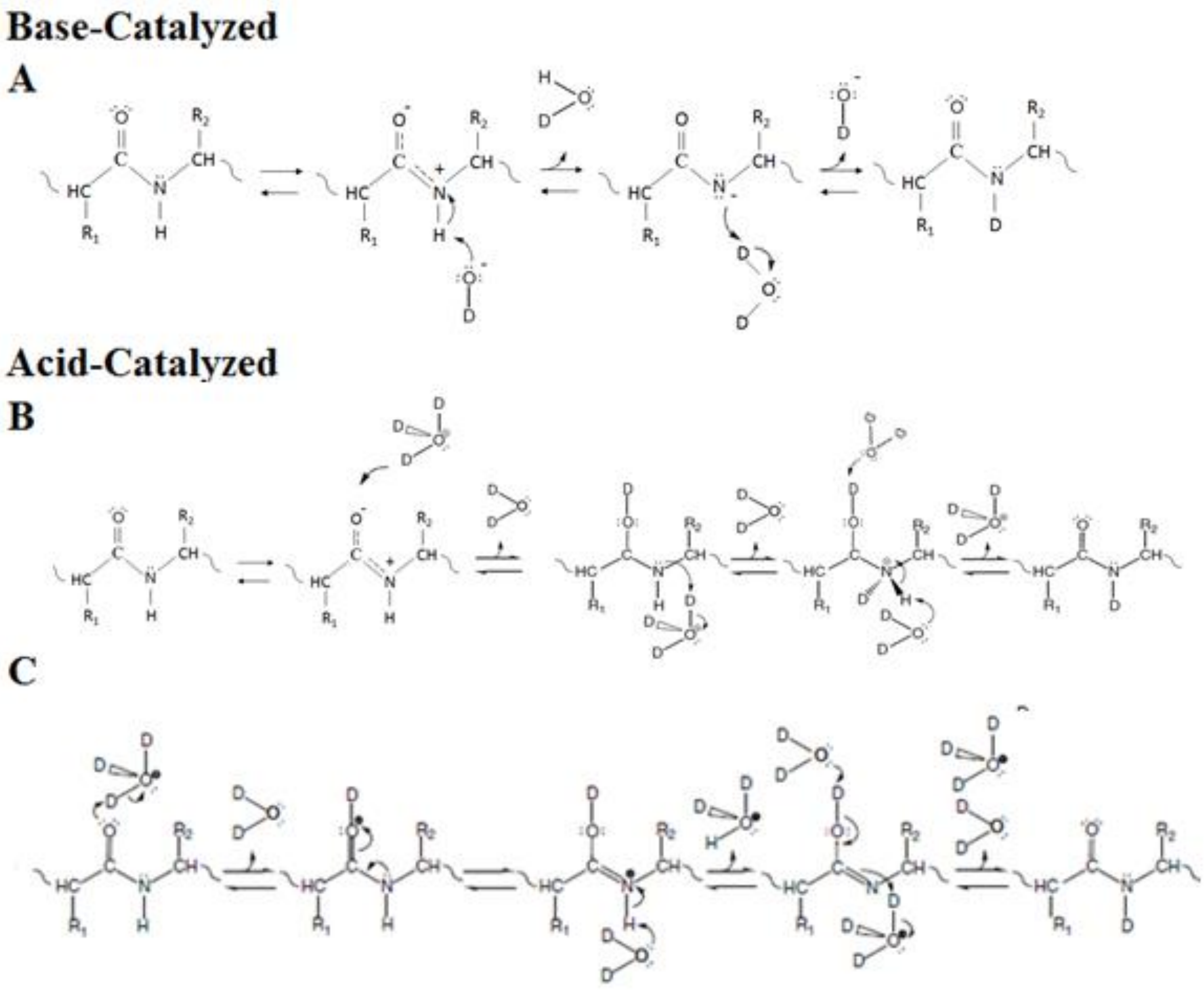


Figure 1.8. Chemistry of base- (A) and acid- (B and C) catalyzed backbone amide HDX mechanisms. Under native conditions (pH 7.0), HDX is mainly driven by base catalysis where the backbone amide is deprotonated by deuterium hydroxide. Acid catalysis can occur via N-protonation (B) and O-protonation (C). N-protonation mechanism is similar to base catalyzed HDX except the protonation order is reversed; protonation of amide backbone is followed by deprotonation. O-protonation occurs through acidification of the amide backbone upon amide oxygen protonation resulting in the formation of imidic acid. This imidic acid is then deuterated to form the deuterated amide backbone. Amides adjacent to electron withdrawing and donating sidechains proceed through O- and N-protonation respectively[58]. Adapted and modified from Weis[59].

As mentioned previously, HDX rates are highly dependent on pH and temperature. For example, the rate constants determined for a poly-DL-alanine peptide at 20°C has $k_{int,H}$ = 41.67 $M^{-1}min^{-1}$, $k_{int,OH}$ = 1.12 x 10^{10} $M^{-1}min^{-1}$, and $k_{int,H2O}$ = 3.16 x 10^{-2} $M^{-1}min^{-1}$. From these data, it is apparent that base catalyzed rates far exceed the other two with water catalysis being negligible[56]. A plot of log(K_{ch}) vs pH generates a V-shaped chevron curve with a characteristic minimum (pH_{min}) at pH 2.5-3.0 at which the acid and base catalyzed HDX are equal[57]. Typically, HDX is performed at pH 7 which is relevant for most proteins and quenched at pH 2.5. This decrease in pH decreases the HDX rate by approximately 4 orders of magnitude[56,57]. Hence by reducing pH, the reaction can be quenched sufficiently such that MS can detect differential changes in deuterium incorporation.

An increase in temperature increases the chemical HDX exchange rates by altering the water ionization constant, k_w (increases the concentration of OH^-). The relationship between K_{ch} and temperature can be explained using the Arrhenius equation,

$$K_{ch}\,(T)=k_{rc}(293)\exp(\frac{-Ea}{R}\left(\frac{1}{T}-\frac{1}{293}\right)) \qquad (4)$$

where $k_{rc}(293)$ is the reference constant, E_a is the activation energy for acid (14 kcal mol^{-1}), base (17 kcal mol^{-1}), and water (19 kcal mol^{-1}) catalyzed exchange, and R is the gas constant (8.134 J mol^{-1} K^{-1}). This relationship is an exponential function[56]. Decreasing temperature from 25 to 0°C, leads to approximately a 14-fold decrease in exchange rate. Hence, finely controlling these two factors, pH and temperature, can greatly reduce the exchange rate as well as the back-exchange reactions[52,56].

In addition to pH and temperature, sequence can also have an effect on K_{ch}, especially inductive or steric effects from neighbouring side chains[56]. Although the effect is not as pronounced as the previous two factors, they should be considered in some cases. For example, polar side chains increase the acidity of the amide backbone through inductive effects hence accelerating the exchange rates. Bulky side chains, however, decreases the rate by steric effects and could result in a 10-20 times slower exchange rate based on pH[56].

1.4.2 HDX- Well Folded Proteins

Most proteins in nature exhibit well-defined tertiary folding with the exception of intrinsically disordered proteins (IDPs). In the case of higher order structured proteins, the major factors affecting HDX rates mainly arise from intramolecular hydrogen bonding and solvent accessibility rather than the effects of pH and temperature[60,61]. HDX rates of amides in unstructured peptides can have a half-life of milliseconds to seconds, whereas in native folded proteins, they can range from milliseconds to years[60]! Furthermore, proteins in solution are not static, but exhibit transient conformational dynamics (breathing motions)[50,62]. During this brief period, some intermolecular H-bonds are broken, making them readily available for HDX. Hence the HDX kinetics for various states of proteins can be expressed by:

$$\mathrm{cl\ (H)} \underset{\mathrm{k_{cl}}}{\overset{\mathrm{k_{op}}}{\rightleftarrows}} \mathrm{op\ (H)} \underset{\mathrm{D_2O}}{\overset{\mathrm{k}_{ch}}{\rightarrow}} \mathrm{op\ (D)} \underset{\mathrm{k_{op}}}{\overset{\mathrm{k_{cl}}}{\rightleftarrows}} \mathrm{cl\ \ (D)} \tag{5}$$

where cl and op corresponds to closed and open states respectively. H and D represents protonated and deuterated states. k_{op} and k_{cl} are the rate constants for the opening and closing states. Proteins in solution are in equilibrium between both states hence the exchange process is mainly affected by the rates of unfolding/folding and its intrinsic exchange rate, k_{ch}. Hence the exchange rate constant (k_{HDX}) of protected amides can be expressed as[62,63]:

$$k_{HDX} = \frac{k_{op}\ X\ k_{ch}}{k_{op}+k_{cl}+k_{ch}} \qquad (6)$$

Under native conditions, the frequency of folding far exceeds unfolding ($k_{cl}>>k_{op}$), so equation (6) could be reduced to:

$$k_{HDX} = \frac{k_{op}\ X\ k_{ch}}{k_{cl}+k_{ch}} \qquad (7)$$

Based on the timescale of the opening and closing events, there are two possible scenarios: $k_{cl}>>k_{ch}$, and $k_{cl}<<k_{ch}$. When $k_{cl}<<k_{ch}$, the rate of chemical exchange is much faster compared to the rate of protein folding. Hence, during the unfolding event, the solvent exposed amide backbones are all exchanged due to sufficient time for the exchange process to occur. In this scenario, the amide exchange rate, k_{HDX}, is highly dependent on the rate of protein unfolding, k_{op}. This mechanism is called EX1 kinetics and equation (7) can be further simplified to:

$$k_{HDX} = k_{op}\ (k_{cl} \ll k_{ch},\ EX1) \qquad (8)$$

Most stable proteins at their physiological conditions, however, exhibit a much faster refolding rate compared to chemical exchange rate, k_{ch} , in other words, $k_{cl}>>k_{ch}$. This is called EX2 kinetics where exchange occurs through multiple opening and closing states of the protein. Hence, EX2 probes the structural fluctuations occurring in the native state. Equation (7) can be reduced to:

$$k_{HDX} = \frac{k_{op}\ X\ k_{ch}}{k_{cl}} = K_{op}\ X\ k_{ch}\ (k_{cl} \gg k_{ch}, EX2) \qquad (9)$$

where K_{op} defined by the ratio of k_{op}/k_{cl} is the equilibrium constant of the opening and closing reaction.

Equation (9) demonstrates that k_{HDX} is proportional to K_{op} and this information can be used to determine thermodynamic properties of protein as shown in equation (10).

$$\Delta G^\circ = -RTlnK_{op} = -RTln(k_{op}/k_{cl}) = -RTln(k_{HDX}/k_{ch}) = RTln(PF) \tag{10}$$

where R is the ideal gas constant, T is the temperature, and PF is the protection factor. A PF=1 represents no protection from exchange such as in the case of unstructured peptides. When PF>1, there is some level of protection from exchange due to the surrounding environment (i.e., solvent protection or the presence of secondary structure)[64,65].

Another advantage of using MS over NMR is its ability to detect EX1 versus EX2 kinetics (resulting spectrum from NMR is an average). For example, Figure 1.9 demonstrates the isotopic envelope evolution as a result of either EX1 or EX2 exchange. EX2 kinetics exhibits characteristic binomial isotopic distribution where its centroid value increases gradually as a function of mixing time. In contrast, regions undergoing EX1 kinetics exhibit bimodal isotopic distribution patterns because of full deuteration during a single opening event. This pattern could be completely separated or partly merged depending on k_{op}[66].

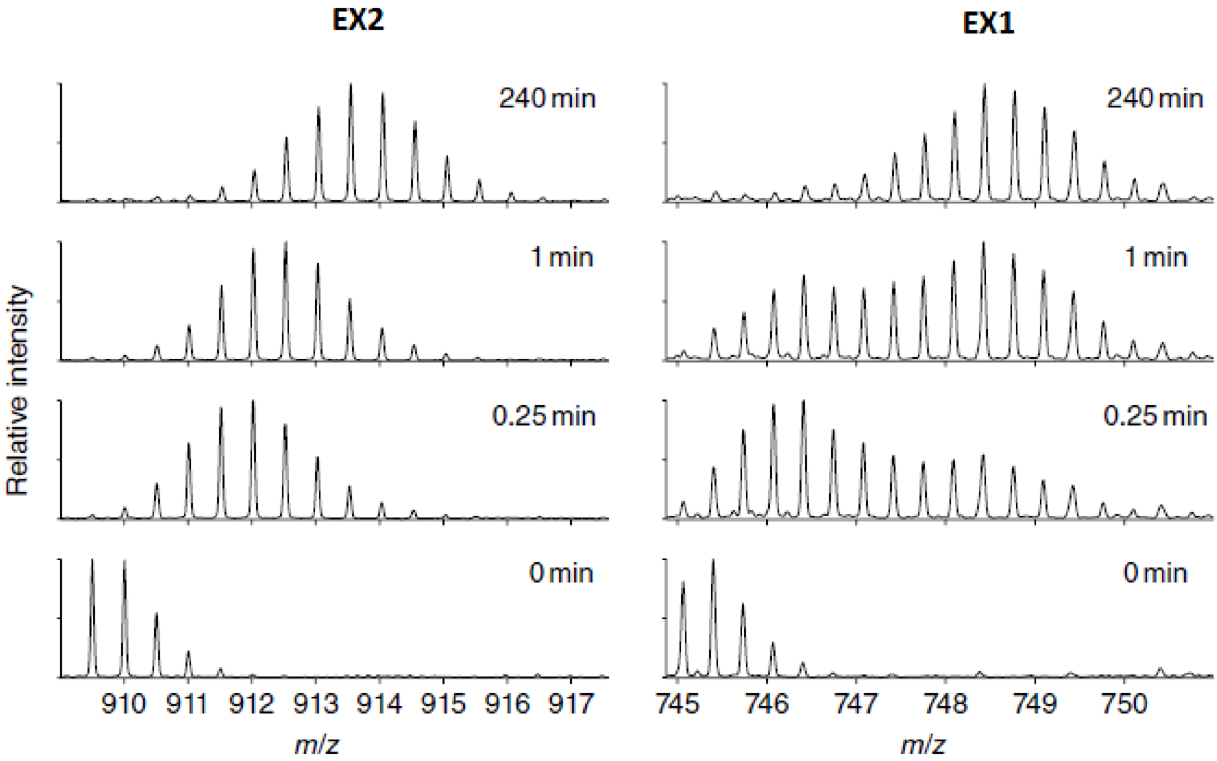


Figure 1.9. MS characteristics of EX1 and EX2 kinetics. EX2 kinetics exhibit a characteristic binomial distribution of isotopic envelope and increases gradually as a function of mixing time. EX1 kinetics is characterized by distinct envelopes and the intensities vary as a function of mixing time. Adapted from Weis[59].

1.4.3 HDX- Disordered Proteins

Intrinsically disordered proteins (IDPs) explore conformational spaces where most of the amide hydrogens are exposed which translates to IDPs transiently occupying its closed state, $k_{op}>k_{cl}$[59,67]. The approximation above for folded proteins assumes that the rate of unfolding is much smaller than rate of folding, $k_{cl}+k_{ch}>>k_{op}$[68]. However, if the interconversion between open and closed states are much faster than the chemical exchange rate as in the case of IDPs ($k_{ch}<k_{cl}<k_{op}$), then the EX1 and 2 regime would never occur (refer to equation (6) and (7)),

contradicting the approximation[59]. To explain the HDX data for IDPs, the Linderstrøm-Lang mechanism must be modified without any approximations and as follows [59,69].

$$[cl] = (1 - f_{op})[\left(\frac{k_\alpha}{k_\alpha - k_\beta}\right)e^{-k\beta t} - \left(\frac{k_\beta}{k_\alpha - k_\beta}\right) e^{-k\alpha t}] \qquad (11)$$

$$[op] = f_{op}[\left(\frac{k_\alpha - k_{ch}}{k_\alpha - k_\beta}\right)e^{-k\beta t} - \left(\frac{k_\beta - k_{ch}}{k_\alpha - k_\beta}\right) e^{-k\alpha t}] \qquad (12)$$

$$[ex] = 1 - [\left(\frac{k_\alpha - k_{ch}f_{op}}{k_\alpha - k_\beta}\right)e^{-k\beta t} - \left(\frac{k_{ch}f_{op} - k_\beta}{k_\alpha - k_\beta}\right) e^{-k\alpha t}] \qquad (13)$$

In equations (11-13), [*cl*], [*op*], and [*ex*] represents the concentrations of closed, open, and exchanged forms respectively. k_α and k_β represent fast and slow processes respectively. The kinetics in equations (11-13) demonstrate biexponential functions and under steady state conditions when $k_{cl}+k_{ch}>>k_{op}$, the chemical exchange rate from equation (13) could be reduced to:

$$[ex] = 1 - \mathrm{e}\hat{}\,\text{-}\left(\frac{k_{op}\,k_{ch}}{k_{cl}+k_{op}+k_{ch}}\right)^{t} \qquad (14)$$

This can be further reduced to explain EX1 and EX2 kinetic regimes using equations (8) and (9)[59,69]. Despite the complexity of the equations, one can still derive insights from HDX of IDPs. For example, bottom-up HDX (peptide-level) generates HDX kinetics that are segment-averaged, which are summations of a series of equations. With enough overlapping peptides, one can potentially deduce exchange rates of individual sites[59]. There have been several studies where variants of the above equations were used to study IDPs[65,65,70–72]. For example, Keppel *et al.* used PF to map the disordered to ordered transition upon coupled binding of CREB binding protein and activator of thyroid and retinoid acid receptors[65].

Residual secondary structures can also be studied using HDX[56,73]. For example, the kinetics will exhibit the same level of deuterium uptake however at an attenuated rate due to formation of transient structures[56,73]. Similarly, when the population of transient structures are ~6% and the interconversion dynamics are slow (1s in closed state), the kinetic rates are comparable with attenuated deuterium uptake[56,73]. These examples illustrate that even small perturbations of order in IDPs can be detected via HDX. These states, however, are invisible in other spectroscopic techniques[74].

1.4.4 Time Resolved HDX (TR-HDX)

Conventional HDX workflows employ labelling times as short as 10s and is well-suited for characterizing changes in structure and conformational dynamics of proteins[75]. However, regions of proteins involved in biological activities are often flexible or unstructured and undergo full deuterium exchange under conventional mixing times[65,76,77]. Hence to probe these weakly structured regions, millisecond timescale HDX should be employed.

The first millisecond HDX measurements were conducted using quench-flow pulse labeling and analysed by NMR[78,79]. Dharmasiri and Smith then introduced millisecond HDX labelling using quenched-flow MS for spatially and time resolved measurements[80]. Konermann and coworkers pioneered continuous-flow millisecond HDX, however the technique at the time lacked spatial resolution[81]. Ultimately, Wilson and coworkers achieved spatial resolution using a 'bottom-up' workflow through the use of a microfluidic device (Figure 1.10)[72]. Briefly, the setup consists of a kinetic mixer for millisecond to second HDX labeling, a static mixer such as a Valco-T mixer for HDX quenching, a microreactor containing acid proteases such as pepsin and/or protease XIII crosslinked to agarose beads for rapid digestion, and an on-chip incorporated electrospray needle[72,82]. The entire process is performed online with labelling times ranging from 40ms to 12s. Labelled proteins are efficiently digested in the proteolytic chamber in under 8s minimizing back-exchange in the absence of cooling. Furthermore, the electrospray needle incorporated at the outlet of the microfluidic chip enables online analysis[72].

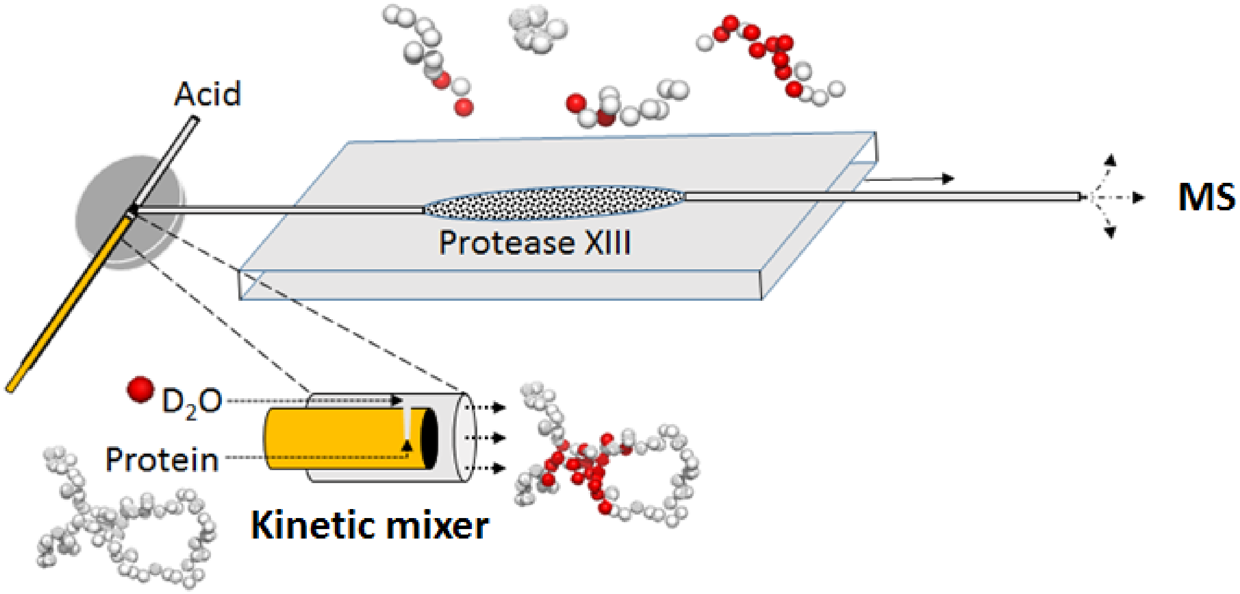


Figure 1.10. A schematic depiction of TR-HDX performed on a microfluidic device. The kinetic mixer and acid channel (pH 2.4) are connected to a Valco T-mixer. HDX occurs in the kinetic mixer and is efficiently quenched in the conjunction of the Valco T-mixer. Labelled protein is then transferred to the downstream proteolytic chamber containing either pepsin or protease XIII covalently linked to agarose beads where digestion occurs. The distal capillary is used as an ESI source to transfer the digested peptides into the mass spectrometer.

TR-HDX with millisecond mixing times can probe the relative conformational flexibility such as changes in dynamics upon ligand or protein binding. An example of early work on the study of weakly ordered regions was on mildly denatured myoglobin[81]. Using subsecond mixing times (7ms to 3.1s), Simmons *et al.* were able to detect five similar folded states, one EX1, and three EX2 transition states providing a detailed picture of conformational ensembles during the millisecond mixing regime[81]. Another example of TR-HDX on weakly structured regions was performed on native cytochrome c[72]. The authors implemented TR-HDX on a microfluidic platform with the bottom-up HDX approach. The results were reported in terms of segment-averaged PFs ranging from 10-36. The unstructured regions that exhibited the highest PFs were those in close contact with the heme cofactor.

TR-HDX has been used to study transient intermediate species during the protein folding events that is invisible by other classical techniques. For example, when ubiquitin was allowed to fold for 40ms, TR-HDX identified a species that acquired a folded-like ESI charge state distribution, yet exhibited a denatured HDX profile[83]. Although these are not structurally resolved, with application of bottom-up TR-HDX workflow, several intermediates have been well resolved for myoglobin, cytochrome c, calcium binding to S100A11 signalling protein, and bacteriorhodopsin[84–86].

TR-HDX has also been implemented to isolate and characterise catalytically active intermediates. For example, a study conducted by Liuni *et al.* showed that resting and acyl-enzyme states of chymotrypsin sampled the same conformational space. Interestingly, however, their rates were different; the acyl-enzyme species was undergoing faster conformational searching during catalysis which the authors described as “intensified dynamics”[87]. This was a novel finding and a new enzyme catalysis model proposed through use of TR-HDX, highlighting the potential of this technique in providing insights to new fundamental mechanisms in enzyme catalysis.

1.4.4.1 TR-HDX- Application: IDDs and IDPs

IDPs lack or have very little secondary or tertiary structures[67]. Under physiological conditions, IDPs sample a wide range of conformational ensembles[67,88]. Despite its unstructured nature, there are small regions that adopt transient secondary structures that are invisible to the conventional HDX workflow[89]. However, TR-HDX can detect these fleeting species due to the fast labelling timescale.

Recently, TR-HDX was applied to study a conformational ensemble shift in the IDP tau upon hyperphosphorylation[70]. Please refer to Chapter 3 for the study and detailed explanation. Briefly, the study implemented TR-HDX on a microfluidic device to probe the PFs of native and hyperphosphorylated tau. Tau hyperphosphorylation represents one of the hallmarks in Alzheimer's disease (AD)[90]. In the native state, substantial protection was observed in the aggregation prone regions called hexapeptides that are buried in the core of the protein. Furthermore, the regions surrounding the hexapeptides also showed attenuation in deuterium uptake, indicative of a protection mechanism. Upon hyperphosphorylation, however, one of the hexapeptides was solvent exposed along with release of the native globular conformation. Exposure of this hexapeptide could potentially provide an aggregation template for the seeding process. This study provided a detailed picture of the conformational shift that occurs upon hyperphorylation as well as evidence for the causative role of tau amyloidosis in AD[70].

In the same year, Keppel and Weis applied millisecond HDX and reported the presence of residual helical structure in the C-terminus of activator of thyroid and retinoid acid receptor (ACTR)[91]. Their analysis approach was different in a sense that instead of fitting HDX kinetic data, they interpolated half-life exchange (time needed to reach 50% deuterium uptake) to quantify the data[91]. The N-terminus of ACTR exhibited faster exchange compared to the C-terminus. The attenuation in C-terminus was rationalized by the presence of aspartic acid which slows the exchange rate via electrostatic repulsion[56]. Nevertheless, other complimentary techniques such as NMR and computational simulation supported the presence of transient helical structure in the C-terminus[91,92].

1.5 IDPs- Current Perspective

Proteins are powerful molecular machines that are constantly at work to keep all of us viable[93]. This reason alone spurred structural biologists to gain insights into their chemistry and biological functions. It has been well established that proteins require motions/dynamics to carry out their functions and these motions sample a wide range of conformational space at different timescales with varying amplitudes giving rise to different conformational states[93,94]. Ultimately, the goal is to characterise proteins' free energy landscape with all relevant conformational states that are catalytically active and their respective kinetics and thermodynamics[94]. Thus far, there is overwhelming data on proteins' structures in their ground states because these are the most stable species in solution and also due to technical limitations of the technologies present at the time. This scope has provided a strong correlation between their well folded three-dimensional structure and their function resulting in the concept of the structure-function paradigm[95].

However, with the advent of high resolution biophysical techniques such as NMR, it is becoming increasingly evident that higher energy states are more relevant to protein function rather than their ground states. This often challenges the structure-function paradigm and lock-and-key mechanisms suggesting that protein function is dictated by their dynamics rather than their structure[96,97]. Proteins at higher energy states contain more "disordered" regions and are highly dynamic exhibiting large degrees of motion, interconverting between different heterogeneous conformers[93].

The lack of structure in IDPs and IDDs can have several advantages over folded proteins. For example, IDPs can provide a larger surface area for interactions with various binding partners hence exhibiting their versatility[98,99]. Furthermore, their malleability enables them to finely tune affinities towards various partners with different degrees of specificity and also present motifs that are more accessible for post-translational modifications (PTMs)[94,100,101]. For example, IDPs sometimes exhibit disorder-to-order transition upon binding with high specificity and low affinity[88,93]. These modes of binding are essential in regulatory and signalling processes which involve transient protein-protein interactions[102,103]. Interestingly, disordered regions comparatively have higher mutation rates without affecting stability or function[93,104]. Furthermore, disorder may help limit overall protein size and crowding effects in cells[105].

1.5.1 IDP- Structure and Function

IDPs and IDDs are disordered mainly due to their primary sequences which are predominantly comprised of charged polar residues as well as glycine and proline with minimal hydrophobic residues[106]. Functionally, IDPs are very diverse due to the presence of multiple binding motifs, encompassing essential biological roles (i.e., transcription, signalling, cell cycle regulation) to pathological roles (amyloidosis resulting from protein aggregation)[94,102,103]. Energetically, IDPs lie high on the free energy landscape and sample numerous conformations that can be functionally relevant, enabling recognition of different binding partners to carry out essential biological functions (Figure 1.11)[94,107]. However, when misregulated either through faulty PTMs or a degradation pathway, IDPs can funnel through an “off-pathway” resulting in misfolded toxic oligomers[107]. These toxic intermediates are problematic because typically hydrophobic residues are solvent exposed, presenting a sticky surface allowing for other monomeric forms to undergo aggregation through intermolecular interactions.

Most of the time, IDPs mediate protein recognition through numerous binding motifs and the resulting complexes facilitate allosteric regulation[108,109]. Furthermore, due to accessibility, IDPs and IDDs are often a target for PTMs that act as molecular switches and also affect their binding affinities towards certain targets[93]. It is also well known that many IDPs interact with their partners in an extended conformation such that it wraps around their targets, increasing the buried surface area resulting in high specificity[98,99]. In addition, IDPs act as scaffolding or “hub” proteins through distinct motifs to integrate and localise signals that are central to protein interaction networks[101,110,111]. Due to their high charged residue content, IDPs are highly soluble and can act as a “detergent” to facilitate solubility of other proteins as well[94]. For example, the disordered region of clusterin (glycoprotein) can solubilise bacteriorhodopsin[112]. However, when hydrophobic residues are exposed to the solvent, there is a tendency for them to aggregate with other monomers resulting in aggregation based diseases such as AD[70,90,94,113]. Furthermore, if their concentrations are too high, they can bind to other physiological targets. Examples include studies of SH2 and PDZ domain containing proteins[114,115]. Hence, the tight regulation of IDPs is essential to maintain cell viability. This is done through help of scaffolding proteins and through degradation pathways (IDPs and IDDs are susceptible due to their unstructured nature)[93,116]. Despite the tight regulation, there are exclusions where IDPs need to be present in high concentrations or for long durations because they become part of a stable protein complex to perform relevant biological functions[117]. An example is the activation of tumor suppressor protein, p53, in response of stress stimuli. In unstressed cells, p53 is bound to its negative regulator mdm2 and is targeted for degradation. However, upon stress, its N-terminus is phosphorylated which results in subsequent p53 activation[118].

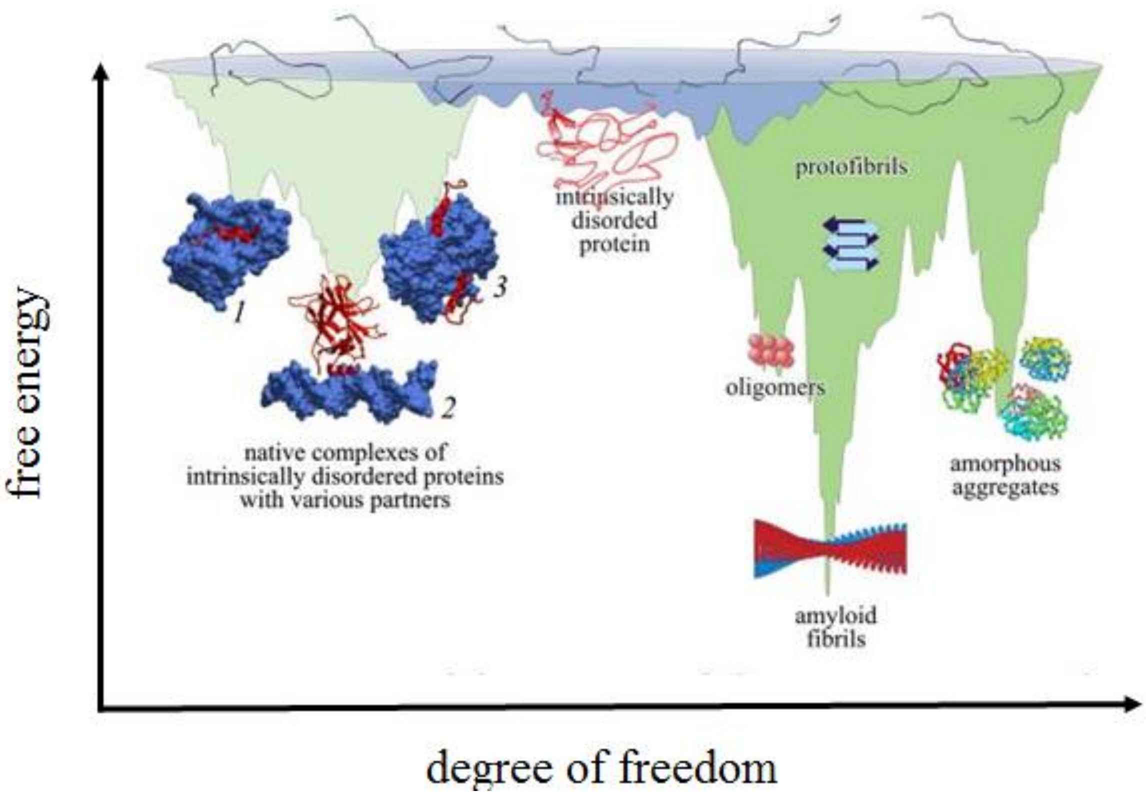


Figure 1.11. Schematic of the protein folding and aggregation landscape of IDPs. The funnel is split into two sides; physiological (light green) and toxic (dark green). Left side represents newly synthesized unfolded protein funneling to its native state through a multitude of intermediate conformations via intramolecular interactions and help of chaperone proteins. Right side represents folding of conformations into toxic oligomers, aggregates, and amyloids through intermolecular interaction when misregulated. Adapted and modified from Turoverov *et al*[119].

1.5.2 IDP association with human diseases

Although IDPs and IDDs are deeply involved in crucial biological functions, their dysfunction can result in a wide range of human diseases[120]. For example, IDP dysfunction and aggregation is the main culprit in numerous neurodegenerative diseases such as AD and Parkinson's. Furthermore, it was found that ~79% of cancer-associated proteins contain intrinsically disordered regions of at least 30 residues[121]. In the case of human papillomavirus (HPV) , it was shown that an increased amount of non-structural proteins (E6 and E7) resulted in a high risk class of HPVs[122]. So why are IDPs prone to development of pathological diseases? Despite IDPs versatility and advantages, a subset of those features are also the key drivers in the disease pathology[120]. For example, the larger surface area of IDPs and IDDs for PTMs make it more susceptible for impaired modifications resulting in abolishment in some of the essential cell signaling processes as in the case of p53 N-terminal transactivation domain (NTD)[120]. Alternative splicing (AS) can diversifies the proteome and higher frequency of AS could be tolerated in IDPs because its functional and structural features are not perturbed[123,124]. However, this also makes IDPs more vulnerable such that it could weaken or abolish essential protein-protein interactions as well as generating fragments that are more prone to forming protein aggregates[43,120,124]. Abnormal AS is associated with myotonic dystrophy, cancer, and AD, just to name a few[43,90,125,126].

The most notable pathological phenomenon for IDPs and IDDs, however, is the formation of protein aggregates due to protein misfolding or “off-pathway folding” as in the case of numerous neurodegenerative diseases and amyloidosis[116,120]. Misfolding can occur due to impaired regulation, missignaling, and misidentification[120]. One of the common features resulting from misfolding is the aggregation or the formation of highly-ordered and stable filamentous protein aggregates known as amyloid fibrils[127]. The fibrils are comprised of cross β-sheets in the core running perpendicular along the fibril axis[127]. Two to six protofilaments are associated laterally or twisted forming larger fibrils 4-13nm in diameter[127]. Structurally, amyloids from different diseases are similar in morphology. However, the aggregation-prone intermediates are vastly different for different IDPs and it is speculated that these are the toxic species in the formation of the diseases[120].

1.5.3 IDP- p53 N-terminal Transactivation Domain and Cancer

p53 is a tumor suppressor protein that plays a major role in cellular stress responses such as DNA damage, hypoxia, and other stresses[128]. It is known that its inactivation either through mutation or deregulation of associated pathways can result in the majority of known cancers[129]. Once activated, it regulates the cellular response of apoptosis, cell cycle arrest, and DNA repair depending on the nature of the stress[128] (Figure 1.12). Structurally, p53 is composed of mainly four distinct domains: NTD, DNA binding domain, tetramerisation domain, and C-terminal domain[128]. p53 becomes activated and stabilised through a series of PTMs[128,130,131]. Most common is phosphorylation of the p53 NTD that is critical for protein-protein interactions in the regulation of its activity and stability[128]. Furthermore, it also plays a major role in dictating modifications of other domains of the protein such as acetylation of the C-terminal domain[130]. The NTD is intrinsically disordered and can provide multiple conformational states facilitating in interaction of numerous binding proteins through the modulation of PTMs. The NTD can be divided into two subdomains: TAD1 and TAD2. Both of these domains are known to contain transient secondary structures in the unbound state, however it has been shown to adopt helical structure upon complex formation with binding partners through coupled folding driven by hydrophobic interactions[132–134]. p53 NTD has been shown to interact with various transcriptional regulatory proteins such as TATA binding protein, CBP/p300, and p62 subunit of general transcription factor II H (GTFIIH)[134–136]. Most of these interactions are dictated through PTMs such as phosphorylation because it modulates the binding affinities of different binding partners that are known to bind at the same sites[129]. For example, double phosphorylation at Ser46 and Thr55 increases its binding affinity to Tfb1 in an additive manner[129,134].

Hence p53 NTD plays an important role in p53 activation and downstream activities. Furthermore, phosphorylation can serve as a molecular on/off switch. Studies from mouse models further highlighted the importance of NTD and its modifications. Mutant knock-in and knock-out studies demonstrated impairment in p53-dependent apoptosis and/or G1/S cycle arrest. For example, homozygous mice containing Ser23Ala mutation, Ser23 (human S20) is an important phosphorylation site for p53 activation, had much shorter life spans and developed B-cell lymphoma[137]. Many of these mutations on the NTD reduced p53 activity by abolishing p53 interactions with other transcriptional coactivators.

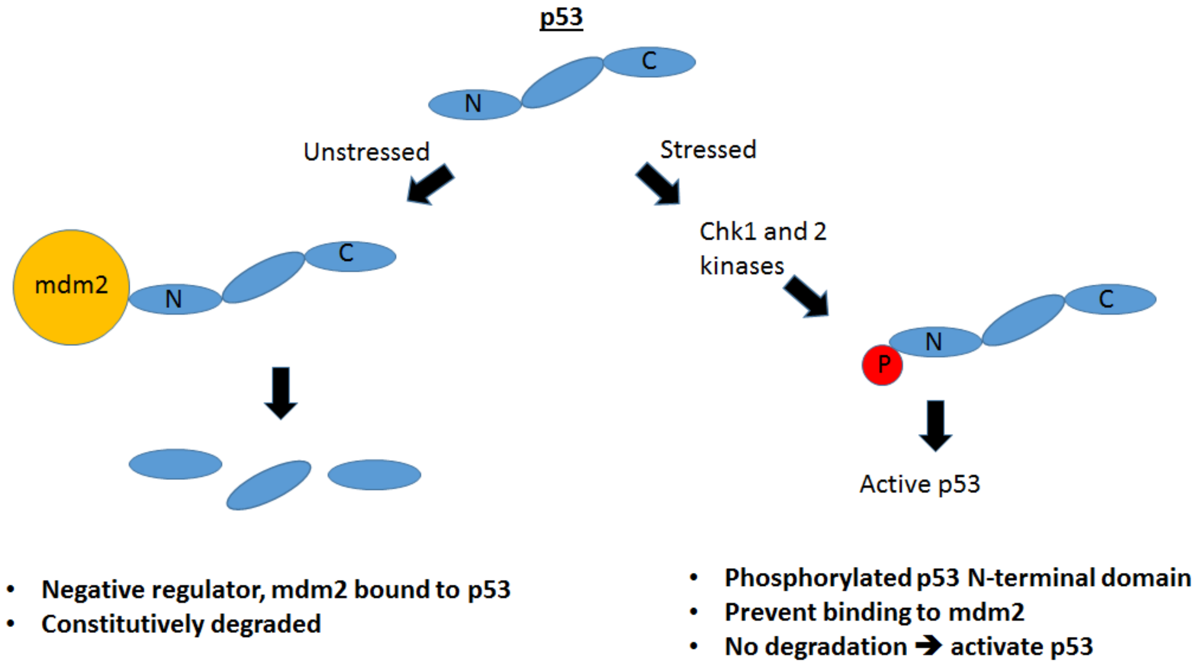


Figure 1.12. Simplified mechanism of p53 activation. p53 is shown in blue oval nodes, N and C represents the N- and C-termini whereas the middle node is the DNA binding and tetramerisation domain. When unstressed, p53 is bound to its negative regulator, mdm2 and is targeted for degradation. Upon cellular stress such as radiation induced DNA damage, p53 is activated by phosphorylation at its N-terminus by kinases such as chk1 and 2. This results in dissociation of mdm2 and stabilization of p53. Once activated, p53 induces downstream pathways involved in cell cycle arrest, apoptosis, and DNA repair.

1.5.4 IDP- Tau and Alzheimer's Disease

Tau is an IDP that has a function in microtubule binding (Figure 1.13)[90]. Under physiological conditions, tau mainly functions in binding and stabilizing microtubules, maintaining the cytoskeleton morphology of the neuronal cells[90,138]. There are total of 6 tau isoforms generated by alternative splicing in humans that range from 352-441 amino acids[139]. These are also intrinsically disordered with minimal secondary structures[138,139]. Structurally htau40, full length tau, is composed of projection N-terminus domain and assembly C-terminus domain[90,138]. The projection domain may play a role in establishing a "clear zone" that is responsible for adhering to other tau monomers or other microtubule binding proteins[140]. The C-terminal domain contains 4 repeats of the same 31-32 amino acid sequences that are important for microtubule binding as well as formation of neurofibrillary tangles (NFTs) in AD[90,138]. Despite the disordered nature of tau, it still adopts a fold that resembles a paperclip where both N- and C- terminals are in contact with the repeat domain regions[141]. Tau consists mainly of hydrophilic and charged residues; an acidic stretch is near the N-terminus followed by the basic domains[90]. Another important feature of tau is the presence of proline rich regions in between the N and C-termini. This region mostly contains Ser-Pro (SP-) or Thr-Pro (TP-) motifs that are readily phosphorylated in AD[90,142].

Tau sparked major interest when its aggregation was discovered in AD and other taupathies[90]. Mutations in tau have been shown to be the cause of other neurodegenerative diseases such as frontotemporal dementia, Parkinson's disease, and progressive supranuclear palsy[143,144]. Now it is becoming more evident through animal models that tau is the bullet in AD pathogenesis whereas Aβ is the trigger[145]. Interestingly, tau alone when misregulated is still able to manifest as a risk factor. For example, *in vivo* studies showed that mice with one or both tau genes knocked out but still overexpressing Aβ did not show any memory impairment. Furthermore, its plaque load as a result of Aβ aggregates were the same as that of a parental strain containing both tau genes. However, when both were overexpressed, the mice manifested severe memory impairment and learning deficits[146].

Tau phosphorylation is required for careful regulation of microtubule binding[147]. However, hyperphosphorylation can have detrimental effects and represents one of the hallmarks of AD as aforementioned. In AD patients, the level of phosphate incorporation is at least four times higher than healthy brains[148]. Tau is an easy target for phosphorylation by multiple kinases and contains at least 45 phosphorylation sites[149]. Most of these sites are targeted by proline-directed kinases (e.g. Glycogen Synthase Kinase-3β (GSK3β), Mitogen-Activated Protein Kinase (MAPK), cyclin-dependent kinase 2 (cdc2), c-Jun N-terminal Kinase (JNK), and cyclin dependent kinase 5 (cdk5)) due to the presence of many SP- or TP-motifs within the tau sequence. However, other non-proline directed kinases such as AMPK related families can also phosphorylate tau[90,150]. Hence tau phosphorylation can be used as a biomarker for AD and thus far few antibodies have been raised against these phosphosites, some examples include PHF1 that recognizes pS396 and 404, AT8 for pS202 and 205, and AT100 for pT212 and S214[151,152].

Upon hyperphosphorylation, tau loses its affinity for microtubules, and has higher affinity to itself[90,142]. This subsequently results in tau self-assembly and aggregation into paired helical filaments (PHF) or NFTs[90]. PHFs are amyloids and are characterized by well-ordered stacked β-sheet rich structures[153]. The propensity for such formation is encoded in repeat regions 2 and 3 containing short hydrophobic residues called hexapeptide motifs[154]. The core of the PHF (~20 residues of repeat regions 2 and 3) is established through axially stacked cross-β structures, however, the remainder of the protein (N-terminus and the rest of the C-terminus) is still highly mobile even in the aggregated state and has been described as a “fuzzy coat”[155,156].

Other modifications such as tau proteolytic cleavage, oxidation, glycosylation and acetylation can also play a role in AD[90]. Cleavage of N- or C- termini by caspases has shown to abrogate the paperclip fold and enhance the aggregation process[157]. Oxidation of the cysteine residues can crosslink to form disulfide bonds with other tau monomers, enabling a faster aggregation process[90]. N-glycosylation can aide in the hyperphosphorylation process[158]. Lys280 acetylation impairs tau binding to microtubules, inhibits tau degradation and also hastens tau aggregation[159].

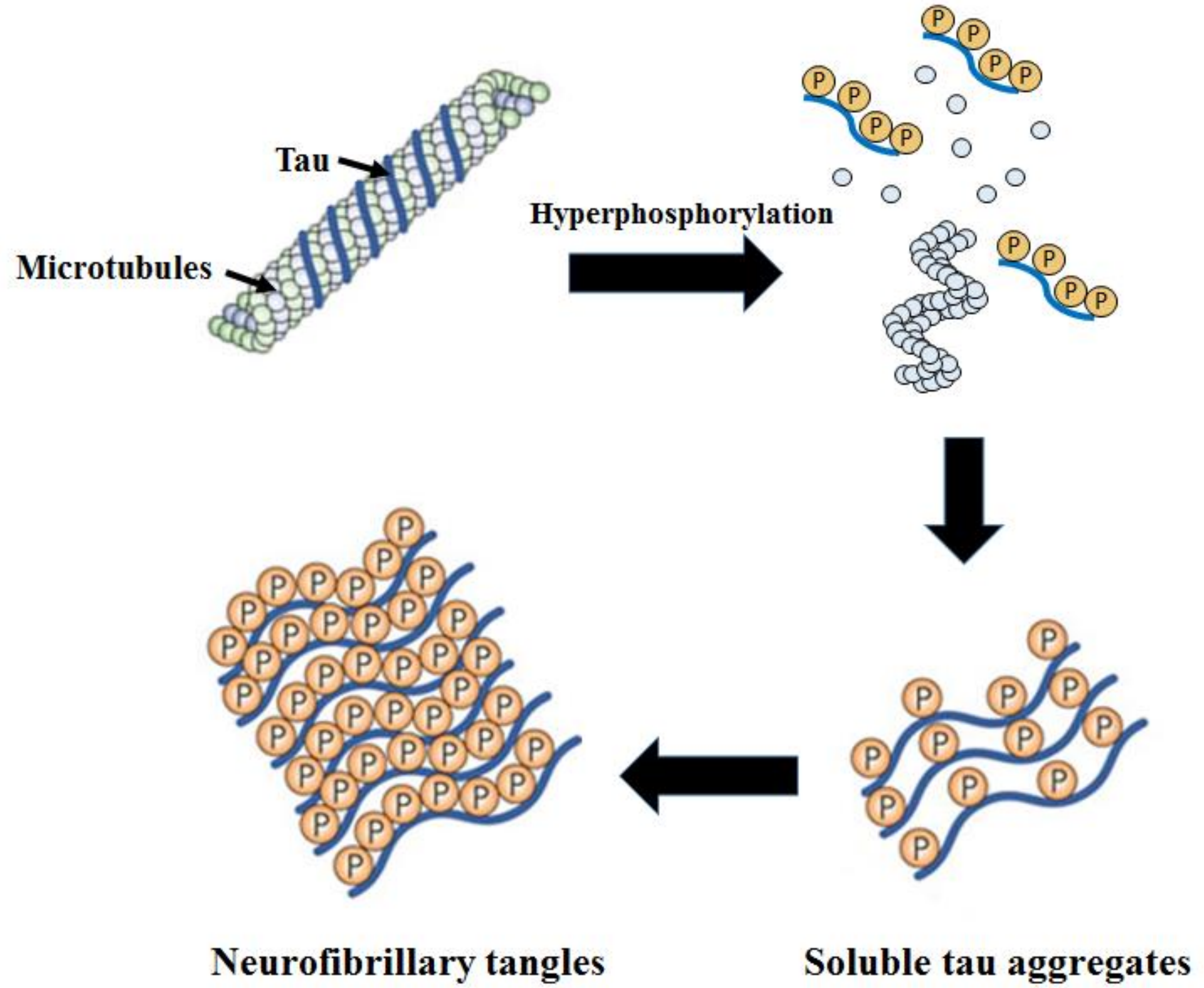


Figure 1.13. Tau: physiological function and mechanism underlying AD. Under physiological conditions, tau binds to microtubules stabilizing the cytoskeleton of neuronal cells. In AD, however, tau is hyperphosphorylated resulting in its dissociating from microtubules and initiating self-aggregation ultimately forming NFTs. Adapted and modified from Citron[160].

1.6 Research Objectives

IDPs and IDDs pose challenges to conventional biophysical techniques due to their intrinsic dynamic nature. TR-HDX, however, represents an alternative powerful structural labeling technique that can probe the dynamic and structural ensembles of these proteins and domains. Chapters 2-4 will implement a TR-HDX microfluidic chip to conduct structural studies on IDPs at sub-second timescales. Chapter 5 will implement gas-phase HDX coupled to DMS technology to study solution phase protein structures.

Chapter 2 focuses on the effect of phosphorylation on the structural ensemble of p53 NTD and its binding affinity in disordered regions in free form as well as upon binding to PH domain of p62 subunit of general transcription factor II H.

Chapter 3 focuses on elucidating the conformational shift and identification of toxic intermediates of tau upon hyperphosphorylation to understand the mechanism driving tau pathology in AD and other tauopathies.

Chapter 4 focuses on the effects of six potential AD drug candidates on the structural ensemble of tau protein as well as probing their binding sites. The aim was to study the mechanism of aggregate inhibition and to provide structural information to further refine and improve the potency of next generation drugs.

Chapter 5 focuses on applying DMS-HDX to study gas-phase protein conformations of folded, partially-unfolded, and unfolded proteins in solution. The aim was to find optimal conditions such that unique DMS-HDX profiles were generated for distinct structural features that can be applied towards rapid qualitative analysis of structural stability.

Chapter 2

Enhanced Binding Affinity *via* Destabilization of the Unbound State: A Millisecond Hydrogen / Deuterium Exchange Study on the Interaction Between p53 and the Pleckstrin Homology Domain

A manuscript of this chapter is in preparation.

- Zhu, S.; Khatun, R.; Lento, C.; Sheng, Y.; Wilson, D. J. Enhanced Binding Affinity *via* Destabilization of the Unbound State: A Millisecond Hydrogen / Deuterium Exchange Study on the Interaction Between p53 and the Pleckstrin Homology Domain

2.1. Summary

Presence of intrinsically disordered domains enables a single protein to engage in a myriad of binding partners, with PTMs such as phosphorylation modulating target specificity and affinity. It is known that phosphorylation can act as a chemical on/off switch for protein complexation, however, this is not a satisfying explanation for enhancement of interactions through coupled folding that are hydrophobically driven, a common phenomenon of IDPs or IDDs through phosphorylation regulated interactions. To determine the effect of phosphorylation on the complexation and its binding affinity in disordered regions, we have examined the

interaction between p53 NTD and the pleckstrin homology (PH) domain of p62 using TR-HDX. Here we demonstrate that phosphorylation did not affect the conformation of bound p53 NTD with PH, but instead the modification increased conformational disorder of NTD in the unbound state. We propose that this increase in disorder upon phosphorylation creates an elevation in the free-energy gap between the free and bound p53 NTD states, resulting in a stronger affinity complex.

2.2. Introduction

It is a widely accepted notion that protein-protein interactions (PPI), whether transient or permanent, are critical to many biological processes such as signal transduction, catalysis, and cell cycle regulation[161–163]. Many of these transiently interacting hubs are enriched in disordered domains whose function is modulated by PTMs that act as molecular switches[130,164]. These covalent modifications can either alter the local chemistry of the protein itself or their targets, ultimately eliciting different effect responses. Although, the structural changes and response effects due to PTM have been widely explored, there is increasing evidence that changes in protein dynamics of IDDs and IDPs upon PTMs can significantly affect PPI dissociation constants (K_d) without a significant change in the native complex ground-state structures[165–167].

p53 NTD is intrinsically disordered and contains many linear motifs that are known to interact with various transcriptional regulatory proteins such as TATA binding protein, CBP/p300, and p62 subunit of GTFIIH, to name a few[134–136]. In unstressed cells, the tumor-suppressor p53 is expressed in low levels and is quite short-lived. However, upon exposure to cellular stress conditions, p53 becomes activated and stabilised through a series of

PTMs[128,130,131]. Once activated, it regulates cell cycle responses such as apoptosis, cell cycle arrest, and DNA repair depending on the nature of the stress. Phosphorylation of the NTD tightly regulates the activity and stability of p53[128].

Interaction of p53 NTD with p62 is important for the recruitment of p53 to the GTFIIH complex to initiate the transcriptional elongation process[168]. Specifically, p53 NTD binds to the PH domain of the p62 subunit[134]. p53 NTD binds to the PH domain with a K_d of 3.2± 0.57 μM in its native form, however when phosphorylated at either position Ser46 or Thr55, the K_d increases by 6 fold to ~500nM and dual phosphorylation at both sites increases the K_d by ~30 fold to 100nM[134]. Binding sites have been identified for p53 NTD/PH domain, with mutagenesis studies identifying Lys18 of the PH domain as necessary for the tighter binding observed for phosphorylated p53 NTD[134]. From structural perspective, this tighter interaction is driven by stabilization of amphipathic helix of p53 NTD driven by hydrophobic interactions[129,134]. However, this observation alone does not provide a satisfying rationale for phosphorylation driven modulation as this helix can be stabilized even in the absence of phosphorylation[165,166].

To examine the effect of phosphorylation-enhanced binding of p53 NTD, the change in conformation and dynamics of p53 NTD upon phosphorylation with Chk2 *in vitro* and binding to the PH domain of p62 subunit was conducted using TR-HDX on a microfluidic chip as shown in Figure 1.10. As mentioned in the introduction, TR-HDX is a structural labelling technique that can detect slightest changes in weak hydrogen bonding network on the millisecond timescale, making it amenable to study residual structures in p53 NTD and solvent accessibility upon binding to PH[59,70,72].

2.3 Results and Discussion

2.3.1 Dynamic analysis of native and phosphorylated p53 NTD

p53 NTD was phosphorylated *in vitro* with chk2, a kinase involved in the p53/p62 interactions, to determine the structural and dynamics changes upon phosphorylation[134]. To localise the phosphorylation sites, the protein was digested using trypsin or pepsin and subsequent enrichment of the phosphopeptides was performed using a titanium dioxide column. Identified sites were Ser6, Ser15, Thr18, Ser20, Ser37, and Ser46. MS/MS spectra for the phosphorylated peptides are provided in the appendix, Figure A1. Ser15, Thr18, Ser20, and Ser37 have been previously identified by Shieh *et al.* 2000, however the remaining sites (S6 and S46) could have been phosphorylated due to a longer incubation time of 120 minutes as opposed to 20 minutes[169]. This was necessary to ensure that phosphorylation goes to completion. Phosphorylation at position Thr18 activates p53 by releasing it from its negative regulator, mdm2 and phosphorylation of Ser46 enhances binding affinity to PH domain by 6 fold as mentioned earlier[134,170]. Ser46 is phosphorylated by ataxia-telangiectasia mutated (ATM) kinase, Dual specificity tyrosine-phosphorylation-regualted kinase 2 (DYRK2), homeodomain interacting protein kinase (HIPK), and AMP-activated protein kinase (AMPK). However, this site is not known to be substrate for chk2[171].

Next, IMS was performed on both native and phosphorylated p53 to investigate their relative cross-sectional areas (CSA) (Figure 2.1). IMS is analogous to electrophoresis where separation is based on CSA of gas-phase ions. Based on the chromatogram (Figure 2.1), native p53 NTD exhibited two peaks; a major peak corresponding to a drift time of 5.87ms and a minor peak at 8.75ms. This illustrates the presence of two gas-phase conformations, with the majority populating a relatively more compact conformation. However, upon phosphorylation, a bimodal distribution was exhibited with equal intensities at both 5.87 and 8.75ms. This indicates a shift to an overall extended conformation of p53 NTD upon phosphorylation, considering native p53 NTD is already in a partially collapsed state[172]. To expand on this finding, TR-HDX was performed on both the native and phosphorylated p53 NTD to localize the areas of conformational change.

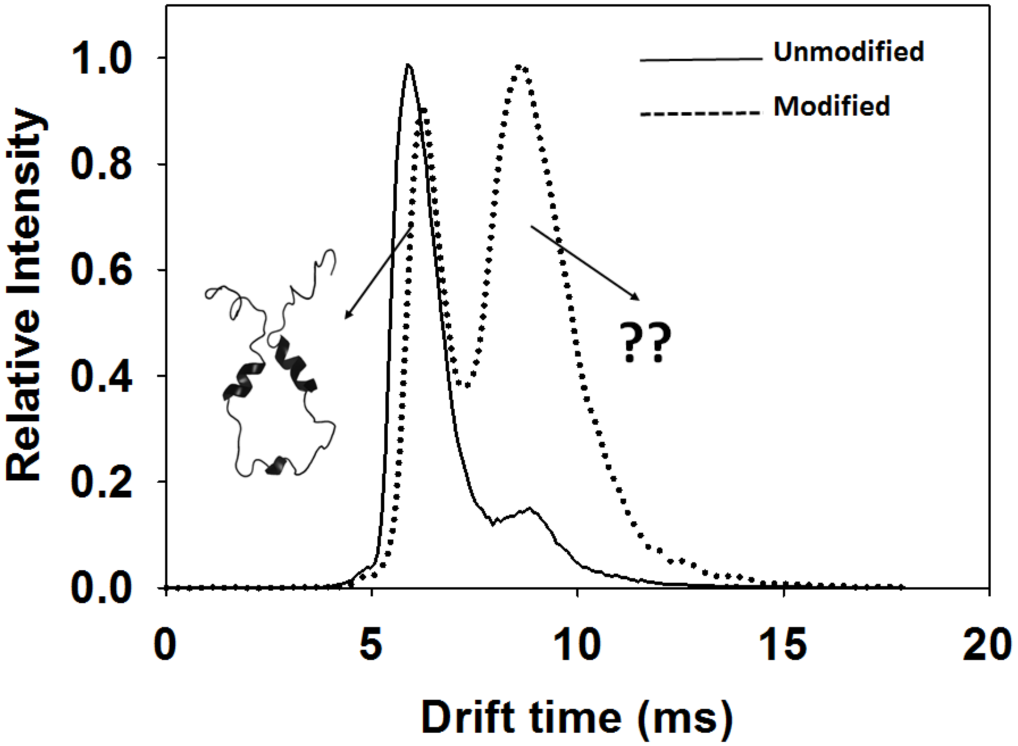


Figure 2.1. Ion mobility chromatograms of native and phosphorylated p53 NTD. Comparatively, native p53 exhibited a more compact conformation (solid line) whereas the phosphorylated counterpart showed a more extended conformation (dashed line).

For TR-HDX, five mixing time points were acquired in triplicates for both native and phosphop53: 0.20s, 0.57s, 0.95s, 1.32s, and 2.06s (Figure 2.2). A sequence coverage of 75% was achieved for both: 8 peptides for native and 16 peptides for phosphop53. This low sequence coverage could be due to short residence time in the proteolytic chamber (typically < 8 s) to avoid significant back-exchange[72,82]. Nevertheless, phosphop53 showed a greater deuterium exchange globally compared with native. This is in agreement with the IMS data, strongly supporting the conclusion that upon phosphorylation, p53 NTD adopts an extended conformation.

Transient secondary strucutres in p53 NTD has been predicted by computational approaches[172–174]. Furthermore, performing secondary structure propensity using Protein Structure Prediction Server showed high helical propensity in residues 16-25, 37-44, and 48-56[172,174]. Mapping the HDX data onto a representative 3D ribbon diagram (generated by PHYRE2 server using PDB 2GS0 as a template) confirmed the presence of residual secondary structures at residues 19-21, 32-43, and 53-59 of native p53 NTD under the fastest mixing time regime of 0.20s as shown by lower deuterium uptake (Figure 2.2). However, when the mixing time was extended, these transient differences were lost since these are short-lived conformations. For example, during 0.57s of mixing, information of the C-terminal end transient helix was lost and at 2.06s, all short lived conformations were not detectable.

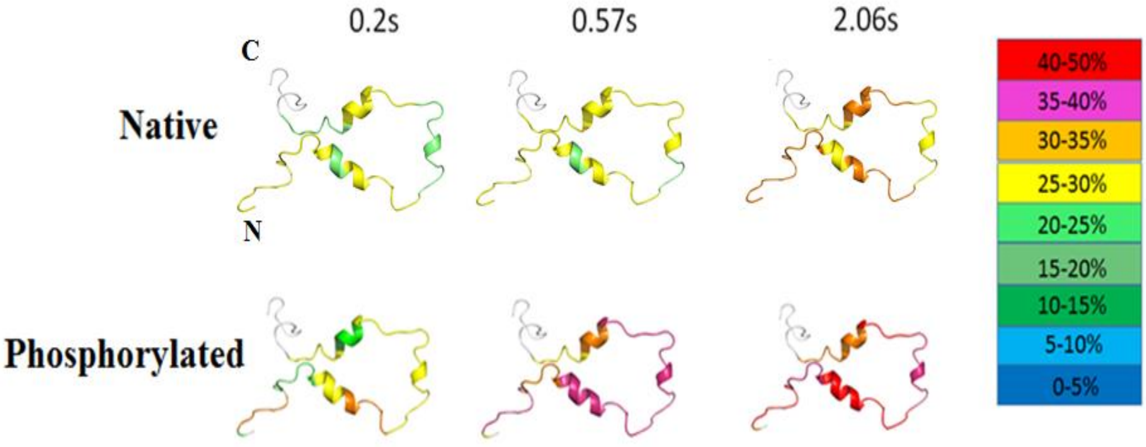


Figure 2.2. HDX data for native and phosphorylated p53 NTD. The data is mapped onto the predicted ribbon model by PHYRE2. Deuterium uptake percentage is represented by the color scheme shown on the right.

Segment-averaged protection factors can provide the relative thermodynamic stability estimation of these residual secondary structures. By dividing a peptide's intrinsic and observed rates, the PF can be calculated. The intrinsic rate is a multiexponential expression that represents a unstructured exchange rate for each residue which can be obtained from the online webserver, Sphere[52,60]. Figure 2.3A and B show the segment-averaged PF for unmodified and phosphop53 NTD respectively (unmodified p53 NTD average PF = 1.89, phosphop53 NTD average PF = 1.16). Comparatively, phosphorylated p53 NTD showed lower PFs indicating an increased conformational dynamics upon phosphorylation. For unmodified p53, the regions predicted to have residual structures which are residues 18-21, 32-43, and 53-59 showed greatest PFs (>2). These are also the regions that exhibited lowered exchange at the earliest HDX mixing time of 0.20s (Figure 2.2). For phosphop53 NTD, however, all the regions exhibited PF below 1.3 indicating a near-complete unstructured ensemble.

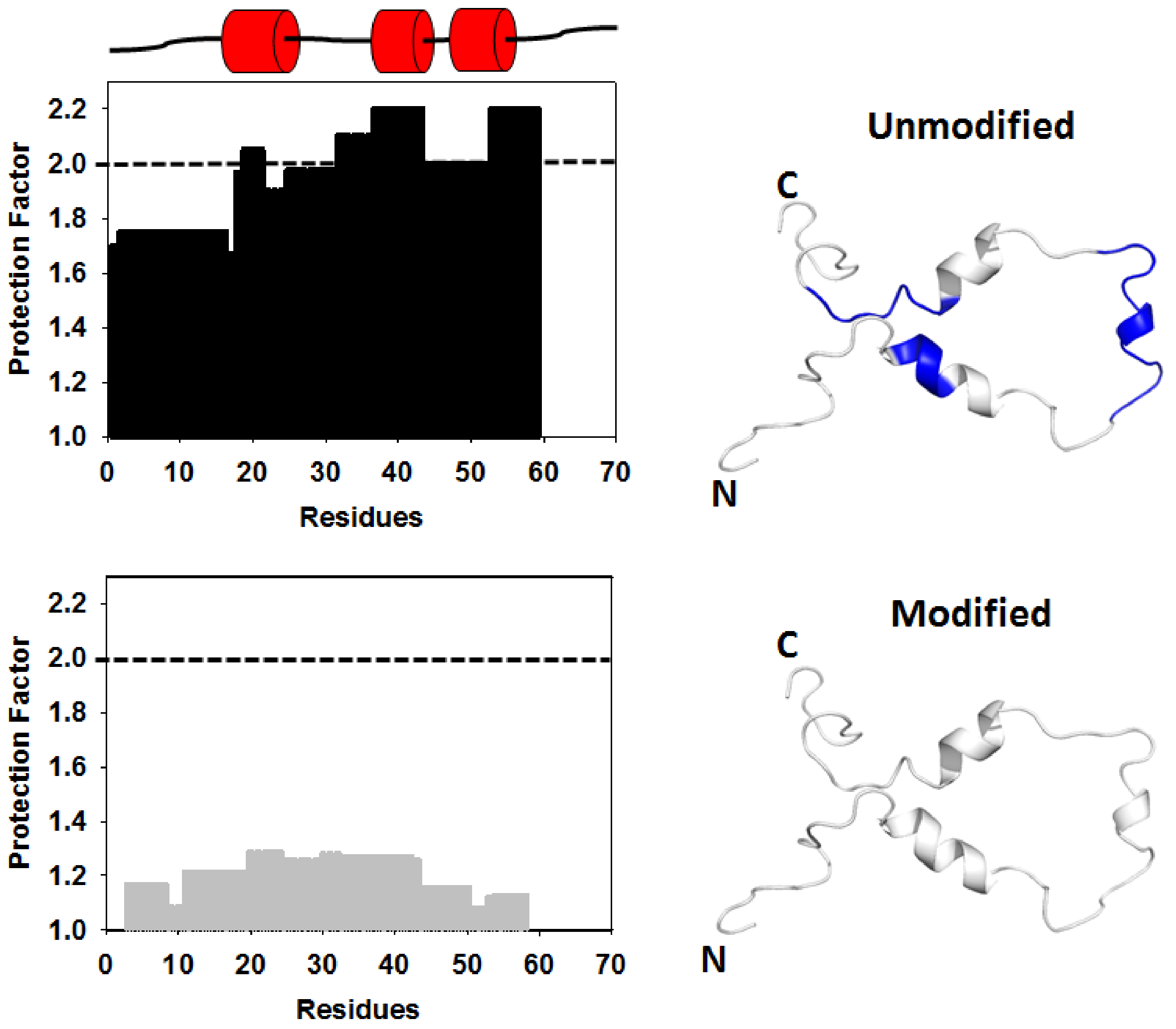


Figure 2.3. A. **Segment-averaged PF for unmodified p53 NTD.** The same data is mapped onto the 3D structure on the right. Regions with high helical propensity are shown in red cylinders aligned with the sequence domains with PF of 2 or higher containing some residual structures. **B. Segment-averaged PF for phosphop53 NTD.** The PF data is also mapped onto the 3D structure on the right.

Although there was an overall increase in deuterium uptake for modified p53 NTD, the C-terminal helix was protected as exhibited by lower deuterium uptake in the kinetic data compared to the rest of the regions throughout all of the time points which strongly suggests stabilisation of the C-terminal helix upon phosphorylation (Figure 2.4). Such local stabilization could play a critical biological role as it allows for the exposure of other linear motifs within the NTD and may facilitate binding with other interating proteins. Based on HDX and IMS data, native and phosphop53 both exhibited distinctive differences in dynamic behaviours in solution. Native p53 NTD exhibited a partially collapsed conformation with the presence of transient helical structures[172]. However, upon phosphorylation, there was an extension of the overall conformation with stabilisation of the C-terminal helix. These distinctive dynamics may also provide evidence for discrepancies in K_d values when binding to the PH domain and insights to p53 being a central hub to its transcriptional activity network.

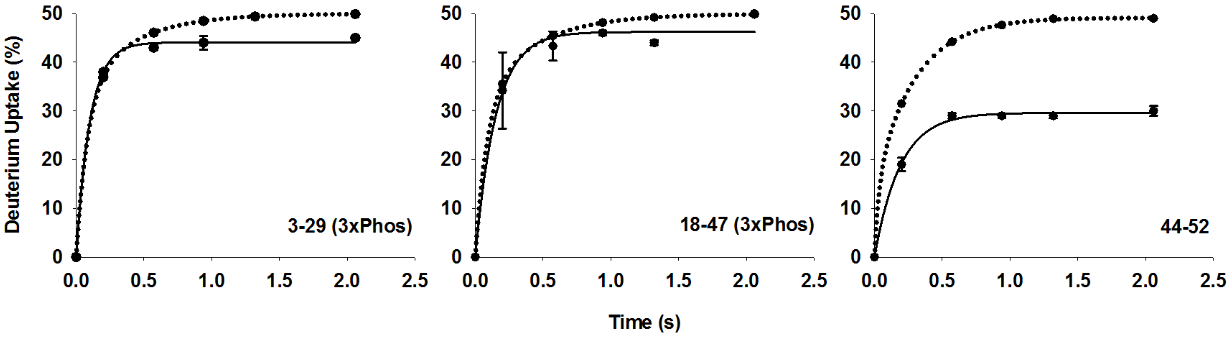


Figure 2.4. Kinetic plots for three sample peptides. 3-29 (3xPhos), 18-47 (3xPhos), and 44-52. Dashed lines represent the intrinsic rate kinetics of the labeled peptides and solid lines represent the experimental data kinetics observed at 3 or more labelling times (n = 3), error bars represent S.E.

2.3.2 p53-PH and phosphop53-PH complex

To investigate the effect of phosphorylation in p53 NTD complexation, the interaction between the (un)modified p53 NTD with PH domain was examined. p53 NTD and PH domain were incubated in equimolar concentrations to achieve complex formation and this was validated using SEC-MS and native agarose gel (Figure 2.5). During SEC-MS, MS did not detect the complex as a whole but rather as individual components suggesting dissociation of the complex during the separation process. Free p53 NTD eluted at a retention time of 53.46 minutes and free PH eluted at 51.4 minutes. Bound p53 NTD of the complex eluted at 51.2 minutes (Figure 2.5A). The fact that bound p53 NTD eluted earlier than the free p53 NTD validated complex formation despite its dissociation during the separation process in the column[175]. This was further supported by the native agarose gel electrophoresis in Figure 2.6B.

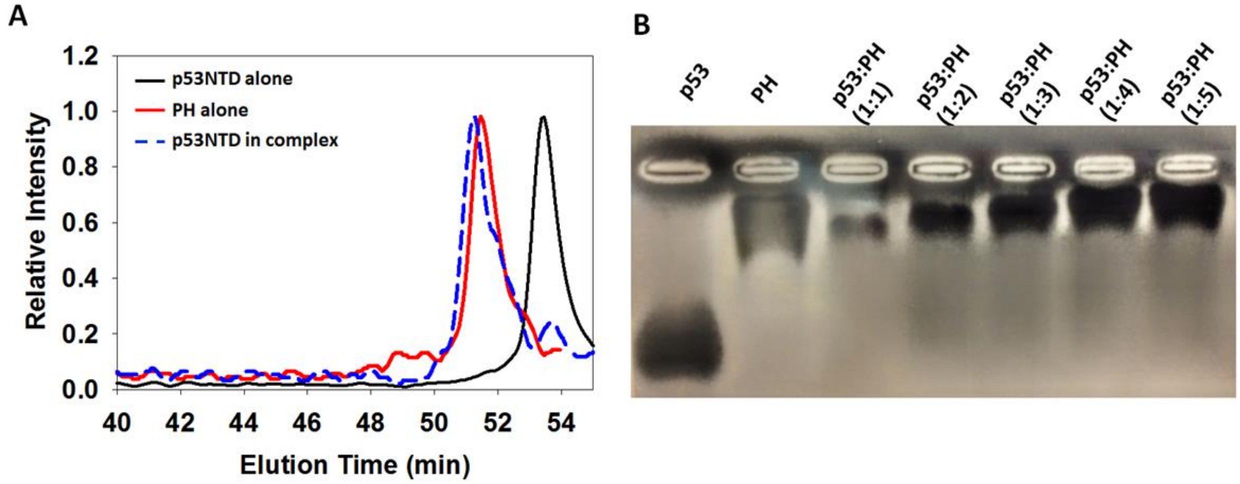


Figure 2.5. A. SEC-MS chromatogram for free p53 NTD (solid black line), free PH (solid red line), and p53 NTD in complex (dashed blue line). A BioSep S4000 column was used with an internal BSA standard (data not shown). Free p53 NTD and PH eluted at 53.46 and 51.4 minutes respectively, whereas bound p53 NTD eluted at 51.2 minutes. **B. Native agarose gel electrophoresis showing the migration of free p53 NTD and PH as well as the complexes at varying ratios**.

Next, the binding site of the complex was determined using TR-HDX and compared with literature data. Equimolar concentrations of 50μM were used to form the complex and the HDX conditions were kept constant as per native p53/phosphop53 TR-HDX. Figure 2.6 shows the segment-averaged deuterium uptake difference plot of the (un)modified complex and (un)modified p53 NTD at 0.95s. Difference plot of native complex and free p53 NTD showed that residues 18-31 and 37-43 exhibited a decrease in deuterium of 2.5 to 5% upon binding to PH. However residues 44-52 exhibited a drastic decrease of 5.5% in the C-terminus. This data is in agreement with previous literature data where residues 47-55 were determined to be the binding site using NMR spectroscopy[134]. HDX on the phosphop53-PH complex was then conducted. Interestingly, the difference plot of the modified complex and modified p53 NTD exhibited a similar HDX profile where residues 18-28, 33-43, and 48-52 exhibited a decrease in deuterium of 2.5 to 5% upon binding and residues 31-32 and 44-47 exhibited a decrease of 5.5% (Figure 2.6). The data indicate that, unlike the unbound species, the 'in-complex' structure of the p53 NTD is independent on phosphorylation and that both native p53-PH complex and phosphop53-PH complex have similar bound conformations, also suggesting that both complex-bound ground states are similar thermodynamically.

To summarize the data, phosphop53 resulted in an increase of disorder upon phosphorylation which corresponds to an elevated free energy state configuration relative to the native as shown by TR-HDX and IMS data[119]. This was further supported when the structure and dynamics of the unmodified and modified complex bound-states were similar independent of phosphorylation. This large energy difference between free native and phosphop53 contributes to tighter binding to the PH domain of the p62 subunit. Horn *et al.* illustrated a very similar case by conducting HDX studies on two variants of human growth hormones (wt-HG and v-HG)[167]. It was shown that v-HG was highly flexible and more solvent exposed compared to wt-HG. Interestingly however, v-HG exhibited a 400 fold increase in K_d relative to wt-HG when bound to HG cognate receptor. Furthermore, both variants' bound states were similar thermodynamically, similar to the case of p53-PH domain complex[167]. An elevation of the free phosphop53 ground state points toward an increase in binding affinity without lowering the energy of the phoshop53-PH complex. Furthermore, Borg *et al.* also demonstrated an increase in binding affinity of IDPs using a "polyelectrostatic" model which suggested that incorporation of multiple charges such as phosphates on a disordered protein can influence its binding affinity towards its partner through net charge effects. Specifically they showed that by altering the net charge of cyclin dependent kinase inhibitor Sic1 from -1 to -6 (increase in phosphate incorporation), it enhanced its binding affinity towards Cdc4 by 100 fold (100 to 1μM)[176].

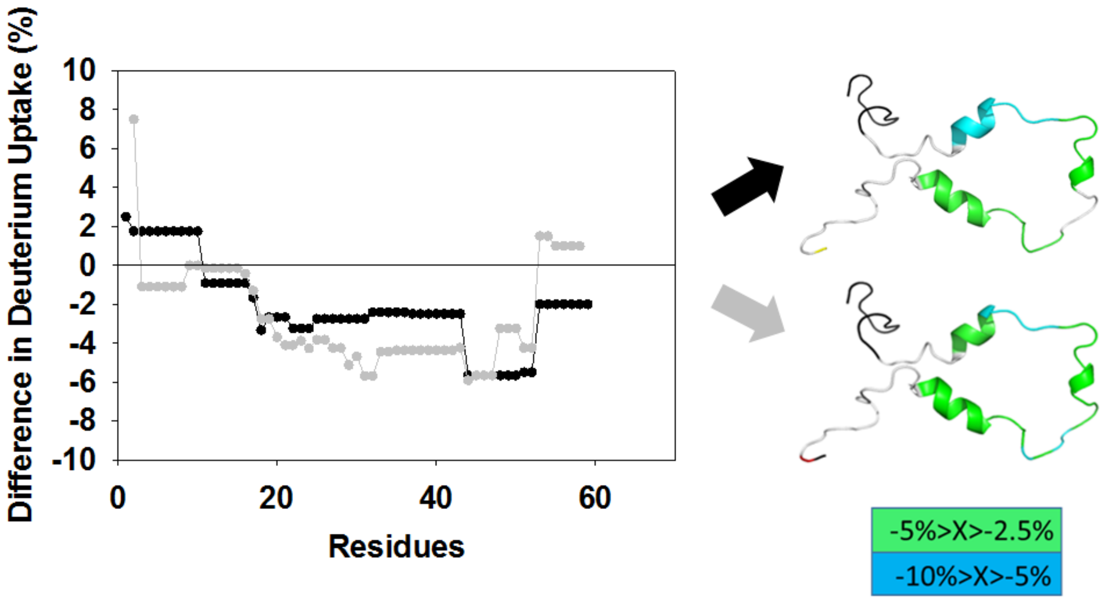


Figure 2.6. Difference HDX plot of unmodified complex with free p53 NTD (black dotted line) and phosphorylated complex with phosphorylated p53 NTD (gray dotted line) at 0.95s of mixing. Same HDX data is mapped onto the 3D structure shown on the right.

2.4 Conclusion

Despite the unstructured nature of native p53 NTD, it still retains a partially collapsed state with the presence of transient helices. Extended conformation of p53 upon phosphorylation not only exposes linear motifs allowing for other proteins to bind for the activation of cellular responses, but also enhances binding affinity to PH domain via a higher ground state energy difference. Despite this enhanced binding affinity, the bound energy states of both native/PH domain and phosphop53/PH domain complexes do not change. HDX and IMS analysis of native p53 NTD, phosphop53, native p53-PH complex and phosphop53-PH complex are powerful complementary techniques highlighting the importance of protein dynamics in molecular recognition especially for disordered proteins and/or domains.

2.5 Materials and Methods

2.5.1 Protein expression and purification. p53 NTD- pET28a construct and Chk2- pET28a plasmids were kindly provided by Dr. Yi Sheng (York University). PH domain was cloned from pDONR223 plasmid encoding general transcription factor II H, polypeptide I gene into pGEX4T2 plasmid. The proteins were bacterially expressed and further purified using Ni Sepharose 6 Fast Flow resin (p53 NTD and Chk2) and Glutathione Sepharose 4B resin (PH-GST construct). Subsequently, the proteins were concentrated using Vivaspin concentrators and buffer exchanged using Zeba desalting columns to remove high concentrations of imidazole or reduced glutathione.

2.5.2 *In vitro* p53 NTD phosphorylation. p53 NTD phosphorylation was carried out by incubating 70uM of p53 NTD with 10uM of Chk2 kinase in phosphorylation buffer (50mM Tris-HCl (pH 7.5), 10mM $MgCl_2$, 5mM DTT and 3mM ATP) for two hours. The kinase was deactivated by boiling the reaction at 80°C for 5 minutes, the precipitated kinase was then centrifuged and discarded.

2.5.3 Phosphopeptide identification and localization. The reaction was digested in the presence of pepsin agarose for 20 minutes at room temperature. Subsequently, the digested peptides were desalted and phosphopeptides were enriched using titanium dioxide resins. The phosphorylation sites were further identified using MS/MS or LC-MS/MS.

2.5.4 In vitro binding studies. To determine the binding of p53 NTD to PH domain, proteins were incubated in equimolar concentrations (50uM:50uM) overnight at 4°C. 30uL of the reaction was then loaded on BioSep-SEC-S4000 HPLC column (5uM resin) coupled to the mass spectrometer along with 5uM of internal BSA standard. Separation was performed at a flow rate of 0.2ml/min in SEC buffer (200mM ammonium acetate, pH 7.5) with a method time of 60 minutes.

2.5.5 TR-HDX chip fabrication. The microfluidic device was constructed as previously described with few modifications [72]. The rapid mixing device and acid quenching lines were connected using a Valco t-mixer externally. The output of the Valco t-mixer was then connected to the PMMA chip for infusion into the digestion chamber. Protein and deuterium were mixed at a 1:1 ratio at flow rates of 2:2 ul/min. Acid flow rate was 12 ul/min. All of the experiments were done in triplicates.

Chapter 3

Hyperphosphorylation of Intrinsically Disordered Tau Protein Induces an Amyloidogenic Shift in Its Conformational Ensemble

A version of this chapter was published in PLoS One:

- Zhu, S.; Shala, A.; Bezginov, A.; Sljoka, A.; Audette, G.; Wilson, D. J. Hyperphosphorylation of Intrinsically Disordered Tau Protein Induces an Amyloidogenic Shift in Its Conformational Ensemble. *PLoS One.* **2015**, *10* (3), e0120416.

3.1. Summary

Tau is an IDP whose primary function is to stabilise microtubules in neuronal axons maintaining the cell's morphology. In AD and other taupathies, however, tau forms insoluble amyloids known as NFTs as a result of hyperphosphorylation of tau monomers. Hence, accurate characterization of these "misfolded" intermediates upon hyperphosphorylation is crucial in elucidating the structural factors driving tau pathology as well as the rational design of potential drugs that could inhibit such formation. However, IDPs pose challenges to structural biologists as they are not amenable to conventional structural biophysical techniques. In this work, TR-HDX is implemented to study the residual structures in tau as well as the conformational shift upon tau hyperphosphorylation. By comparing the unmodified and modified ensemble, a specific conformational bias was defined that provided the structural implication for the enhancement of amyloidogenic propensity. Representative structures for both native and hyperphosphorylated tau were generated using a low computational modeling approach that is in agreement with the TR-HDX profiles.

3.2. Introduction

The physiological function of full length tau (htau40) and it's isoforms is to bind to microtubules, thus stabilizing the cytoskeletons of neuronal cells[177,178]. Tau is of great interest as it was found to be one of the hallmarks in AD (the other being dysfunction of Aβ peptide) as well as its role in other neurodegenerative disorders such progressive supranuclear palsy, Parkinson's disease and frontotemporal dementia known as tauopathies[90,179]. There have been numerous studies on Aβ and it is also becoming evident that misregulation of tau, such as erroneous PTMs, is involved in AD pathology, however, the mechanisms behind its driving force is poorly understood[90,180,181]. Nevertheless, formation of amyloids as a result of misfolded tau and Aβ seems to be a common theme in AD. For tau, hyperphosphorylation by GSK3β has emerged as a common occurrence in AD[182–184]. However, without structural information of hyperphosphorylated tau, it remains elusive whether this state is a causative agent or a protective mechanism against early toxic prefibrillary aggregates[180,182,185,186].

IDPs and IDDs have been a challenge to structural biologists as they are not amenable to the current gold standard biophysical techniques such as X-ray crystallography and NMR[88]. IDPs and IDDs often fail to crystallise for X-ray diffraction purposes and due to its lack of well-defined native structure, NMR signals often broaden out and are clustered in a narrow chemical shift window making its analysis extremely challenging[187]. Nevertheless, small IDPs can be studied by NMR for identification of transient residual structures that have implications in binding to biologically relevant partners and probing weak tertiary structures[188,189]. However, studying larger IDPs (>200 residues) can be exceedingly challenging[189].

In this work, a “bottom-up” TR-HDX workflow on a microfluidic platform is implemented to probe residual structures and conformational shift in native and hyperphosphorylated tau upon GSK3β phosphorylation to better understand tau pathology from a structural perspective. Unlike conventional HDX, TR-HDX is highly sensitive to weak hydrogen bonding and solvent accessibility making it an ideal candidate for probing transient residual structures in IDPs[91,190]. Data is presented as deuterium uptake as well as PF which is a ratio of the peptide’s observed rate and its equivalent random coil intrinsic rate on the millisecond timescale. Exposed regions would have higher deuterium exchange and a rate that is closer to its intrinsic rate (lower PF) whereas transiently protected regions will have an attenuated deuterium uptake and hence lower rate (higher PF).

3.3. Results and Discussion

3.3.1 TR-HDX on native tau

IDPs such as tau are difficult to study using conventional high resolution biophysical techniques hence alternative low to mid resolution techniques were implemented to study these proteins. For example, CD measurements established that tau lacked a stable secondary structure and radius gyration analysis by small angle X-ray scattering (SAXS) also confirmed presence of little to no tertiary structures[191]. Förster Resonance Energy Transfer (FRET) measurements suggested a conformational ensemble that resembles a “paperclip-like S” global fold where the N- and C-termini are transiently interacting with the repeat domains[141,192] . In 2009, Mukrasch and colleagues provided the most detailed structural data on native tau using NMR. Despite the challenges associated with peak assignments and low spectral dispersion, they were able to assign locations of potential residual secondary structures. With implementation of TR-HDX, we provide the first detailed view of the tertiary structural biases in the native and hyperphosphorylated tau ensemble.

The protein and D_2O were injected at flowrates of 1µl/min and 3µl/min respectively in the rapid mixer to achieve a maximum label of 75% (1:3). The labelled protein is then quickly quenched by acetic acid at pH 2.4. The average residence time of the labeled protein in the proteolytic chamber ranges between 6-8 seconds. The resulting digested peptide profiles for native and hyperphosphorylated tau produced 77.1% and 71.7% sequence coverage respectively (Figure 3.1). The full list of digested peptides for native and hyperphosphorylated tau are listed in Table B1 in Appendix B. There were a total of 75 peptides from the native protein and 71 from phosphotau. Although pepsin is known to preferentially and efficiently cleave peptidic bonds of bulky hydrophobic residues, the linear shape of the digestion profiles (Figure 3.1B and C) indicates it can non-specifically cleave charged residues as well[193]. There were 19 phosphorylation sites in the phosphotau peptide list, with five that could not be localised by MS/MS (Table B1 and Figure B1 in Appendix B).

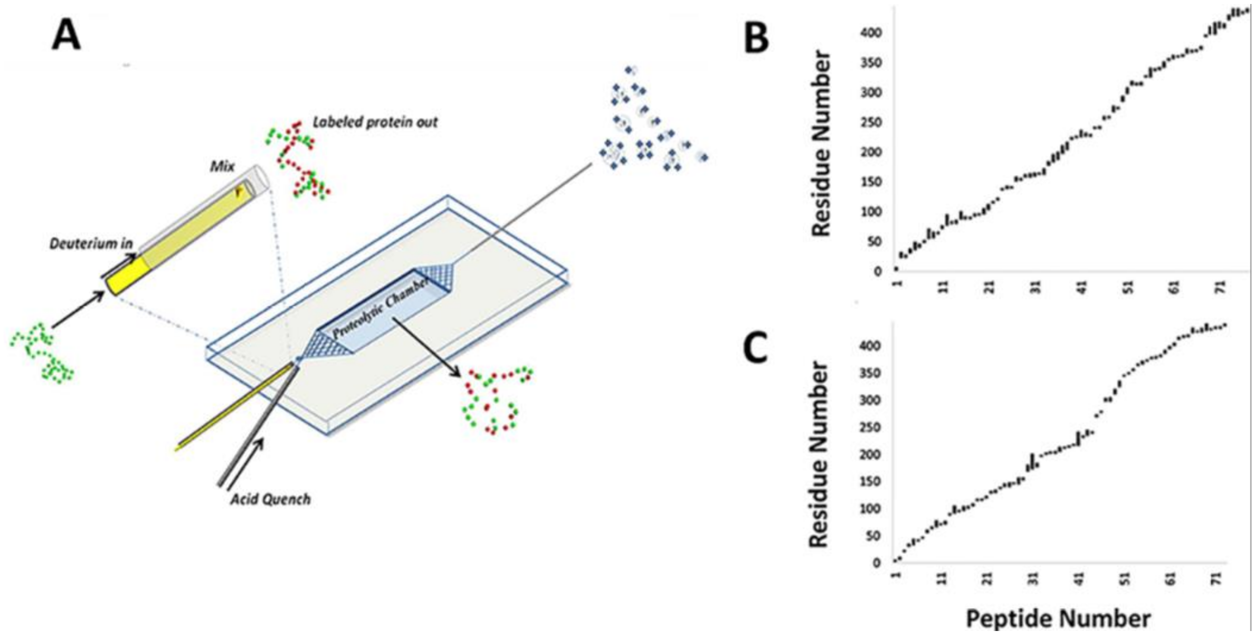


Figure 3.1. TR-HDX microfluidic chip with native and hyperphosphorylated tau digestion profiles. A. Schematic depiction of the TR-HDX microfluidic chip using “bottom-up” workflow. Inset shows the rapid mixer composed of two concentric capillaries, protein is introduced in the glass capillary (yellow) and D_2O is introduced at the gap between the glass and metal (light grey) capillaries. Both solutions mix efficiently at the notch region, labeled as mix. Different mixing times can be achieved by adjusting the position of the glass capillary. After the labeling step, the solution is then quenched at the junction of the rapid mixer and the “acid quench” line with acetic acid followed by digestion in the proteolytic chamber. The resulting peptides are then electrosprayed into the mass spectrometer. **B and C.** Digestion profile for native and hyperphosphorylated tau respectively. There were total of 75 peptides from the native and 71 peptides from hyperphosphorylated tau. Adapted from Zhu *et al*[70].

Figure 3.2 shows the process of TR-HDX data analysis for four sample peptides. Figure 3.2A shows mass spectra of deuterated peptides at the corresponding labeling times and the deuterium uptake for each is determined by fitting observed isotopic distributions to theoretical values generated by a custom built program[72]. The best fit distributions are shown as filled black dots with the associated deuterium uptake percentage. Once deuterium uptake for all time points are analysed, a kinetic plot is generated as a function of mixing times, as shown in Figure 3.2B. In a well-folded protein, the deuterium uptake kinetic plots can often exhibit “bi-phasic” behaviour corresponding to both structured and unstructured regions [91]. However, in this case, all of the kinetic plots fit to a single exponential function, which is not surprising due to the intrinsically disordered nature of tau. By dividing a peptide’s intrinsic rate, k_{int} and observed rate, k_{obs}, PF’s can be deduced as listed in Figure 3.2B. A peptide’s intrinsic rate is dependent on primary sequence and can be calculated from a multiexponential function. All the peptides and their associated PF are listed in Table B1. In conventional HDX analysis, the first two residues of the N-terminal amino acids are assumed to undergo rapid back-exchange (deuterium to hydrogen) during the proteolytic step (usually 5 minutes) and are not included in the intrinsic rate calculation. However, since the residence time in our proteolytic step is only 6-8 seconds as opposed to 5 minutes, we could incorporate the second amino acid of the N-terminus[65,72].

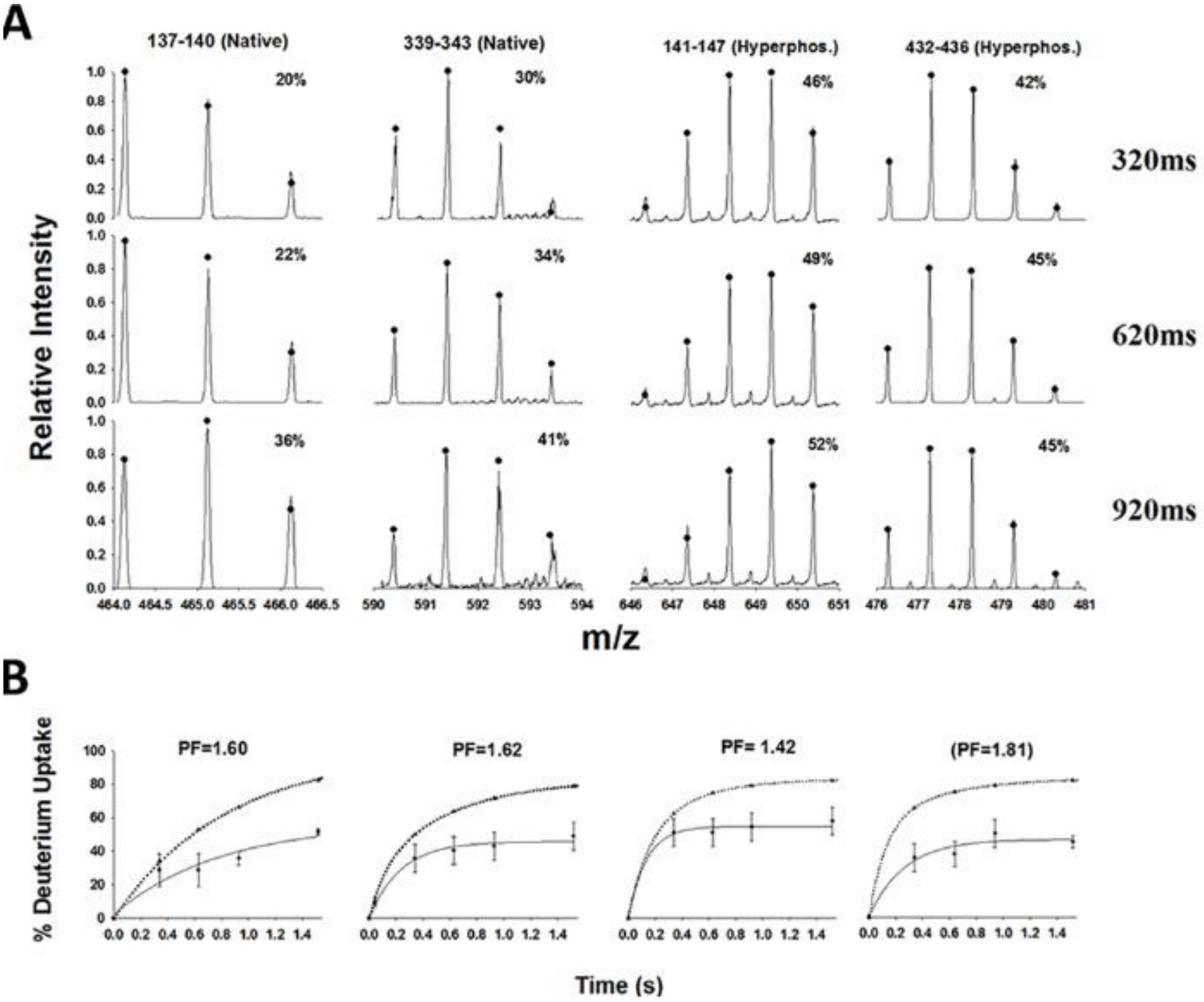


Figure 3.2. TR-HDX data analysis of native and hyperphosphorylated tau. A. Raw TR-HDX spectra for four sample peptides at three different mixing times. Each deuterated spectrum is fit to the theoretical uptake (filled dots) with corresponding percentage (inset). **B.** Kinetic plots for the respective peptides. Each are plotted to single exponential functions: solid for observed profile and dashed for theoretical profile. Adapted from Zhu *et al*[70].

Figure 3.3 shows the deuterium uptake profile for native (top bar graph) and hyperphosphorylated tau (bottom bar graph) at 1.52 seconds of mixing time lined up with the domain structure derived from the chemical shift analysis by Mukrash and colleagues[188]. Globally, native tau took up much less deuterium than hyperphosphorylated tau exhibiting weak but strong protection throughout the protein. Interestingly, only 6.7% of the covered sequence showed negligible PF that is reflective of a random coil whereas 14.6% of the sequence showed moderate PF of 10 and higher. There were no regions that exhibited a PF associated with stable secondary structures which was to be expected. Overall, PF ranged between 1 to 70 with a global average PF value of 8.1, which indicates the presence of a weak internal hydrogen bonding network and/or solvent protection consistent with weakly structured globular proteins. When tau was allowed to mix with D_2O for 10s, the protein was 90% exchanged and under 2 minutes of mixing, it was fully deuterated.

Low deuterium uptake at the N- and C-termini, proline rich and hexapeptide domains are in good agreement with the FRET measurements suggesting an S-fold conformation: contact between N-terminus to proline domain and C-terminus to repeat domains corresponding to hexapeptide motifs one and two (H1 and H2)[192]. Hexapeptide motifs are mainly comprised of hydrophobic residues that have high β-structure propensity and in NFTs, hexapeptides form the core of the fibrils[194]. Hence low deuterium uptake in H1 and H2 can be thought of as a protective mechanism against aggregation and is reflective of its non-amyloidogenic nature in native tau. Although the HDX data agrees well with global structural data provided by NMR, regions that had been predicted to contain residual structures did not show any decrease in deuterium uptake[188]. Especially in the case of the β-strands and the polyproline helices as these do not participate in an internal hydrogen bonding network. Only one of the two helices (C-terminus), exhibited drasticly low levels of deuterium uptake which is most likely involved in the S-fold conformation in its tertiary fold.

3.3.2 TR-HDX on hyperphosphorylated tau

Hyperphosphorylation was achieved by incubating native tau with GSK3β in the presence of ATP for 36 hours resulting in a heterogeneous mixture of phosphotau species (23.9% of predicted sites for all kinases and 51% for GSK3β only). Figure B1 shows the MS/MS spectra of the phosphorylated peptides. Furthermore, Figure B2 shows the denaturing SDS gel migration shift of phosphorylated tau after 5 and 30 hours. All of the localised phosphorylated sites match the S/T-XXXX-S/T consensus motif for GSK3β except for T69 and S305 which have been identified *in vitro* by Hanger and colleagues[149]. Many of the phoshorylated sites were also epitopes in NFTs aggregates such as S214, S202, S396, and S404[195,196]. Hyperphosphorylated tau was much more amyloidogenic compared to its native state, forming fibrils with NFT morphology within 48 hours whereas tau exhibited amorphous aggregates after 8 days (Figure B2).

Prior to performing TR-HDX, the solution was filtered through 0.2μM filter disks to ensure no aggregates were present. Furthermore, TEM was also performed on the filtered solution confirming the absense of soluble aggregates. The HDX profile for hyperphosphorylated tau at 1.52 seconds is shown in Figure 3.3 (bottom bar graph). Compared to the native profile, phosphotau exhibited a drastic increase in deuterium uptake, especially in the N- and C-termini and the hexapeptide II region hinting at an extended conformation. Previous low resolution structural studies have shown similar findings where there was an overall extension of the tau conformation upon hyperphosphorylation[197,198]. H1 did not show any change whereas residues 326-352 and 114-124 showed decreases in deuterium uptake. Residues 326-352 harbour repeat domains 3 and 4 that are known to bind to microtubules in its native state. Decreased deuterium at this region upon hyperphosphorylation could infer to formation of residual secondary structure or could be due to shielding from solvents due to the rearrangement of the tertiary structure. Nevertheless, this protection could be linked to tau's low affinity to microtubule binding upon hyperphosphorylation. Residues 114-124 have been shown to contain 18% of the helical structure in its native state, however, this was not reflected in our studies[188]. Interestingly, upon hyperphosphorylation this region showed a drastic decrease whereas the adjacent regions showed increases in deterium uptake, strongly suggesting a formation of the helical structure in the phosphotau ensemble. The regions corresponding to the residual structure on the C-terminus showed a similar phenomenon: increase in deuterium uptake. This could be associated with the release of the global S-fold conformation. H2 showed the greatest increase in deuterium uptake, going from being strongly protected in the native state to almost completely exposed in the hyperphosphorylated state. This exposure of H2 could potentially play a role in increased amyloidogenic propensity.

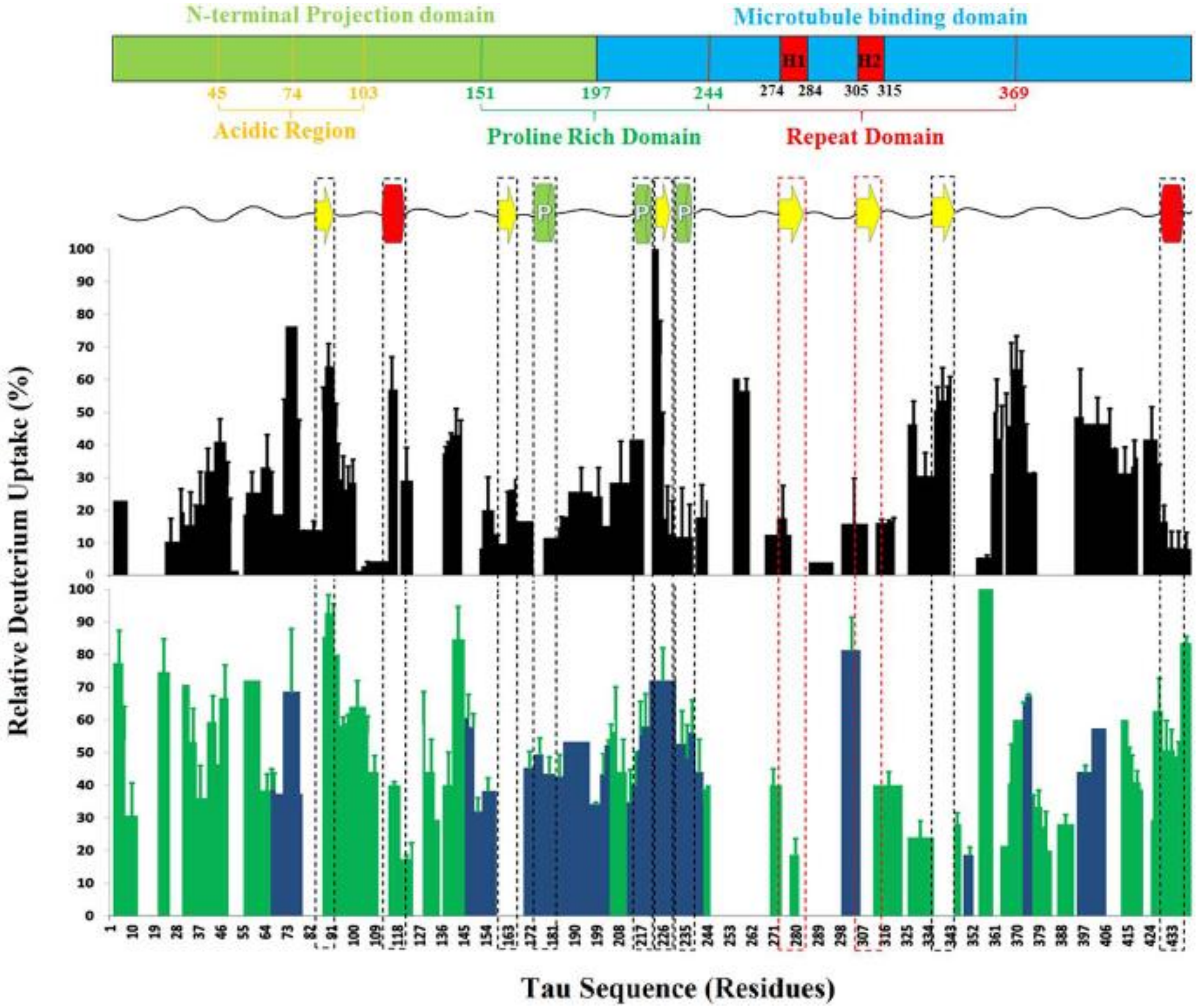


Figure 3.3. TR-HDX profile for native and hyperphosphorylated tau at 1.52 s of mixing lined up with the primary sequence of tau and the domain structure derived from NMR chemical shift index analysis[188]. Native HDX profile is shown in the top plot (black bars) and hyperphosphorylated tau HDX profile is shown in the bottom plot (green and blue bars). Blue and green bars represent the phosphorylated and non-phosphorylated peptides respectively. Both plots are lined up directly below the secondary structure map derived from NMR where yellow arrows represent segments with high β-sheet propensity, red cylinders represent regions with high α-helical propensity and green cylinders represent transient polyproline helices[188]. Adapted from Zhu *et al*[70].

3.3.3 Structure Modelling

In order to conclusively validate the above assumptions, a molecular simulation was implemented to characterize the structural ensemble of native and hyperphosphorylated tau using the HDX data. To do so, Framework Rigidity Optimized Dynamics Algorithm New (FRODAN) was used to generate theoretical ensembles for both states[199]. FRODAN can build models by exploring broad regions of the conformational space which is critial for IDPs as they also sample these wide spaces. To evaluate the structures, Volume-Area-Dihedral Angle Reporter (VADAR) was used to calculate the backbone amide solvent accessiblity of the potential models with the deuterium uptake in the HDX data[200]. The degree of the match is determined by the Pearson coefficient, R, between the models and the HDX data as shown in Figure 3.4. Figure 3.4A and B correspond to models for native and hyperphosphoryated tau respectively. IDPs can sample a wide array of different ensembles and this was reflected in the models generated; no single model correlated well with the TR-HDX data with R_{max} being ~0.3 for both states. Nevertheless, the most representative structures generated contained all the features predicted in our qualitative HDX data. In the native ensemble, the high ranking models ($R>0.275$) exhibited compact structures with the expected global S-fold and sequesteration of both hexapeptides. In the hyperphosphorylated tau ensemble, the high ranking models ($R>0.16$) exhibited extended conformations as expected with a few structural nodes corresponding to regions with relatively low deuterium uptake. Many of these models also showed exposure of H2 and protection of H1.

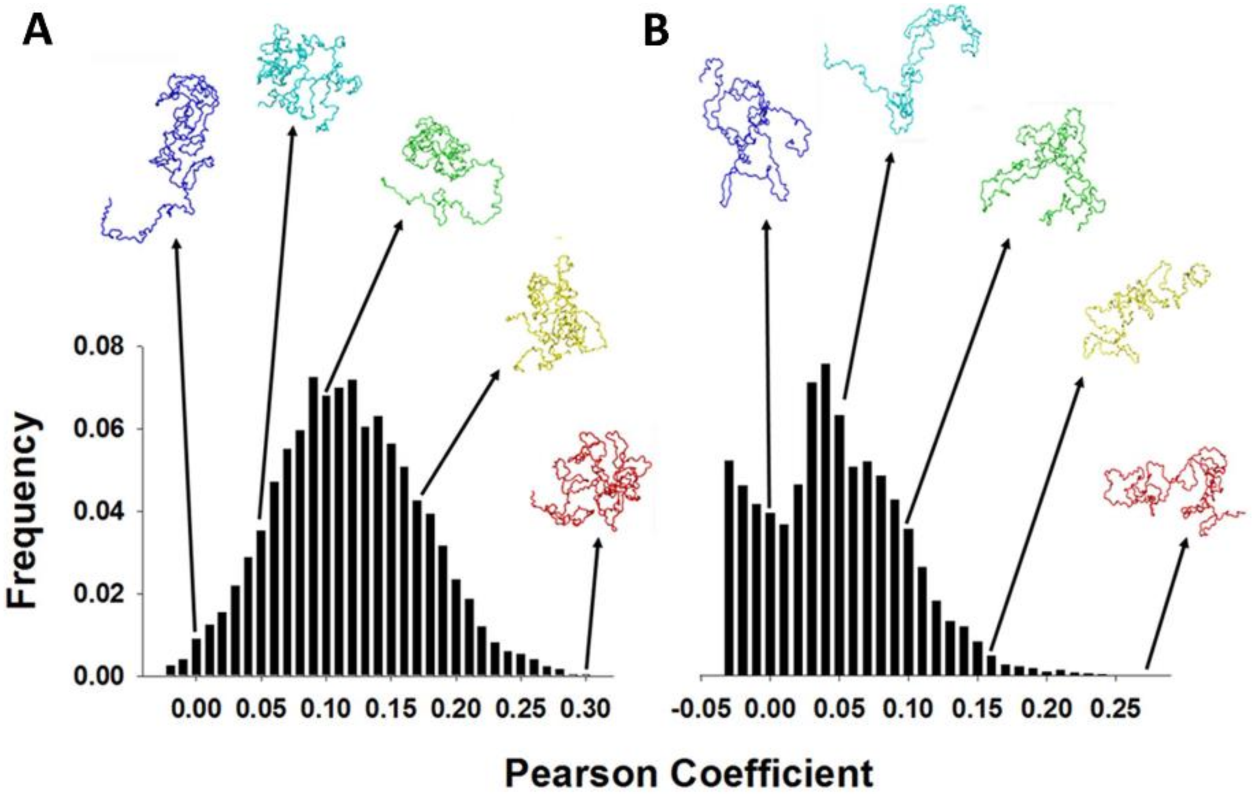


Figure 3.4. Histograms representing candidate structures of both native and hyperphosphorylated tau within the FRODAN ensembles that are in good agreement with the HDX data. A. Distribution representing native tau ensemble with R_{max}=0.3. **B.** Distribution representing hyperphosphorylated tau ensemble with R_{max}=0.29. Adapted from Zhu *et al*[70].

Following the structural ensemble generation, HDX data (Figure 3.3) was mapped onto the highest scoring structures (Figure 3.5). Although, the models are not 100% reflective of the HDX data, they still provide a detailed and accurate picture of the conformational shift underlying the increase in amyloidogenic propensity upon hyperphosphorylation using GSK3β.

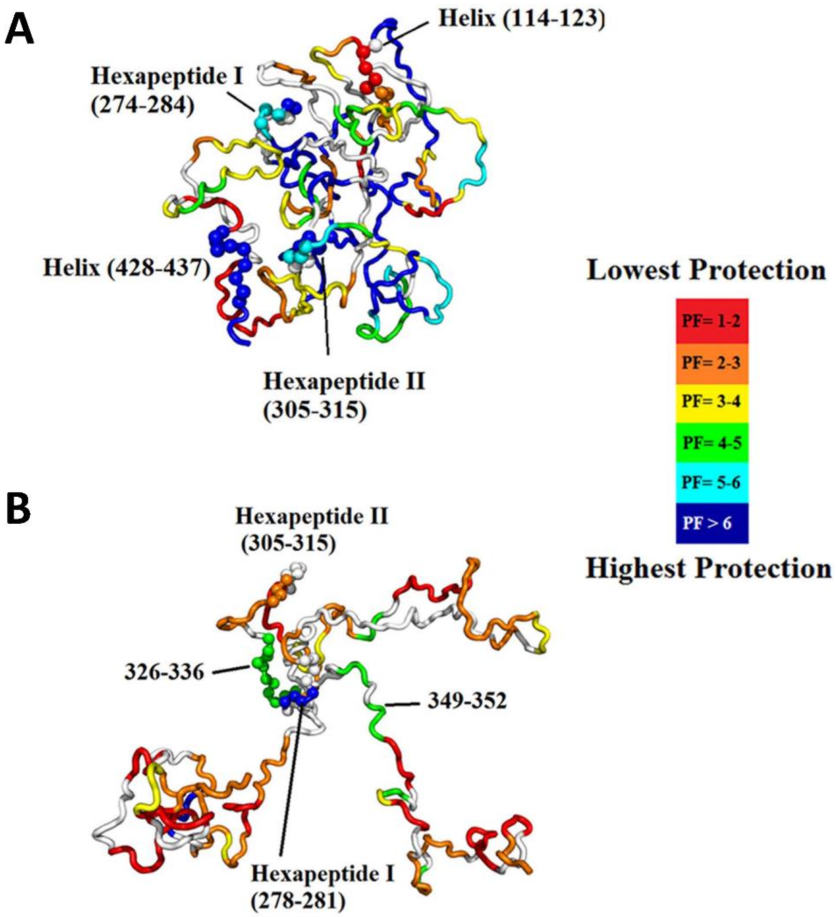


Figure 3.5. TR-HDX data mapped onto the most representative structures from native and hyperphosphorylated tau ensemble. A. PF was plotted as a rainbow scheme onto the most representative structure from the native tau ensemble where blue to red represents highest to lowest protection. Native tau exhibited a global S-fold with both hexapeptides protected in the core of the protein. **B.** PF data mapped onto the most representative structure from the hyperphosphorylated tau ensemble. Upon hyperphosphorylation, there was an overall extension of the conformation with the release of the termini, exposure of H2 and formation of residual structures between hexapeptide I and residues 326-336. Adapted from Zhu *et al*[70].

3.4. Conclusion

Many attempts have been made to study large IPDs such as tau, however due to the intrinsic nature of these proteins along with technical limitations of high-resolution structural techniques, structural biologists have resorted to using low- to medium- resolution techniques such as CD, SAXS, floresence, and FRET. In this study, a detailed view of the conformational shift of full length tau driven by GSK3β hyperphosphorylation was obtained. There has been a debate over whether hyperphosphorylation plays a causative or preventative role in AD, however, there is more evidence towards the causative role as many global structural changes can be rationalised as increasing amyloidogenic propensity. Our data demonstrated that hyperphosphorylation results in release of its native S-fold to a more extended conformation with exposure of H2 whereas H1 still remains sequestered. Exposure of H2 could act as a template for early aggregation stages. These data provide the structural implication for the causative role of hyperphosphorylated tau in AD and also illustrates the power of TR-HDX coupled to a microfluidic chip as a powerful analytical technique for studying IDPs for the characterisation of pathogenic conformations.

3.5. Materials and Methods

3.5.1 htau40 protein expression and purification. Tau was expressed and purified as previously reported with minor changes[43]. Briefly, pET-29b plasmid containing htau40 construct was transfected into *E. coli* BL21 cells for protein expression. Bacteria were grown at 37°C and was induced with isopropyl beta-d-thiogalactopyranoside (IPTG) for an additional 3 hours when

OD_{600} was between 0.5-0.8. Cells were pelleted using centrifugation at 6000xg for 30 minutes at 4°C. The pelleted cells were then resuspended in resuspension buffer: 20mM MES, 0.2mM $MgCl_2$, 5mM DTT, 1mM PMSF, pH 6.8. Resuspended cells were then sonicated on ice for 20 minutes at 15s on/30s off. The sonicated lysate was then boiled for 20 minutes to rid of other proteins apart from tau. The boiled lysate was then centrifuged at 40,000xg for an hour at 4°C to pellet cell debris and denatured proteins. The supernatant containing tau was dialysed into cation exchange loading buffer (buffer A): 20mM MES, 50mM NaCl, 1mM $MgCl_2$, 2mM DTT, 1mM PMSF, pH 6.8. Tau was then purified on an ÄKTA FPLC purifier system using a strong cation exchange column (SP sepharose FF). Unbound proteins were washed away using 5 column volumes (CV) of buffer A. Other proteins were washed away at increasing salt concentrations. Tau protein eluted between 150mM to 250mM NaCl. Purified tau was pooled and concentrated to ~500µM using Vivaspin concentrator tubes. Finally, the stock protein was aliquoted and placed at -80°C for long term storage.

3.5.2 GSK3β protein expression and purification. GSK3β-pGEX plasmids were kindly provided by Dr. Jim Woodgett. Expression was similar to htau40 with minor modifications[201]. The pelleted cells were resuspended in wash buffer: 30mM Tris-HCl, 100mM NaCl, 0.1mM EGTA, 0.1mM EDTA, 0.25mM DTT, pH 7.3. Lysates were prepared from sonication which were subsequently pelleted at 40,000xg at 4°C. The supernatant containing GSK3β was then loaded onto a GST-gravity column. Unbound proteins were washed away with 5 CV of wash buffer. GSK3β was eluted using wash buffer with 10mM reduced glutathione. The eluted protein was concentrated and buffer exchanged using wash buffer to remove the glutathione. The stock protein was then aliquoted and stored at -80°C with an additional 20% of glycerol.

3.5.3 Tau phosphorylation *in vitro*. Purified tau was incubated with GSK3β (20:1 respectively) in phosphorylation buffer containing 50mM Tris-HCl, 10mM $MgCl_2$, 5mM DTT and 3mM ATP, pH 7.5 at 30°C for 30 hours[149,202].

3.5.4 Transmission Electron Microscopy (TEM) analysis of tau fibrillary aggregates. Native and hyperphosphorylated tau were incubated for 8 days and 48 hours respectively at room temperature for formation of fibrillary aggregates. Philips EM201 Accelerating Electron Microscopy was used to acquire TEM images. 5µl of protein was blotted on a 100 mesh copper grid and were stained with 2% uranyl acetate. TEM was operated at 80kV and images were observed under magnification at a range of 30-100k.

3.5.5 Identification of tau using native MS. Concentrated tau was buffer exchanged into 50mM ammonium acetate, pH 7.0 using a 5kDa MWCO dialysis cassette. Buffer exchanged tau was further diluted with ddH_2O reaching a final concentration of 20µM and identified using the Synapt G1.

3.5.6 Fabrication of TR-HDX microfluidic chip. The kinetic mixer and TR-HDX microfluidic chip containing the proteolytic chamber were fabricated as previously described with minor modifications[72,82]. Both were fabricated using a VersaLaser via CO_2 laser ablation. A proteolytic chamber with dimensions of 28.5mm x 4.9mm x 0.4mm was etched on a poly(methyl methacrylate) (PMMA) chip. Rapid mixer and acid quenching capillaries were integrated by soldering them onto the chip at the microchannels. Pepsin agarose was activated in 1M HCl for an hour and was subsequently distributed evenly onto the proteolytic chamber. A thin silicon rubber pad was placed between the chip and a custom built clamp was used to create a tight seal. The reactor was then flushed with 10% acetic acid, pH 2.4 to remove residual 1M HCl. The setup was interfaced to a QStar Elite QToF.

3.5.7 HDX of tau and hyperphosphorylated tau. 1µl/min of the protein (100µM) was mixed with 3µl/min of D_2O in the rapid mixer, achieving mixing ratio of 1:3. Mixing times between 42ms to 12s were achieved by manually adjusting the position of the mixer. The labelled protein was quenched efficiently by acetic acid at the junction of the microchannel, and subsequently digested in the proteolytic chamber. To localise the phosphorylation sites, hyperphosphorylated tau was tryptically and/or peptically digested. The phosphopeptides were then purified and enriched using the Pierce TiO_2 Phosphopeptide Enrichment Kit followed by MS/MS.

3.5.8 HDX and kinetic analysis. The deuterium uptake of individual peptides were calculated as a percentage using a custom built FORTRAN software. Theoretical intrinsic rates of the corresponding peptides were calculated using an online web tool called Sphere[56]. The observed and intrinsic deuterium uptake at different time points were plotted to yield kinetic data. All of the deuterium uptake profiles of individual peptides were fit to single exponential functions to generate observed, k_{obs} and theoretical, k_{int} rate constants. The ratio of the rate constants, k_{int}/k_{obs} yields a PF for each peptide which is a semi-quantitative measurement that illustrates the degree of protection of a particular peptide segment in the conformational ensemble. It is unclear what the effect of phosphates has on the HDX rates hence phosphopeptides containing more than 25% contribution from backbone amides adjacent to phosphorylation sites are neglected.

3.5.9 Structure Model. PDB file of the native tau structure derived from NMR studies were kindly provided by the Zweckstetter group[188]. Using this as a template, an initial native tau structure was generated based on the Pearson coefficient of the comparison between the TR-HDX data and calculated solvent accessibility. The hyperphosphorylated tau structure was then generated using the selected native structure with the highest Pearson coefficient by forced backbone intermolecular repulsion. The resulting PDB files were then translated using MMTSB and WHATIF server was used to add the hydrogen atoms. FRODAN v1.0.1 was implemented for molecular simulation which generated ~30,000 candidates. Whenever RMSD reached 1.0Å, a PDB snapshot is created. For each snapshot, accessible surface area (ASA) was then calculated using VADAR v1.4 using backbone nitrogen atoms only. N-terminal residues exhibited high ASA values hence was excluded from the calculation. The quality of the ASA calculation was evaluated using the Pearson correlation coefficient again.

Chapter 4

Time-Resolved Hydrogen Deuterium Exchange (TR-HDX) Reveals the Structural Basis of Amyloidogenesis Inhibition of Tau Protein by Alzheimer's Drug Candidates.

A version of this chapter has not been published yet

- Zhu, S.; Deng, B.; Sweeting, B.; Taylor, M.; Reed, M.; Wilson, D. J. Time-Resolved Hydrogen Deuterium Exchange (TR-HDX) Reveals the Structural Basis of Amyloidogenesis Inhibition of Tau Protein by Alzheimer's Drug Candidates.

4.1. Summary

Many current drug developments towards AD are driven solely at targeting Aβ, however, it is becoming evident that both Aβ and tau act synergistically towards the pathology of AD. Hence, one approach would be to target both proteins simultaneously. In collaboration with Treventis (a Toronto based biotechnology company), we employed TR-HDX to investigate the influence of six potential tau drug candidates on their anti-aggregation mechanism as well as to probe their binding sites. Our findings indicated that all six drugs were bound to one of the two aggregation prone regions, hexapeptide II. Furthermore, drugs with stronger potency induced tighter core collapse of tau upon binding. With the detailed structural information provided here, we can refine and improve the potency of next generation drugs targeting AD.

4.2. Introduction

AD is a chronic neurodegenerative disease that mainly affects the elderly and poses one of the greatest health care challenges of the 21st century[203,204]. The two hallmarks of AD are the formation of extracellular amyloid plaques (misfolding of Aβ) and intracellular formation of NFTs as a result of hyperphosphorylated tau protein[145]. Thus far, there have been great strides in designing drugs to temporarily alleviate the symptoms of the disease such as cholinergic drugs[205–207]. However, there remains an urgent need for an effective therapy[204]. Much of the drug development in this field, whether its small molecule drugs or antibodies for immunotherapy, is driven by the amyloid hypothesis mainly targeting Aβ pathology. However, there is increasing evidence that Aβ and tau protein do not act in isolation but synergistically and that it would be more beneficial to target both proteins simultaneously[208,209]. For example, immunotherapies targeting Aβ alone yielded less than desired results as many showed no efficacy in clinical phase III trials despite reducing the Aβ burden, reinforcing the notion that both Aβ pathology and taupathies need to be targeted[210–212]. Tau targeted therapies are still in their infancy with many in early clinical phase trials. Immunotherapy does offer advantages such as guaranteed high doses of antibodies with few follow up visits. However, administering the right doses with minimal adverse effect can be challenging, time consuming and costly. An alternative approach would be the rational design of small molecules that can target a common epitope sites of metastable amyloidogenic intermediates such as for Aβ and tau. Although toxic intermediates from different amyloidogenic proteins can be structurally different, they all undergo significant conformational changes to form a metastable intermediate with common structural features, such as manifestation of aggregation prone regions to self-assemble into amyloids[213]. For example, transthyretin must first unfold to transitions from its native to amyloidogenic states whereas Aβ must fold to exhibit the amyloidogenic conformation[213]. Hence it may be possible to target all amyloids through these metastable intermediates.

Many efforts have been taken to accurately characterize misfolded toxic tau and Aβ intermediate structures in the hopes of designing anti-amyloid drugs to target not only AD but also other amyloidogenic neurodegenerative diseases. However due to the disordered nature and size of these proteins, many of the high resolution biophysical techniques such as X-ray crystallography and NMR have not been amenable, making rational drug design extremely difficult[187]. It is crucial to accurately determine the structure of these toxic intermediates for effective rational drug design. In chapter 3, we implemented TR-HDX on a microfluidic chip along with low computational modelling to probe the conformational shift induced by hyperphosphorylation of tau by GSK3β[70]. We were able to provide a detailed snapshot of the conformational biases that enhanced tau amyloidogenic propensity. Unlike conventional HDX platforms which mainly probe stability of hydrogen bonding networks in well-structured regions, TR-HDX is able to probe protein dynamics of weakly structured regions in the millisecond to second timescale ideal for IDP structural studies[72].

In this work, we employed TR-HDX MS to investigate the influence of six novel AD drug candidates that have been shown to inhibit tau NFT formation *in vitro* along with a negative control shown to have no effect. The data presented here provides detailed structural information of tau-drug binding as well as a mode of anti-aggregation mechanism, which can further our understanding and improve the potency of next generation drugs to inhibit potentially all amyloidogenic neurodegenerative diseases.

4.3. Results and Discussion

There is increasing evidence that Aβ and tau both play a role in AD with Aβ being the trigger and tau being the bullet in the disease pathology[145]. Drug molecules that can target both have been suggested to be more beneficial and effective. However, due to the disordered nature of tau, it has been challenging for current techniques to gain insights into its three-dimensional structural properties. Obtaining structural and dynamic information upon potential drugs binding could further refine the drug development stage to generate more potent next generation drugs. An alternative strategy, TR-HDX was implemented to probe the binding sites of six potential AD drugs as well as their modes of binding.

Native tau has been shown to have a global S-shaped conformation from both FRET and NMR measurements where N and C-termini are transiently in contact with aggregation prone regions, hexapeptides I (275-280) and hexapeptide II (306-311) such that those regions are protected from solvent exposure[188,192]. Furthermore, TR-HDX studies from the previous chapter has also validated this notion[70].

Prior to conducting TR-HDX on the test compounds, a peptide sequence library was established using the same TR-HDX experimental setup to keep the parameters consistent. The only difference was using ammonium acetate buffer instead of deuterium. The digested peptides were further confirmed by online MS/MS and offline LC-MS/MS. The sequence coverage was ~70% with the majority of missing peptides resulting from the C-terminus. Our replicated native tau HDX profile agrees very well with the previously published global native tau models whereby the majority of N-terminus peptides showed a decreased deuterium uptake (~30%)

along with the two hexapeptides, especially hexapeptide I (<10%). Most of the C-terminus peptides were not covered hence it is difficult to speculate for this region.

4.3.1 Negative control

First, TR-HDX was conducted on the negative control compound incubated with tau. Figure 4.1A demonstrates the difference in deuterium uptake plot (drug bound HDX minus native tau HDX) as a function of peptides organised in a sequential manner (N- to C-terminus) for two mixing time points: 0.36s (left) and 1.52s (right). The repeat domains of tau are highlighted in green bars and the peptides containing hexapeptide I and II are highlighted in black rectangular boxes. The criteria for significant change upon drug binding were that the differences must be at least 5% or greater and that the difference must also surpass 3σ, standard deviation (blue dotted lines). For example, at both mixing time points, peptide 328-336 is considered to have undergone a significant change as it showed an increase in deuterium uptake of more than 5% and it bypassed the 3σ mark. Figure 1B shows the same difference data mapped onto a representative three-dimensional structure of tau[70,188]. Tau repeat domains and hexapeptides are shown in stick and sphere representation respectively. The degree of change is labelled as a rainbow colour scheme as shown. Overall, the majority of the sequence covered showed no observable differences upon binding to the negative control compound, indicative of the same conformational dynamics as the native tau ensemble. Silver staining gel and IC_{50} by Enzyme-Linked ImmunoSorbent Assay (ELISA) (courtesy of Treventis) were also conducted to see if there were any aggregation inhibition effects as shown in Figure C1 in Appendix C. In the silver staining gel, there was no reduction in the level of aggregates even when the compound was spiked up to 60µM. Furthermore, the ELISA showed an IC_{50} value of ~400µM indicating a negligible inhibition effect.

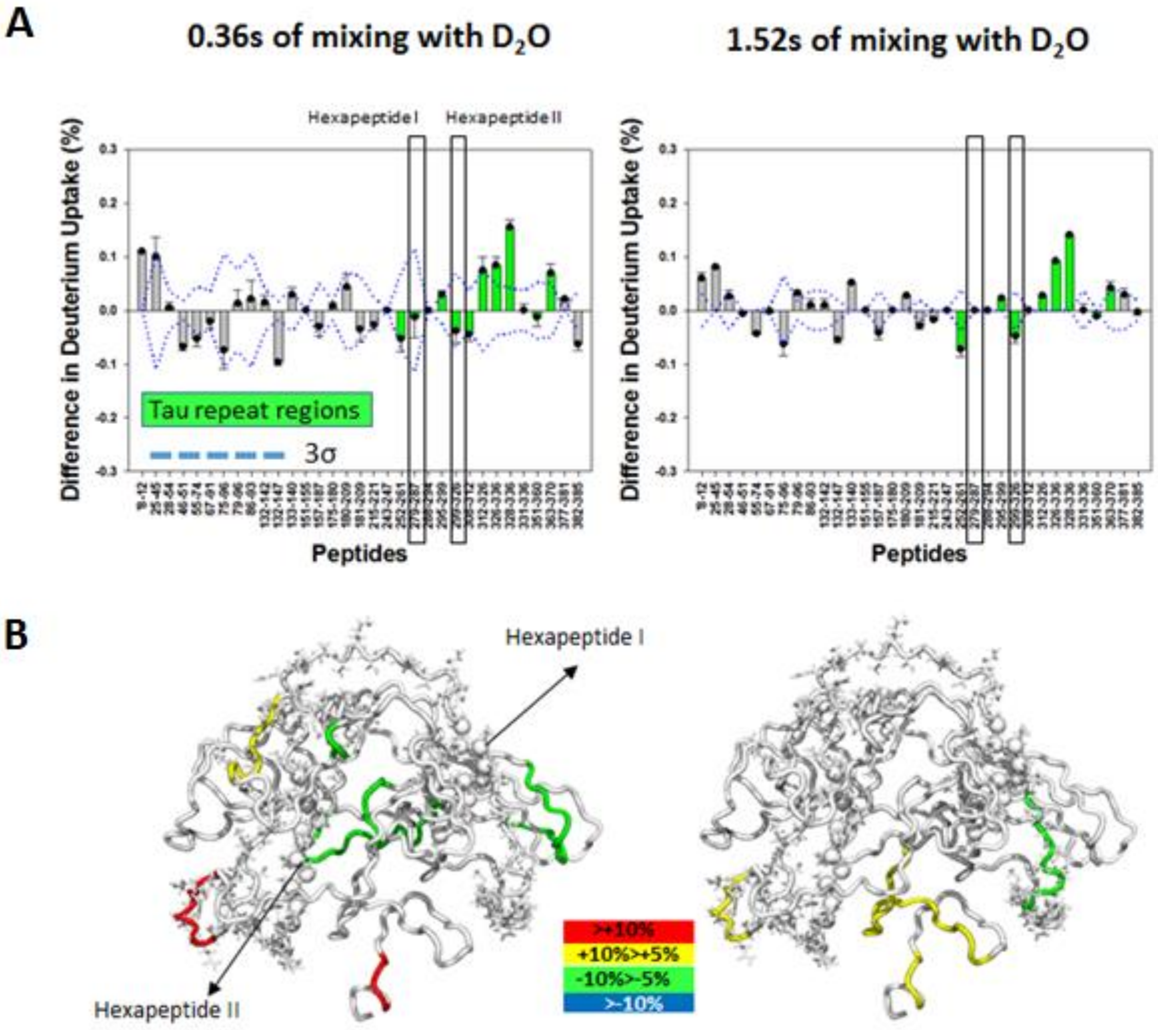


Figure 4.1 HDX difference profiles of tau upon binding to the negative control compound. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.3.2 TRV-I

Next, six compounds were tested using TR-HDX under the same experimental conditions. The data for these compounds are presented from greatest to lowest potency (TRV-I to VI). Figure 4.2 shows the data for TRV-I (most potent compound, IC_{50}~0.71uM) bound to tau. Globally, there was an overall decrease in deuterium uptake upon binding to the compound for both mixing time points. Specifically, residues 288-294 adjacent to hexapeptide I (279-287) and residues 299-326 containing hexapeptide II showed the most drastic decreases. As the mixing time increased to 1.52s, the degree of global attenuation persisted across the sequence except for residues 299-326 (-17% to -14%). Translating these data onto 3D structures (Figure 4.2B), hexapeptide II remained sequestered at both mixing time points. The core of the protein (peptides 28-54 and 157-187) at 0.36s of mixing showed a decrease in deuterium uptake of more than 10% (highlighted in blue). However, as the mixing time increased to 1.52s, the effects were less drastic as demonstrated in green. The data suggests that the compound is binding to the hexapeptide II region and subsequently induces conformational change, specifically collapse of the core. This collapse may help retain hexapeptide II in the core hence preventing or slowing down the aggregation process. A follow-up binding experiment was performed by incubating tau with synthetic peptide 306-336 and the binding was validated by native MS. IC_{50} value for this compound was ~0.71µM and its inhibition effect was also reflected in the silver stain gel (Figure C1); as the concentration was spiked, the total level of aggregates decreased demonstrating its efficacy.

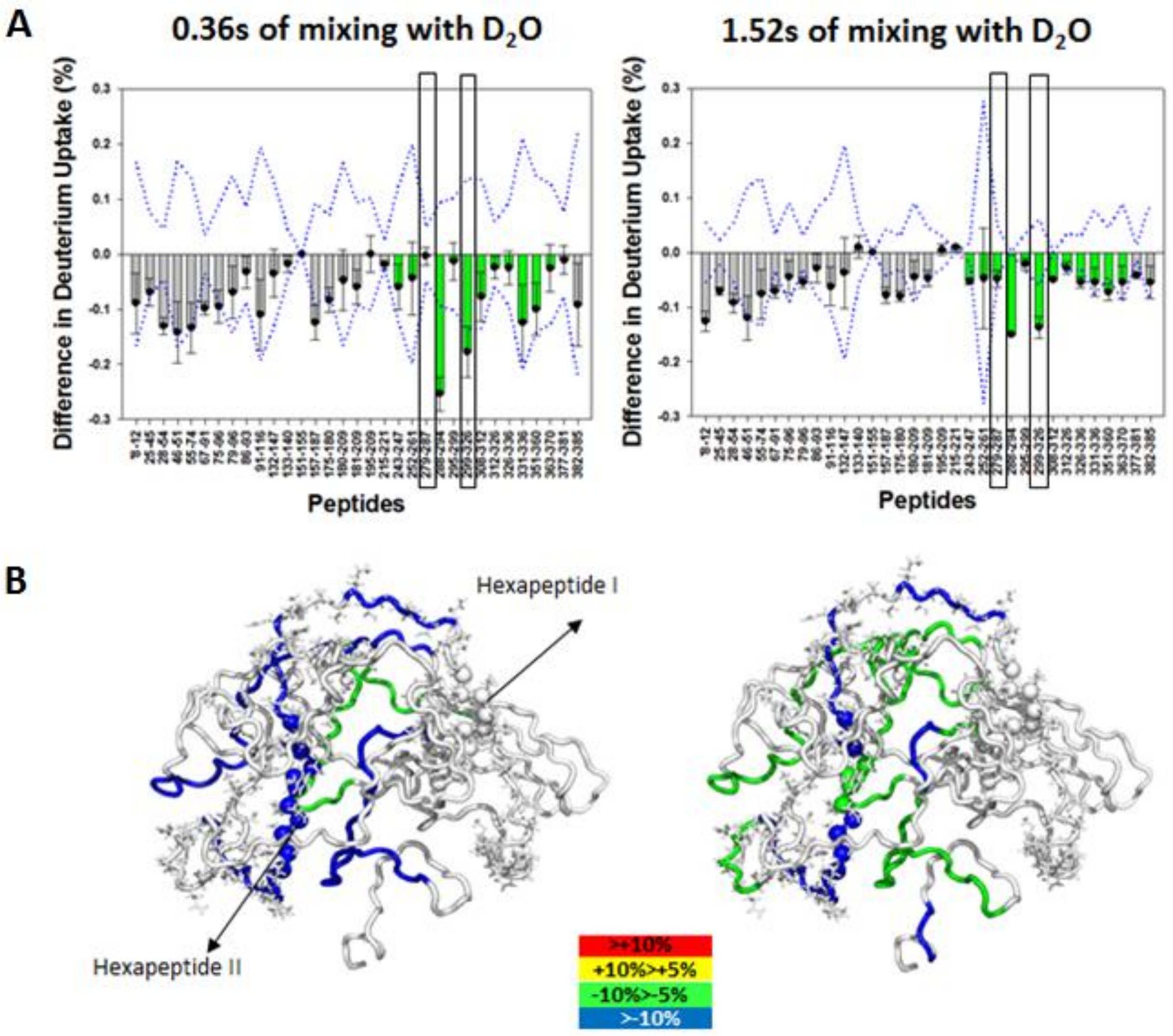


Figure 4.2 HDX difference profiles of tau upon binding to TRV-I. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.3.3 TRV-II

When tau was incubated with TRV-II (IC_{50}~2.42µM, Figure C1), resulting HDX profiles for both time points were similar to HDX profiles for TRV-I where there was an overall global decrease in deuterium uptake as shown by Figure 4.3. However, there were some noticeable differences; for example, at 0.36s of mixing with D_2O, only the core of the protein, specifically peptide residues 25-96 and 157-209 showed a decrease in deuterium uptake of more than 10%. Hexapeptides I and II were not sequestered as they were within 3σ rendering them insignificant. Interestingly, as the mixing time increased to 1.52s, two peptides 279-287 and 299-326 harbouring both hexapeptides I and II respectively, exhibited a decrease in deuterium uptake (slightly more sequestration of hexapeptide I) along with the core. One speculation of why sequestration of the hexapeptides occurs at later mixing times could be due to weak binding of the molecule which is also demonstrated by the "flattened" IC_{50} curve (Figure C1). To test whether TRV-II also binds to hexapeptide II, a similar binding experiment was conducted as TRV-I and the resulting native MS confirmed their binding. This supports the notion that TRV-II is a weak binder but more interestingly, that the allosteric effects are much greater. Although such scenarios are uncommon, Li and colleagues showed somewhat similar results when phenylalanine hydroxylase became more dynamic upon binding to its substrate, phenylalanine, in its active site along with allosteric effects[214]. The active site did not exhibit a decrease in deuterium uptake upon binding but instead showed a drastic increase[214]. In their case, they used a substrate that was rapidly turning over, here we used an inhibitor that was rapidly yet weakly binding (high k_{off}). To fully validate this hypothesis, Isothermal Titration Calorimetry (ITC) could be performed to extract the kinetic binding constants.

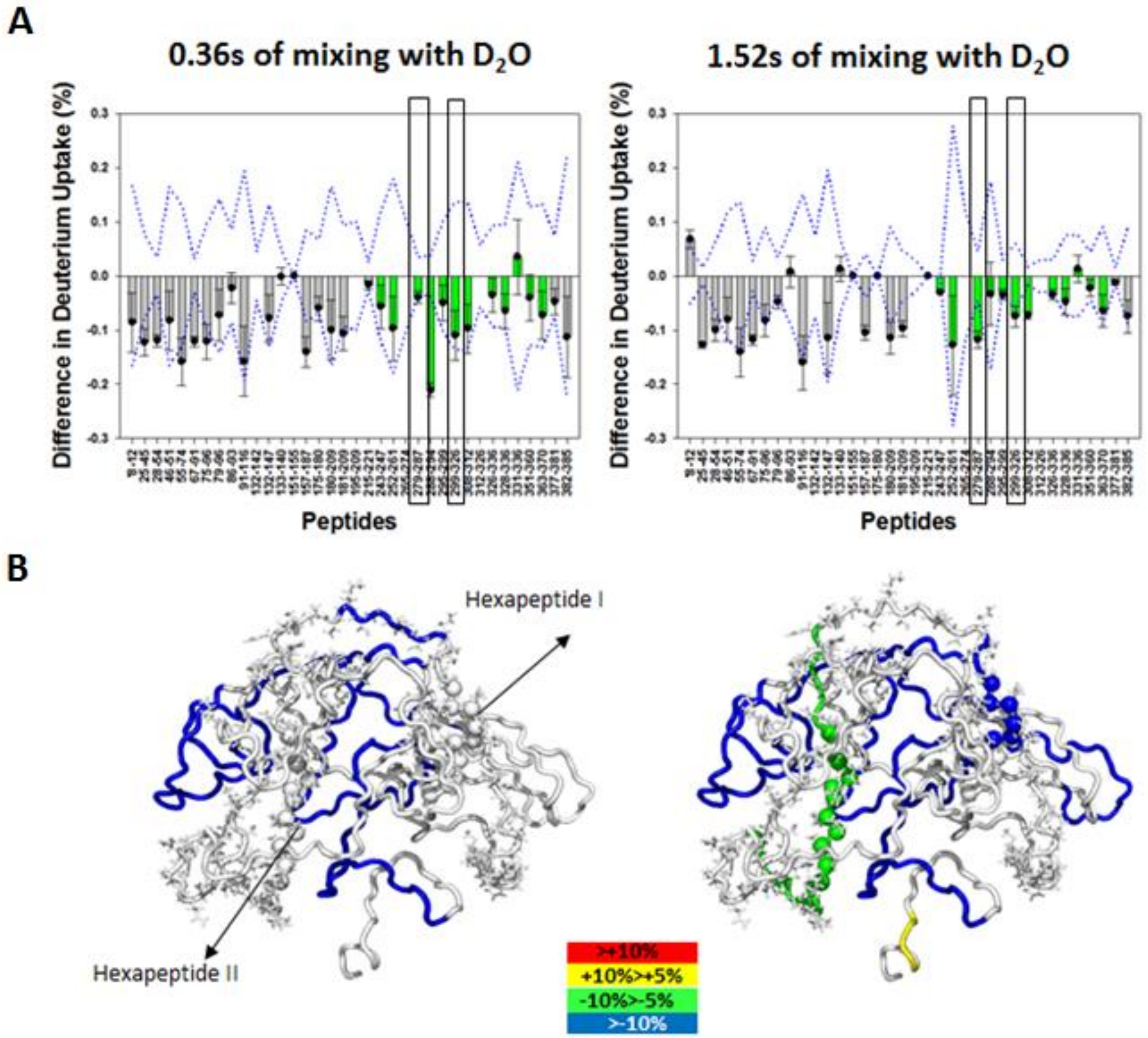


Figure 4.3 HDX difference profiles of tau upon binding to TRV-II. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.3.4 TRV-III

Incubating tau with TRV-III also yielded a global decrease in deuterium uptake as shown by Figure 4.4. Although TRV-III is less potent compared to TRV-II, their IC_{50} values are comparable: 3.07uM and 2.42uM respectively (Figure C1). However, the degree of global decrease seemed more pronounced for TRV-III as shown in Figure 4.4A, residues 28-54, 67-96, and 157-209 were sequestered to more than 10% in deuterium uptake. Furthermore, hexapeptide II (residues 299-336) showed the most drastic decrease in deuterium uptake for both time points (~ -30%). This strongly suggests that TRV-III binds to the hexapeptide II region which was subsequently verified using native MS. Hexapeptide I also exhibited a decrease in uptake but to a lesser extent and only at an earlier mixing time point, indicative of transient sequestration. As mentioned earlier, TRV-II and TRV-III's IC_{50} values are comparable however the HDX profile for TRV-III showed a greater core collapse upon binding. Furthermore, the IC_{50} curve for TRV-III is steeper than that of TRV-II's suggesting that TRV-III might be a better aggregate inhibitor than TRV-II (Figure C1). This is shown in Figure 4.4B where the degree of core collapse is almost identical to that of TRV-II, but with the hexapeptide II aggregation prone region also sequestered in the case of TRV-III.

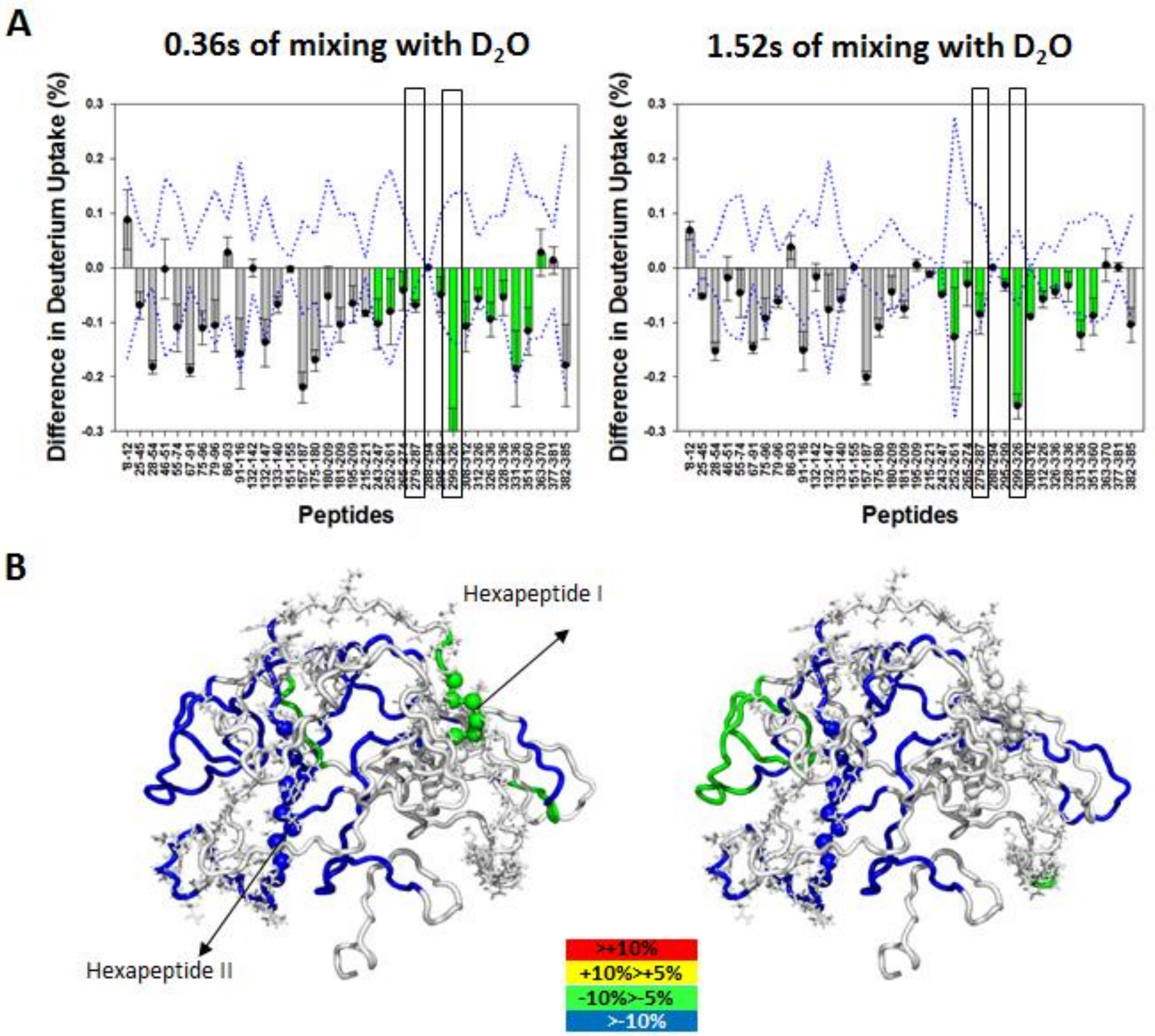


Figure 4.4 HDX difference profiles of tau upon binding to TRV-III. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.3.5 TRV-IV

TRV-IV has an IC_{50} value of 3.6μM (Figure C1) and upon binding to tau, resulted in a global decrease in deuterium uptake at both mixing time points resembling data of TRV-III bound tau (Figure 4.5A). Similarly, residues 299-326 was the peptide with the greatest decrease in deuterium uptake at both time points (Figure 4.5A and B), hinting at the drug binding site which has been validated in subsequent binding experiments. At 0.36s of mixing, the core of the protein exhibited strong protection against labeling. However, as the mixing time was extended to 1.52s, the same change in the core of the protein was attenuated to a lesser extent (5-10% as opposed to 10-15%). These data suggest that TRV-IV binds strongly to the hexapeptide II region as shown by consistently low deuterium uptake upon binding (residues 299-326). This subsequently results in strong core protection of tau rendering it to be less dynamic. However, this induced compact conformation was transient as shown by the negligible change in structure compared to the native state at 1.52s of mixing (Figure 4.5B). Previous HDX data on the three potent inhibitors (TRV-I to III) showed that the induced compact conformation was more persistent as shown by continuous attenuation even at 1.52s of mixing.

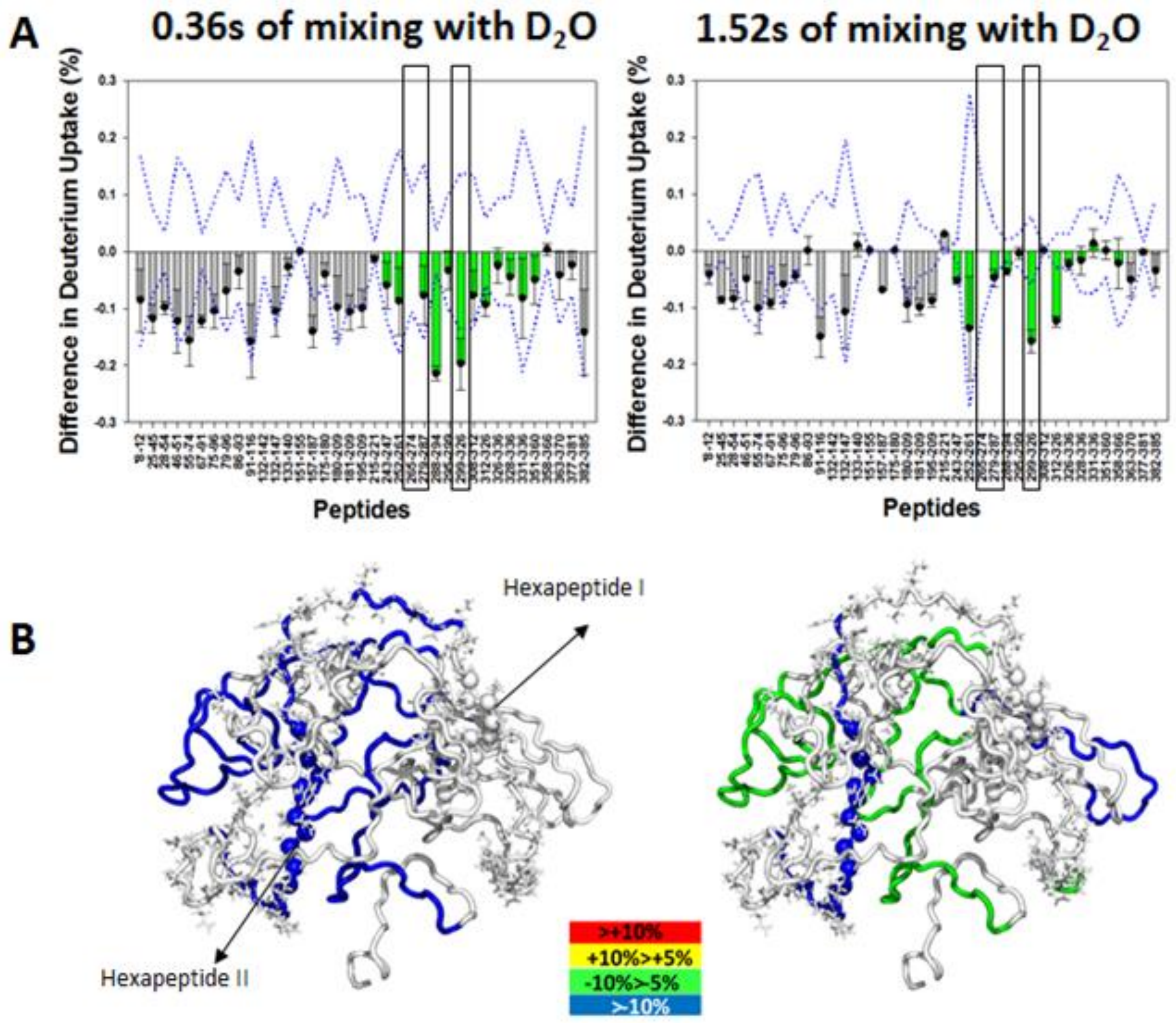


Figure 4.5 HDX difference profile of tau upon binding to TRV-IV. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.3.6 TRV-V

IC_{50} conducted on TRV-V resulted in a value of 22.2μM as shown in Figure C1 (much weaker than the previous drugs) and the resulting difference HDX plot still showed a global decrease in deuterium uptake (Figure 4.6A). Unlike the previous compounds, where the decrease was pronounced throughout the sequence, most of the peptides showed less than 10% decrease (Figure 4.6A and B). The only peptide that showed consistent protection, although not so strong at 1.52s of mixing, was 299-326 (hexapeptide II) which also happens to be the binding site for the drug inhibitor. Residues 25-54 and 67-91 of the core showed moderate protection (~ -8%) at 0.36s of mixing. When the mixing was extended to 1.52s, only residues 67-91 showed protection. This indicates that TRV-V is specific but has a high k_{off} towards hexapeptide II of tau which is a common property of IDPs[93]. Furthermore, the resulting allosteric effects are not as pronounced as the previous drugs which could explain the lower potency of TRV-V.

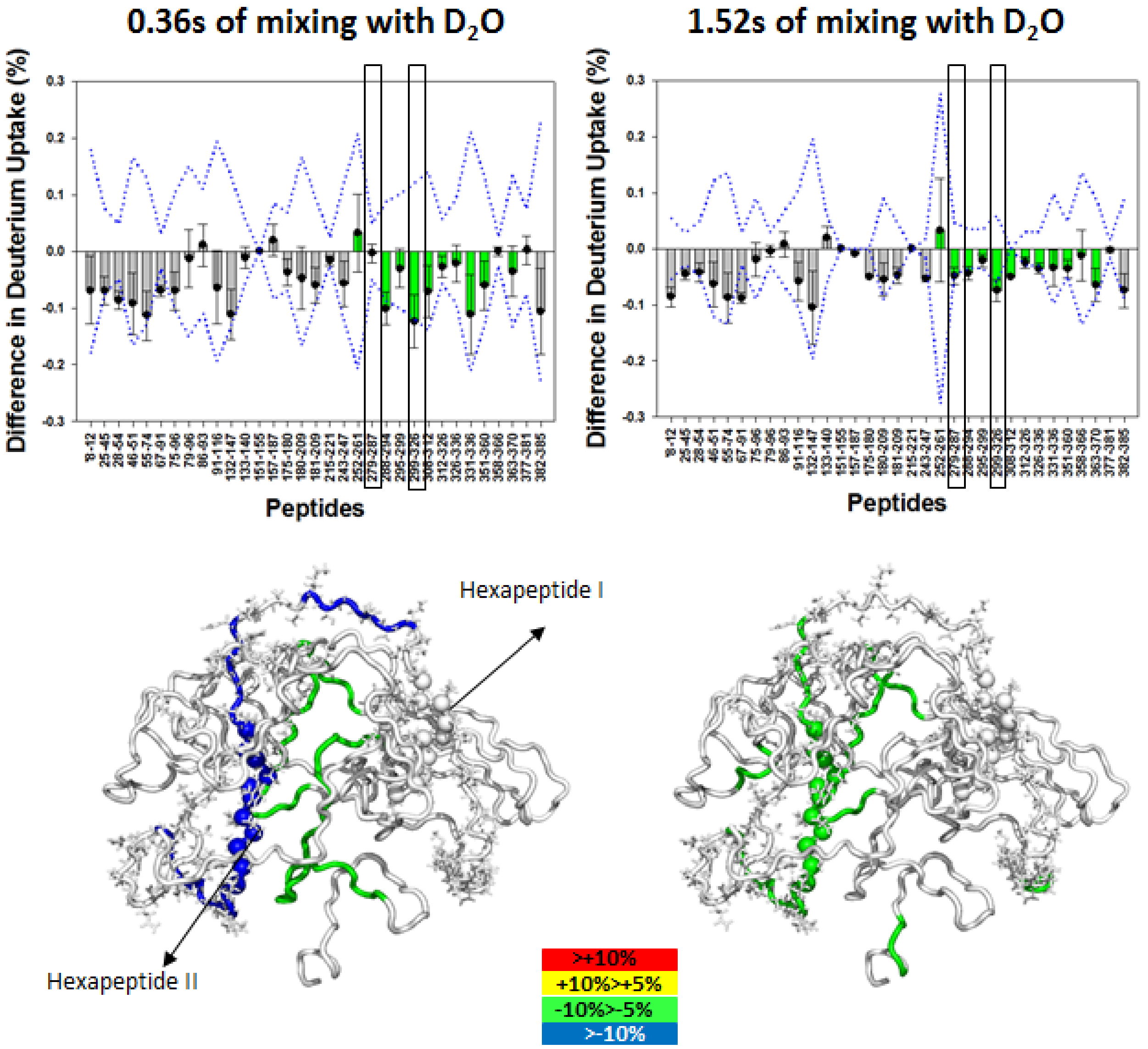


Figure 4.6 HDX difference profile of tau upon binding to TRV-V. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.3.7 TRV-VI

The last compound, TRV-VI has an IC_{50} value of 51.34μM (Figure C1). The resulting difference in HDX at both mixing times showed an increase in deuterium uptake in many peptides upon binding, however they were rendered insignificant as many failed to pass 3σ (Figure 4.7A). The only peptide that showed continual decrease was 299-326 which includes hexapeptide II. Mapping the data onto the representative 3D structure showed no observable conformational changes except for hexapeptide II (Figure 4.7B). Unsurprisingly, this region was also the binding site of the drug validated by native MS. TRV-VI only binds to hexapeptide II region with high affinity without inducing any conformational changes unlike TRV-I to V hence only the aggregation prone motif is sequestered whereas the rest of the protein remains as dynamic as native tau.

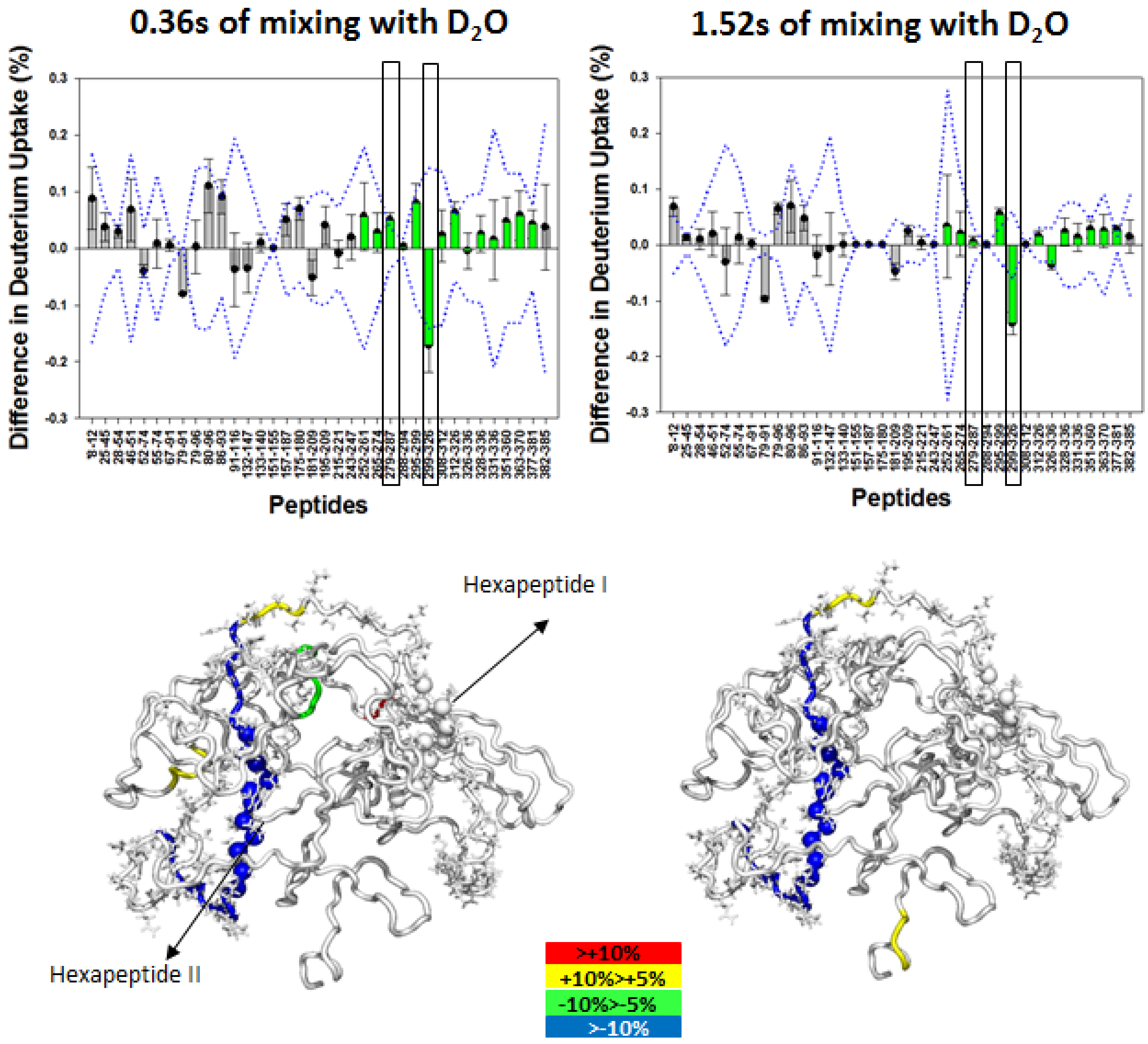


Figure 4.7 HDX difference profile of tau upon binding to TRV-VI. A. HDX difference plot at two mixing time points (0.36s and 1.52s). Tau repeat domains (R1-R4) are shown in green bars and hexapeptide regions are highlighted in black rectangular boxes. Blue dotted lines represent the 3σ threshold. **B.** HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown.

4.4. Conclusion

When tau was incubated with the negative control compound, there was no observable differences in deuterium uptake. However, when incubated with experimental compounds that showed aggregate inhibition effects (ELISA and silver staining gel experiments), there were two noticeable differences: (1) all the compounds were bound to hexapeptide II motif region (confirmed by native MS) and (2) the higher the potency of the drug, the greater the protection of the core from deuterium uptake. The least potent drug, TRV-VI, however only showed binding to the hexapeptide II region without inducing internal core collapse. The summarized data is shown in Figure 4.8.

Previously, we have shown that upon tau hyperphosphorylation using glycogen synthase kinase 3β (GSK-3β), there was a structural ensemble shift from its native global S-fold into a more extended conformation with exposure of the hexapeptide II motif to the solvent. This release has been implicated in amyloidosis by providing an aggregation template[70]. The data presented in this work agrees very well such that binding of the drugs sequesters hexapeptide II and the higher potency drugs induce a stronger internal collapse such that overall extension of the conformation is prohibited. Based on the findings, it appears that conformational change (internal collapse) upon drug binding plays a bigger factor in the mechanism of anti-tau aggregation than simply the drug binding. Perhaps the diminished dynamics along with more compact conformations limits its radius of capture and the binding of the inhibitors to hexapeptide II prevents the initial "aggregation-priming" step[93,215].

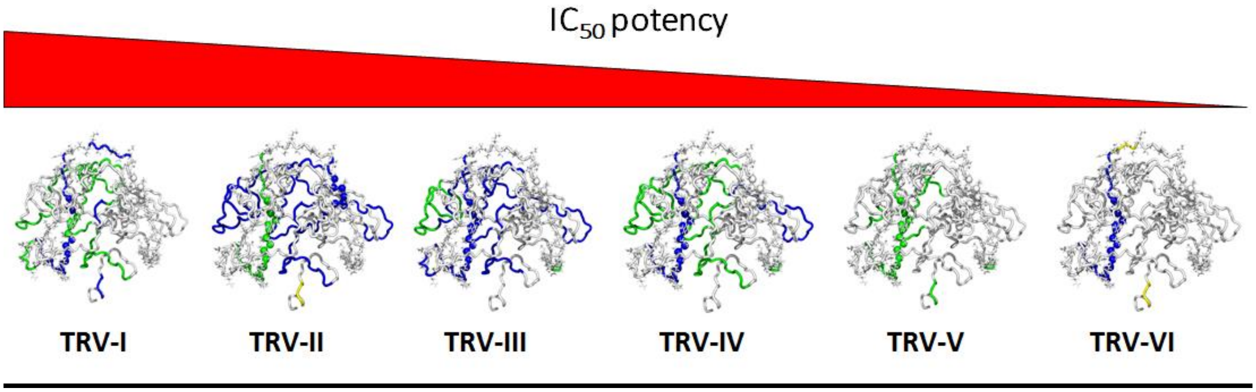


Figure 4.8 Summarized TR-HDX profiles of six drug candidates as function of IC_{50} potency. TR-HDX data plotted onto the representative structure of tau with differences highlighted in simplified rainbow scheme colours as shown. Compound names are assigned at the bottom of the corresponding structures.

4.5. Materials and Methods

4.5.1 Protein expression and purification. Tau expression and purification was carried out as previously reported[70]. Purified proteins were concentrated using Vivaspin concentrators which were subsequently aliquoted and stored at -80°C. Small molecules were kindly provided by Treventis.

4.5.2 Crosslinking protease XIII to agarose beads. Protease XIII from *Aspergillus saitoi* was crosslinked to N-hydroxysuccinimide (NHS) activated agarose beads with minor modifications; linking and quenching buffers were maintained at pH 5.0. Furthermore, crosslinking was performed overnight at 4°C.

4.5.3 TR-HDX microfluidic chip fabrication. The PMMA microfluidic device and rapid mixing module were constructed as previously described with a few modifications[70,72]. Briefly, the proteolytic chamber with a dead volume of 50ul (without the addition of crosslinked protease XIII to the agarose beads) was etched on the device using CO_2 laser ablation. The rapid mixing module and acid quench lines were connected using a Valco stainless steel T-mixer externally. The output of the Valco mixer was then connected to the proteolytic chamber using a stainless steel capillary. A thin silicone rubber pad was interfaced between the active and blank PMMA chip to create a tight seal using custom built clamps. The chip was then interfaced with a QTOF mass spectrometer. 25uM of tau was buffer exchanged into 100mM ammonium acetate buffer, pH 7.0. Subsequently desalted tau was then pre-incubated with 50uM of small molecule for one hour at room temperature prior to the HDX experiments. Protein and deuterium were mixed at a 1 (2ul/min):1 (2ul/min) ratio. Acid flow rate was 12ul/min. Labelling times ranging from 320ms to 3s were achieved by manually adjusting the position of the mixer. All the experiments were conducted in triplicates. Custom-built software was used to analyse raw TR-HDX data.

4.5.4 MS binding studies. For native MS binding assays, 25uM of tau in ammonium acetate buffer, pH 7.0 was incubated with 50uM of compound for an hour at room temperature. 10% acetonitrile was added to the solution prior to native ESI-MS.

Chapter 5

Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution

A version of this chapter was published in Journal of the American Society for Mass Spectrometry:

- Zhu, S.; Campbell, J.L.; Chernuschevich, I.; LeBlanc, J.C.Y.; Wilson, D. J. Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution. *J. Am. Soc. Mass Spectrom..* **2016**, *27* (6), 991-999.

5.1. Summary

DMS, a variant of IMS technology is routinely used to pre-filter and study small molecules from complex mixtures. However, application of DMS to study larger macromolecules such as proteins remains unexplored apart from its analogue, FAIMS. In this work, DMS combined with gas-phase DMS-HDX is employed to probe the gas-phase conformations of folded, unfolded, and partially folded proteins in solution. With optimised conditions, our findings suggest that proteins with distinct structural features in solution exhibited unique DMS-HDX profiles. Although the results provided in this study give a coarse-grain picture of the global protein features, with proper implementation this technique can be

utilized as a rapid qualitative assay to characterize structural stability of proteins in solution that could be of use in biopharmaceutical settings.

5.2. Introduction

IMS is routinely used in the field of MS to separate gas-phase ions based on their CCS areas. It is increasingly adopted in the field of structural biology for separation and analysis of protein structures, dynamics, and complex macromolecular topologies[13,216]. DMS is a variant of the IMS approach that acts as a pre-filtration mechanism of small molecules in atmospheric pressure prior to entering the mass spectrometer[26,217]. Post ionization, ions are pre-selected based on the difference between their high- and low- field mobilities defined as ΔK. To achieve such separation, an asymmetric rf voltage known as the SV is applied orthogonal to the direction of the ion flow which subsequently results in a net displacement trajectory of the ions. It is possible, however, to nullify such lateral displacement for individual ΔK by applying a specific DC offset known as CV, allowing a stable trajectory through DMS into the mass spectrometer[21,26].

One characteristic difference for this approach compared to the conventional drift tube or TWIMS is that the separation parameter, ΔK, is influenced by various factors that are intrinsic to the molecule rather than simple CCS areas. For example, interactions of the modifier (transport) gases with molecules of varying properties such as polarity and size can have a substantial effect on ΔK. This can be advantageous as various modifier gases, usually volatile chemical modifiers, can be optimised and implemented to allow for high resolution separation of species with identical CCSs[28,218–220]. A schematic cross section of DMS is shown in Figure 5.1. Despite its high separation capabilities, quantitative measurements such as determination of CCS areas of molecules are challenging due to “clustering” effects of the modifier gases as well as other factors associated with the alternating fields, especially at high field and temperature. Large proteins exhibit strong molecular dipole moments during their transit in DMS due to field induced alignment. This can be problematic as the resulting eluted CV profiles show a bimodal distribution corresponding to folded and unfolded (field induced) conformations[24,30]. However, this effect is negligible for smaller proteins (<12kDa). Thus, little effort has been made to study larger proteins using DMS since the correlation between DMS acquired gas-phase structural information was poor compared to the liquid-phase protein ensemble[22,29,34,221].

HDX is a chemical labeling reaction that is dependent on the structure of the protein whereby exposed and labile hydrogens undergo exchange with deuterium in the solvent[222]. Unlike liquid phase HDX, where the reaction is predominantly base-catalyzed, the gas-phase HDX reaction mechanism is poorly understood[223,224]. Nevertheless, it is viewed as a sensitive probe and has been combined with other mobility techniques such as TWIMS to study gas-phase protein structures[225,226]. Gas-phase HDX has been used in tandem with FAIMS (variant of DMS) to characterise different gas-phase protein conformations, with labeling occurring post FAIMS separation. Despite separation and characterisation of a multitude of different gas-phase conformations, the results were uninterpretable in the context of translating it to the solution structure[25,34]. For example, many of the co-populated conformations within the well folded protein exhibited wildly different CV values. It would have been understandable if this was the case for unfolded conformations as they can be vastly different. However, these substantially different CV values of the well-folded conformations suggest that much of the heterogeneity must have been generated during the FAIMS separation[34,227]. Thus, FAIMS-HDX has been shown to be a powerful tool to generate and study heterogeneous gas-phase conformations, but poor in correlating it to liquid phase structures.

DMS and FAIMS separation conditions are analogous hence if the labeling were performed post separation in DMS (like FAIMS), similar results would be obtained (little to no correlation with the solution phase). However, liquid-phase structural information might be retained if the labeling occurred simultaneously to the protein separation as it transits through DMS. In this report, a set of model proteins were tested using DMS-HDX to determine experimental conditions that can ultimately reflect the resulting DMS-HDX profiles to liquid-phase structures. The model proteins selected encompassed a wide range of solution structure characteristics including native cytochrome c and holo-myoglobin (well-folded), hyperphosphorylated tau (partially folded), and native tau (intrinsically disordered)[70,141]. The proteins were subjected to and investigated under different sets of instrumental conditions with the addition of HDX reagent in the modifier gas and/or throttle gas (DMS/MS interface). Ultimately, a set of conditions were determined where DMS-HDX can be used as a rapid yet reliable assay for correlating liquid-phase structure information with the resulting DMS-HDX profiles.

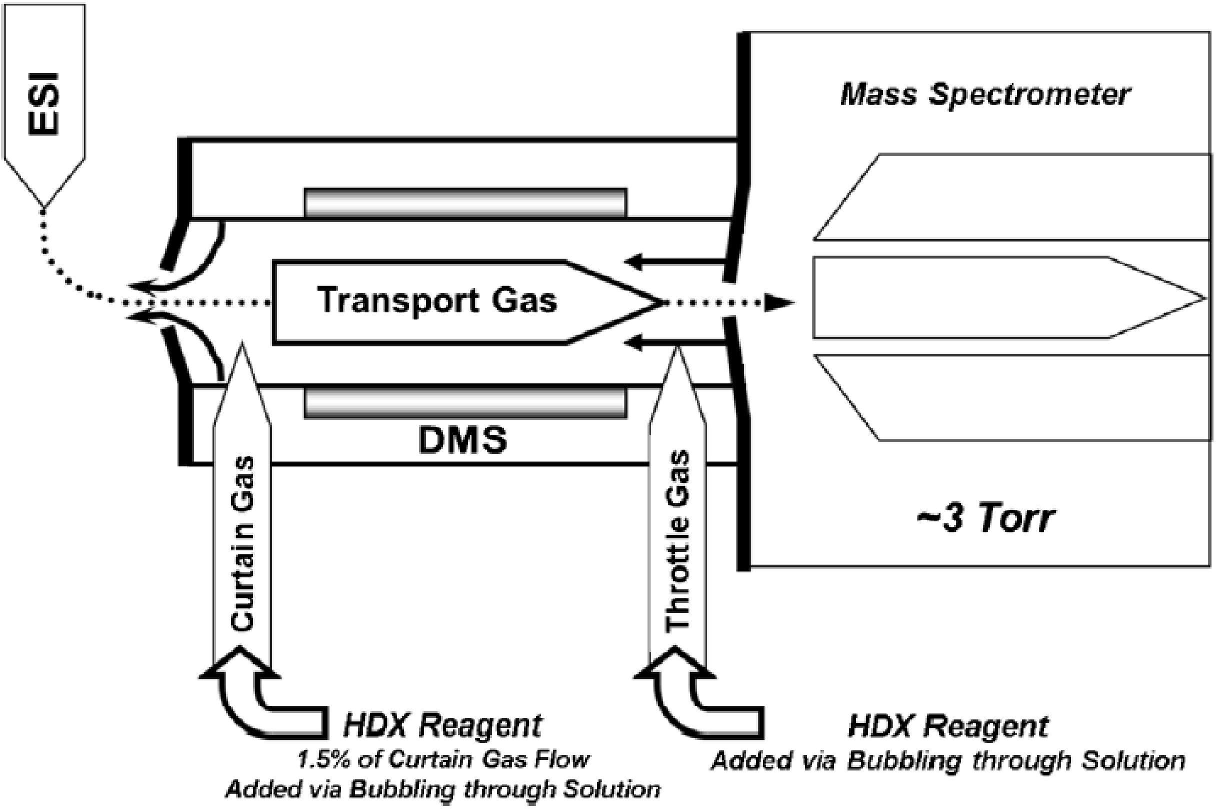


Figure 5.1. Schematic representation of the DMS cell with all the components depicted. Adapted from Shneider *et al*[26].

5.3. Results and Discussion

5.3.1 Tau and hyperphosphorylated tau DMS profiles

Previously, FAIMS has been implemented to characterise differential mobility profiles for well-defined native protein structures such as cytochrome c and myoglobin[30,34]. However, IDPs such as tau are unexplored analytes as no previous studies have been attempted on such proteins[70,188]. Tau is an IDP under native conditions hence can be studied under the same conditions as cytochrome c and myoglobin. Its unfolded nature is shown in the DMS-MS spectrum in Figure 5.2A. Wide and high charge state distribution in the mass spectrum is characteristic of IDPs. Furthermore, broad DMS profiles for each charge state are also expected due to strong pendular alignment effects for large proteins (46kDa) and this was exhibited by the bimodal distribution in the CV profile (Figure 5.2A). Initially, different modifier gases were used and optimised such that the best DMS separation was achieved for the set of proteins studied. Different volatile gases such as methanol and acetonitrile did achieve a certain degree of separation, however, it was 1:1 (ND_4OD:CH_3OD) that achieved the best result across all the proteins. Figure D1 shows the effect of different modifier gases on cytochrome c, myoglobin, native and hyperphosphorylated tau and is found in Appendix D.

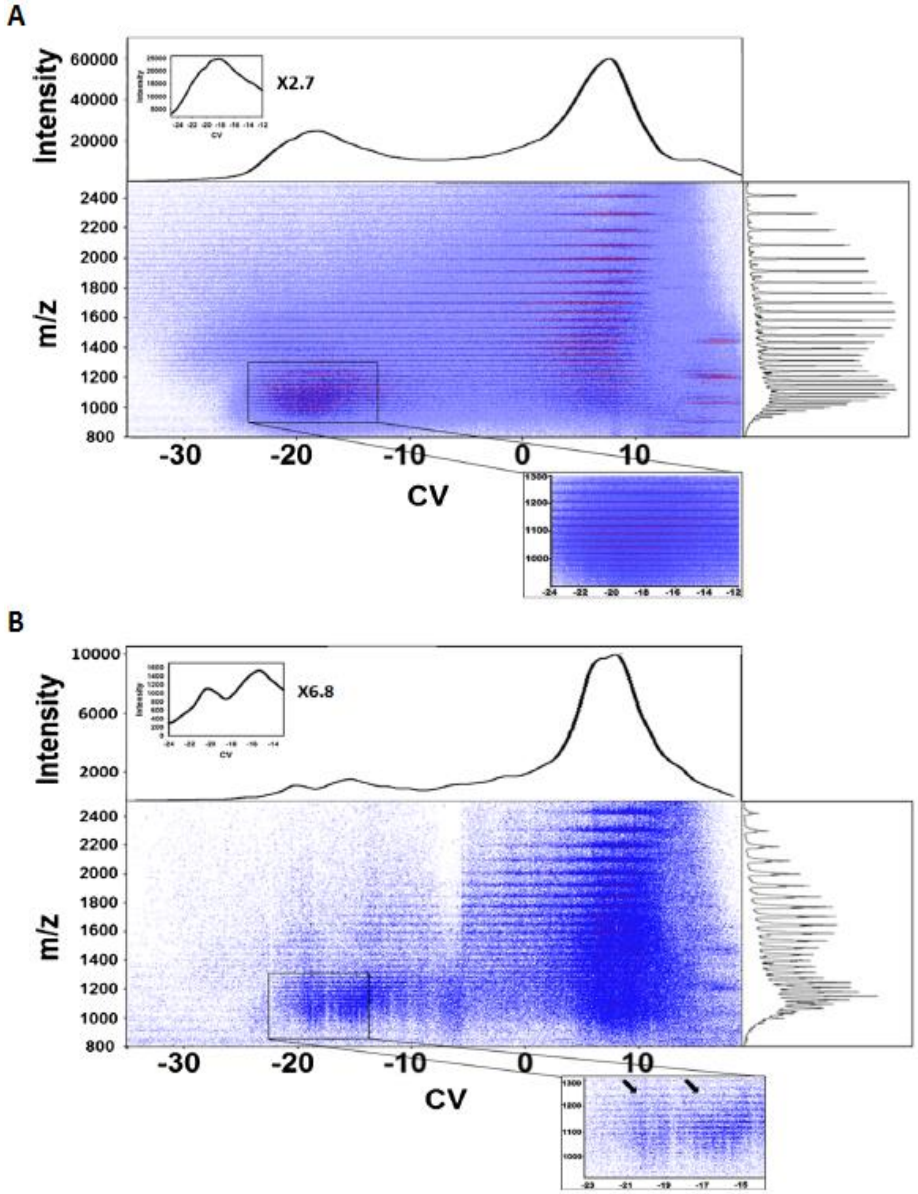


Figure 5.2 Ionogram of DMS-HDX profiles for native and hyperphosphorylated tau protein. (A) DMS-HDX profile ionogram for native tau protein. DMS profile panel is shown above the ionogram and the mass spectrum is displayed in the panel to the right. Inset shows the zoomed in high-charge and low CV regions. **(B)** DMS-HDX profile ionogram for hyperphosphorylated tau protein. DMS profile panel is shown above the ionogram and the mass spectrum is displayed in the panel to the right. Inset represents the zoomed in high-charge and low CV regions and shows peak splitting upon hyperphosphorylation, indicated by black arrows. Adapted from Zhu *et al*[228].

Tau hyperphosphorylation by GSK3β induces an overall conformational shift where it involves lengthening of the protein with establishment of several structural nodes[70]. Hence hyperphosphorylated tau represents a “partially folded” protein model in this study. The mass spectrum resembles native tau as shown by the DMS-MS profile in Figure 5.2B. However, there is one observable difference in the high charge/low CV region where the peaks were split into two. It could be due to the presence of new structural features (i.e., extended conformation with development of structural nodes) or potentially due to the consequence of differing clustering behaviour with phosphates. This open interpretation highlights the challenges in drawing meaningful conclusions from CV profiles. However, gas-phase DMS-HDX could provide more information that can address the issue of whether the separation was structure dependent or due to clustering effects.

5.3.2 DMS-HDX on well-folded proteins

In a typical planar DMS experiment, a transport gas (N_2) with/without modifier gas is incorporated at the entrance of DMS for separation[31,229]. As for resolution enhancement, a throttle gas can be introduced at the interface of DMS and the mass spectrometer orifice[26,230]. HDX reagents can be introduced at these inlets for labeling experiments. Addition through transport gas lines enables reactions to occur “in-transit” during the DMS separation whereas addition through the throttle gas line enables “pulse labeling” post DMS separation. Few HDX studies have been conducted on DMS through the pulse labeling approach (i.e., post DMS separation) as this would provide structural insights arising from the separated species. Furthermore, this approach is not complicated by the clustering effect in addition to gas-phase

HDX while separation occurs. Pulse labeling can be achieved by bubbling HDX reagents such as ND_4OD or D_2O vapors by N_2 gas through the throttle gas line.

Although pulse labeling provides important gas-phase structural information, the HDX vs CV profiles exhibited weak correlations between CV and deuterium uptake. For example, adjacent charge states of well-folded protein, cytochrome c, showed wildly different CV vs HDX profiles. A similar experiment was conducted using ND_4OD as an HDX labeling reagent in the throttle gas line and the resulting profiles for charge states +6, +7, and +8 were completely different as shown in Figure 5.3A and Figure D2 (Appendix D). These data reiterate the fact that FAIMS separation with HDX studies yield no interpretable or reliable correlation between gas-phase and liquid-phase protein conformations[25,34,227]. Hence it is clear that different conformations arising from DMS separations do not retain their solution structure information.

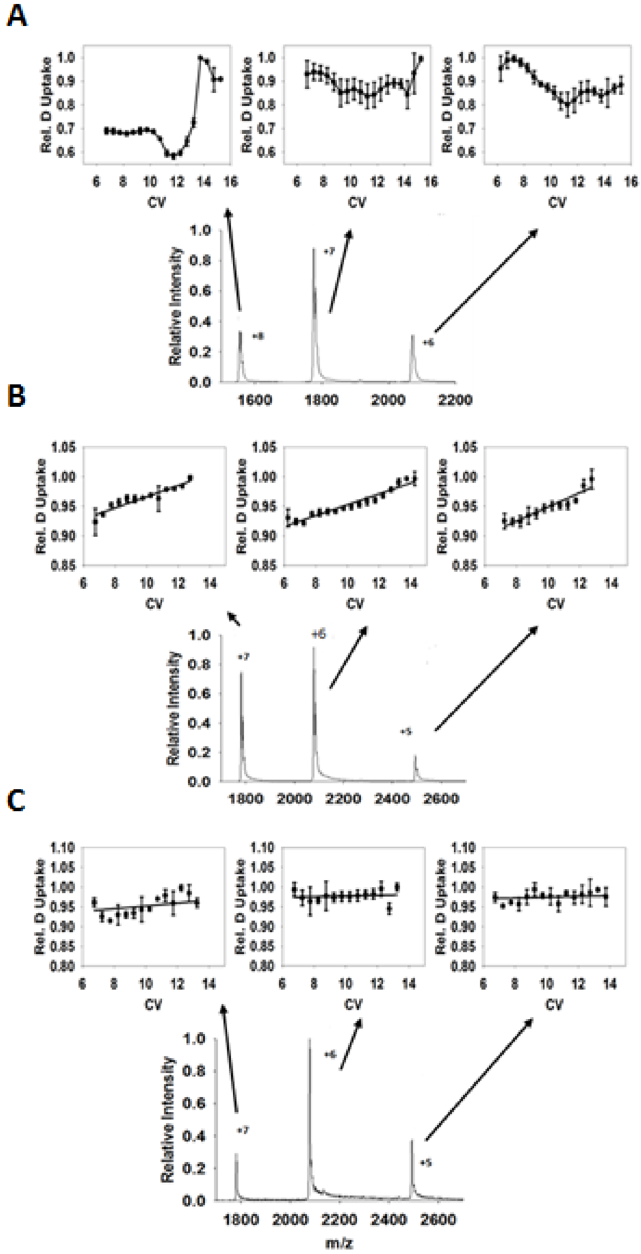


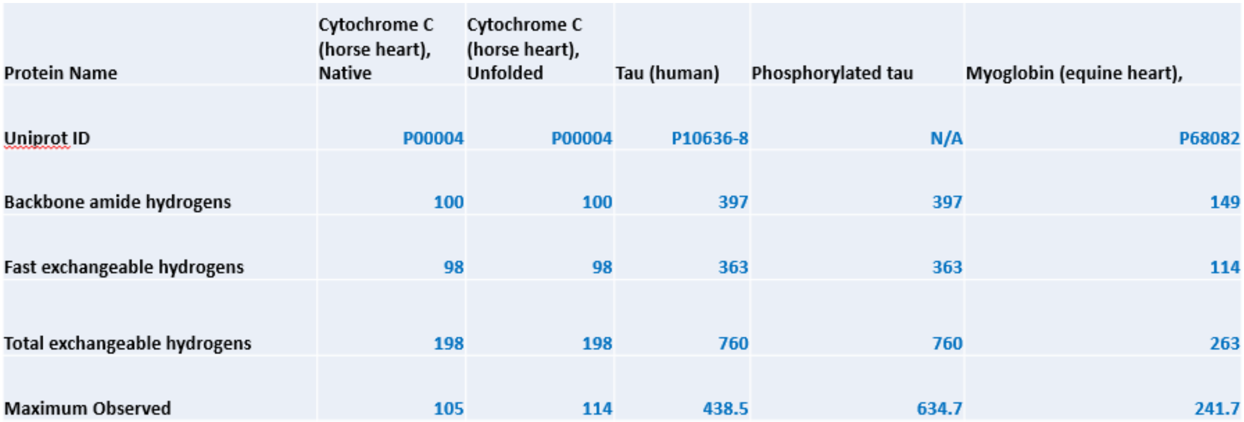
Figure 5.3 Mass spectra of native and unfolded cytochrome c with DMS-HDX profiles of corresponding charge states. (A) Native cytochrome c spectrum with deuterating agent (ND_4OD:CH_3OD) added in the throttle gas line. Inset shows the DMS-HDX profiles for individual charge states. **(B)** Native cytochrome c spectrum with ND_4OD:CH_3OD added in the transport gas line. Inset shows the DMS-HDX profiles for individual charge states. **(C)** Acid-induced unfolded cytochrome c spectrum with ND_4OD:CH_3OD added in the transport gas line. Inset shows the DMS-HDX profiles for individual charge states. Adapted from Zhu *et al* [228].

An alternative “in-transit” labeling approach was implemented where 1:1 ND_4OD:CH_3OD was added at 1.5% (v/v) along with N_2 in the transport gas line. The residence HDX labeling time during transit was ~16ms which is much longer than the sub-ms “pulse-labeling” in the throttle gas line. The resulting mass spectrum showed a decrease in overall charge states, with a substantial increase in deuterium uptake (average deuterium uptake for +6 cytochrome c was 110). This drastic increase in deuterium uptake is not surprising since ND_4OD is a much more reactive HDX reagent than D_2O. Interestingly, the HDX vs CV profile yielded a positive linear correlation for all charge states as shown in Figure 5.3B. Slope values were 1.1×10^{-2}, 1.0×10^{-2}, and 1.0×10^{-2} for the +5, +6, and +7 charge states, respectively with $R^2_{min} = 0.95$. Another interesting observation is that DMS profiles for different charge states of native cytochrome c spanned a wide CV range (between 6-16V) representing a broad set of gas-phase conformations. However, based on the relative deuterium uptake profile in Figure 5.3B, they are all within 0.9-1.0 suggesting similar conformations in this wide CV range. It is somewhat expected as native proteins even in their ground state still exhibit some degree of flexibility. Although these conformers can be HDX distinct, however, the overall ensembles are very similar. To test if this correlation was unique for well-folded proteins in solution, another experiment was conducted on native holo-myoglobin. The resulting data yielded the same positive linear slope in the HDX vs CV profiles (Figure D3). Hence planar DMS with HDX reagents introduced at the transport gas line can be used to provide a coarse-grained picture of solution phase protein structure.

5.3.3 DMS-HDX on unfolded proteins

Folded proteins generated a linear sloped HDX vs CV profile. To test if this “in-transit” labeling technique can also provide a unique profile for unfolded and partially-folded proteins, the same experiment was conducted on acid induced unfolded cytochrome c, IDP tau, and molten globular hyperphosphorylated tau. Table 5.1 provides the structural characteristics for these model proteins.

Table 5.1. Overview of the structural and HDX characteristics of model proteins used.



Protein Name	Cytochrome C (horse heart), Native	Cytochrome C (horse heart), Unfolded	Tau (human)	Phosphorylated tau	Myoglobin (equine heart),
Uniprot ID	P00004	P00004	P10636-8	N/A	P68082
Backbone amide hydrogens	100	100	397	397	149
Fast exchangeable hydrogens	98	98	363	363	114
Total exchangeable hydrogens	198	198	760	760	263
Maximum Observed	105	114	438.5	634.7	241.7

Figure 5.3C shows the profile for acid-induced unfolded cytochrome c. Typically unfolded proteins exhibit broad, high charge distributions, however, here we observed low charge states due to the strong charge stripping effect coming from highly basic ND_4OD in our 1:1 ND_4OD:CH_3OD modifier gas. This resulted in the shift of charge distribution to lower charge states with +5 being the center. Charge stripping was observed for all model proteins but was most pronounced for unfolded proteins. Although, the charge state distribution of unfolded cytochrome c was similar to native cytochrome c, the resulting DMS profiles were different as shown in Figure 5.3C. Unlike native, DMS profiles of individual charge states of unfolded

cytochrome c exhibited a flat line with slopes of $1.0 \times 10{-}3$ (+5), $1.2 \times 10{-}3$ (+6), and $3.5 \times 10{-}3$ (+7), $R^2_{min} = 0.92$. This is not surprising as unfolded species populate a wider conformational space higher in the energy folding funnel, yet still exhibit an equivalent ensemble which was also reflected in the gas-phase conformers. Although folded and unfolded cytochrome c have distinct DMS-HDX profiles, one can argue that the unfolded cytochrome c DMS profile could be due to different solvent conditions (i.e., 1% acetic acid, pH 2.4). Hence to explore this hypothesis, we conducted the same experiment on IDP tau using the same solvent conditions as native cytochrome c. The resulting DMS-HDX profile was essentially flat (Figure 5.4A), similar to that of unfolded cytochrome c. This strongly suggests that the "flat" profile is a unique characteristic to unfolded proteins, whether it's induced or native.

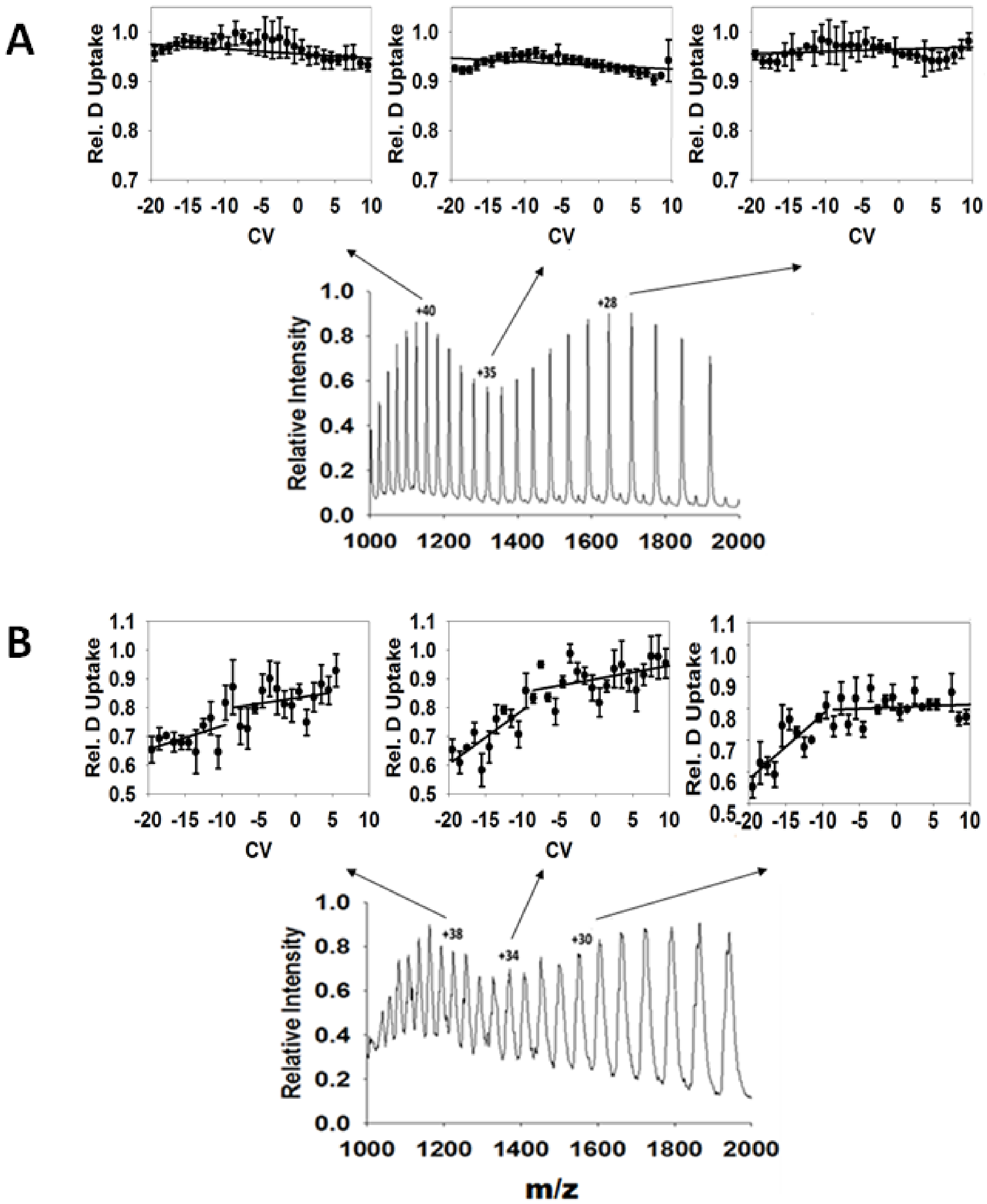


Figure 5.4 Mass spectra of native and hyperphosphorylated tau with DMS-HDX profiles of corresponding charge states. (A) Native tau spectrum with deuterating agent (ND_4OD:CH_3OD) added in the transport gas line. Inset shows the DMS-HDX profiles for individual charge states. **(B)** Partially-folded hyperphosphorylated tau spectrum with ND_4OD:CH_3OD added in the transport gas line. Inset shows the DMS-HDX profiles for individual charge states. Inflection point occurred at -10 CV. Adapted from Zhu *et al*[228].

5.3.4 DMS-HDX on partially-folded protein

Finally, to investigate the DMS-HDX effect on partially-folded protein, hyperphosphorylated tau phosphorylated by GSK3β was used. Previous solution HDX has shown that this species is more elongated yet structurally constraint, containing small pseudo-structural nodes compared to native tau[70,197]. When subjected to the same "in-transit" DMS-HDX conditions, the resulting profiles exhibited a "biphasic" appearance as shown in Figure 5.4B. One might speculate that this appearance is due to a saturation effect, however, that is not the case as the theoretical maximum for tau is 760 and the exchanged sites were 635 (well below the maximum). Thus, this "biphasic" profile strongly indicates a mixture of folded (exchange non-equivalent) and unfolded (exchange equivalent) states with exchange non-equivalent states populating the CV regions between -20 to -10V and the exchange equivalent states populating the CV regions above -10V.

Solution phase conformations may not be the same as the gas-phase conformations as a result of the DMS separation shown by the HDX pulse-labeling DMS experiments. It can, however, control the set of gas-phase conformers populated during the separation shown by the "in-transit" HDX experiment. This observation is consistent with kinetic energy loss cross sectional studies conducted by Wright and co-workers where they reported that gas-phase protein ions retain a "memory" of their solution conformations[231]. By utilising these unique DMS-HDX profiles that are characteristics of well-folded, unfolded, and partially folded proteins, one can conduct a rapid qualitative assay for studying liquid phase protein stabilities such as mutants and post-translationally modified proteins.

5.4. Conclusion

DMS, a variant of IMS technology, has been widely implemented to separate and study small molecules such as isobaric or chiral species. However, this technology is still in its infancy when it comes to studying intact proteins. In this work, planar DMS-HDX experiment was conducted to characterize and compare liquid phase conformational states of native folded, unfolded, and partially folded proteins to gas-phase counterparts. Although harsh conditions in DMS perturbs the solution structure, our results show and support the notion that gas-phase protein ions can preserve a "memory" of their liquid phase conformations under optimised experimental conditions, in this case, application of ND4OD:CH3OD (1:1) as a modifier gas in the transport gas line. Although the findings in this work provide a coarse-grained picture of global folding characteristics, a rapid qualitative assay for protein stability can be achieved with further fine-tuning of this analysis. For example, relative conformational stabilities of different mutants or post-translationally modified species can be studied using the slopes from DMS-HDX profiles.

5.5. Materials and Methods

5.5.1 Protein preparation and tau phosphorylation. Tau was purified as previously reported[43,70]. Please refer to the Materials and Methods section in Chapter 3 for details. Tau phosphorylation was carried out by incubating 100µM of tau with 500 units of GSK3β in the presence of 3mM ATP at 30°C for 30 hours. The phosphorylated tau was buffer exchanged into 50mM ammonium acetate buffer overnight and then filtered through 10kDa MWCO Vivaspin concentrator filters to rid of peptides resulting from protein degradation overnight. Native cytochrome c and myoglobin was prepared by dissolving salt-free lyophilised powder into 50mM ammonium acetate. Unfolded cytochrome c was prepared similarly with addition of 1% acetic acid reaching a final pH of 2.6. Tau/phosphotau and cytochrome c/myoglobin was diluted to 20µM and 5µM respectively prior to direct infusion into the mass spectrometer.

5.5.2 DMS and DMS-HDX experimental conditions. All the experiments were conducted on a modified quadrupole triple time-of-flight mass spectrometer equipped with a planar DMS chamber between the mass spectrometer orifice and ESI source (Figure 5.1)[26,232]. Modifier gases including HDX reagents were incorporated into the transport gas via the curtain gas line at a total flow rate of 25L/min and/or through the throttle gas line via bubbling. All samples were injected at 7µM/min and ESI voltage was maintained at 4800V with a source temperature of 37°C. Nebulizer and auxiliary gas were maintained at a constant pressure of 10 and 80 psi respectively for tau and phosphotau, and for cytochrome c and myoglobin it was 10 and 30 psi respectively. DMS temperature was maintained at 275°C. SV of 3800V and CV scanning of 0.03V increments were used throughout the experiments for all the proteins. There were two types of HDX labeling experiments: labeling during separation and labeling post separation. For gas-phase HDX post separation, D_2O or ND_4OD was bubbled in the throttle gas at the interface between the DMS and mass spectrometer orifice[230]. For labeling during the separation, 1:1 ND_4OD:CH_3OD modifier gas was bubbled at 1.5% (v/v) in the curtain gas line along with N_2 gas[230]. Analogous experiments with non-deuterated counterparts were also conducted.

5.5.3 Data analysis. PeakView v2.1 was used for data analysis. The mass spectra were averaged over 1.0 CV window for tau and phosphotau and 0.5 CV for cytochrome c and myoglobin. For example, mass spectrum for tau at CV = 6V is the average between 5.5 and 6.5V. Smoothing function was used to produce a cleaner spectrum for better analysis and reliability. Poor quality peaks were omitted from the analysis. The centroid peaks of the deuterated protein peaks were subtracted with the non-deuterated counterpart for determination of deuterium uptake. All of the experiments were conducted in triplicates.

Chapter 6

Conclusions and Future Work

6.1. Conclusions

The work presented here highlights the importance of MS as a critical analytical technique to study IDPs and IDDs which have been a challenge to structural biologists due to their inapplicability to conventional biophysical techniques such as X-ray and NMR. Chapter 1 focused on the fundamentals of the techniques implemented in this work as well as the concept of IDPs and IDDs. Fundamental techniques such as hybrid mass spectrometers, front-end separation (LC, IMS and DMS), and HDX for protein labelling are described. This is followed by a brief introduction to the fundamentals of p53 NTD and tau protein.

Chapter 2 focused on the study of how phosphorylation affects structure and conformational dynamics of the intrinsically disordered p53 N-terminal domain (p53 NTD) and how this modification drives complexation with the PH domain of p62 using TR-HDX. Our results demonstrated that phosphorylation did not affect the conformation of bound complex (p53 NTD to PH domain) as shown by similar HDX profiles for both unmodified and phosphop53 in the bound state. Interestingly, this modification increased the dynamics of free p53 NTD in the unbound state. This increase in disorder creates a wider free energy gap between the unbound and bound states which could explain the higher binding affinity when p53 was phosphorylated. This study demonstrated the power of TR-HDX and IMS as a complimentary technique that can be used to highlight the importance of protein dynamics in PTM driven molecular recognition for IDDs.

One of the hallmarks of AD apart from formation of extracellular plaques is the formation of intracellular amyloids known as NFTs as a result of tau hyperphosphorylation. Hence understanding the conformational shift upon tau hyperphosphorylation is critical in elucidating the driving force behind tau pathology. In Chapter 3, TR-HDX and computational modeling was implemented to probe this shift in conformational bias upon hyperphosphorylation by GSK3β. The HDX profile for native tau was in good agreement with previous NMR and FRET data suggesting a paperclip-like "S" shaped fold. However, upon hyperphosphorylation, there were two interesting observations. First is the release of the global "S" fold with the formation of new intra-hydrogen bonding networks within the microtubule binding region. Second is the exposure of one of the aggregation prone regions, hexapeptide II motif, providing a "sticky" template to initiate the aggregation process. The data provided a detailed structural view of the causative role of hyperphosphorylated tau in enhancing amyloidosis propensity in AD as well as highlighting the power of TR-HDX as a powerful analytical technique in probing pathogenic conformational shifts in IDPs such as tau.

Many of the current drug developments for AD targets Aβ aggregates. However, it is becoming evident that misfolded tau is the bullet in AD formation whereas Aβ acts as the trigger. Hence designing potent small molecule inhibitors that can target tau or both proteins simultaneously is more desirable. In Chapter 4, TR-HDX was implemented to investigate the structural effect on tau upon binding to six potential AD drug candidates with varying potencies from Treventis (a Toronto based biotechnology company). Our results demonstrated that all six drugs were bound to hexapeptide II motif regardless of their potency. Interestingly, however, higher potency drugs induced a tighter core collapse of tau upon binding, rendering tau less dynamic. Taken together, these results indicated that conformational change (internal collapse) upon drug binding may play a larger role in the anti-tau aggregation mechanism rather than drug binding. Perhaps the induced compactness and dynamics upon binding limits its radius of capture to aggregate with other tau monomers. Furthermore, molecules binding to hexapeptide II prevents the initial "aggregation-priming" step by preventing its exposure. These results once again highlighted TR-HDX as a powerful technique in elucidating the anti-aggregation mechanism of IDPs. With detailed structural information, drug development stages could further refine and improve the potency of next generation drugs targeting AD.

DMS is mainly used for the pre-filtering of small molecules. However, its application in studying biomacromolecules remained unexplored except for its analogue, FAIMS, which is widely used to separate and characterize gas-phase protein conformations. Despite numerous conformations generated in the gas-phase, these do not truly reflect solution phase structures. In Chapter 5, DMS coupled to gas-phase HDX (DMS-HDX) with optimized experimental conditions was employed to probe and link gas-phase protein conformations to their liquid-phase counterparts. Three different protein models were used: folded (cytochrome c and myoglobin), unfolded (acid induced cytochrome c and native tau), and partially folded (hyperphosphorylated tau). With incorporation of ND_4OD: CH_3OD (1:1) as a modifier gas, our results indicated proteins with distinct solution phase characteristics exhibited unique DMS-HDX profiles as a function of their DMS transmission. For example, unfolded proteins exhibited a "flat" constant deuterium uptake profile, folded proteins exhibited a positive linear correlated deuterium uptake profile, and partially folded protein exhibited a mixture of the two. This technique with further optimization could ultimately be used as a rapid qualitative assay to measure structural stabilities of liquid phase proteins and can be useful in biopharmaceutical settings such as studying different mutants and post-translational modified proteins.

6.2. Future Work

TR-HDX has proven to be a powerful technique to study IDPs and offers many advantages over other conventional structural techniques. TR-HDX can characterise short-lived species such as toxic intermediates, as in the case of tau, which are invisible to conventional techniques such as NMR, X-ray crystallography, and classical HDX techniques. Furthermore, weakly structured regions as well as hydrogen bonding networks of proteins in solution can be probed. Despite the attractive advantages, there are several areas that can be further improved to elevate our current TR-HDX microfluidic system.

Currently, this method operates on a continuous flow system without any separation capabilities and the resulting mass spectra contain overlapping peptide peaks which can be challenging for analysis especially for larger protein digestions. Although, front-end separation could be achieved with IMS, gas-phase proton exchange could occur depending on the MS system. For example, this is a common occurrence in the ion guide of Waters G2 instruments[233]. An alternative approach is to couple TR-HDX to a DMS system and this could potentially be more powerful than IMS in filtering out chemical noise and contaminants. For example, at a constant SV of 3800V, PEG contaminants transmit at a CV of 10V whereas peptides can transmit anywhere between 0-8 CVs.

Another caveat is the need to desalt the proteins into MS compatible buffers such as ammonium acetate prior to conducting TR-HDX experiments. The majority of cytosolic proteins are soluble at low concentrations; however, transcription factor and membrane-associated proteins easily precipitate out in these buffers. Hence it is crucial to retain proteins in buffers mimicking their physiological conditions. Wilson and Konermann have demonstrated a rapid desalting process on a microfluidic chip that is capable of removing salt concentration by one order of magnitude[234]. This can be incorporated onto our microfluidic device prior to the ESI needle to enable efficient peptide desalting. In addition, proteins containing multiple disulfide bonds, such as antibodies, are resistant to proteolytic digestion. Currently our lab is working to resolve this issue by incorporating tris(2-carboxyethyl)phosphine (TCEP) coupled to agarose beads prior to the digestion chamber.

Targeting both tau and Aβ can be more effective as an AD treatment therapy. Unsurprisingly, a subset of the six compounds from Treventis (Chapter 4) have dual inhibition effects on both tau and Aβ aggregates. Hence, it would be interesting to probe the structural effect of the inhibition mechanism on both proteins simultaneously as well as determining their epitope sites using TR-HDX. By doing so, rational design of more potent drug inhibitors could be achieved by targeting these common epitopes. Furthermore, it will also be interesting to conduct the same set of experiments in Chapter 4 using hyperphosphorylated tau as it represents one of the hallmarks of AD. Recently, a study conducted by Meier *et al.* demonstrated that hyperphosphorylated tau associates more closely to ribosomes in AD patients' brain than the controls, leading to a decrease in RNA translation[235]. This aberrant tau-ribosomal interaction impaired protein synthesis which subsequently resulted in reduction of proteins that are involved in synaptic plasticity[235]. From a structural perspective, the exact mechanism of how hyperphosphorylated tau stalls the protein synthesis remain elusive. Hence, it would be interesting to use TR-HDX to study the structural factors behind how these interactions promote neuronal damage.

The ultimate goal is to achieve high-throughput TR-HDX from a whole cell lysate, however challenges such as occurrence of isobaric peptides from different sets of proteins can hinder the HDX analysis. This can potentially be resolved by incorporating the aforementioned optimised dimensions on the microfluidic chip. For example, deuterated proteins can be reduced and digested efficiently with subsequent online desalting on chip. DMS can then pre-filter complex peptide mixtures and confidently identify peptides using data independent acquisition (DIA) methodologies. Upon knowing the protein contents in the lysate mixture, one can perform “targeted TR-HDX” on different sets of proteins. An example is to study the synergistic effect of both tau and Aβ in neuronal cells where one can monitor the conformational ensemble towards a pathogenic pathway or differences in conformations upon spiking with different inhibitor molecules or antibodies against pathogenic misfolded species.

The DMS-HDX technique can also be extended to study structural characteristics of disease relevant p53 mutants. The majority of these mutations occur in the DNA binding region resulting in the perturbation of essential DNA interactions. Furthermore, these mutants are highly unstable and unfold easily even at body temperatures [236]. Hence it would be interesting to probe the structural features of different p53 mutants using the characteristics of DMS-HDX profiles. Although the results generated provide a global picture, DMS-HDX can perform a high-throughput qualitative assay to filter out ineffective drugs during the drug development stages.

References

1. Donnan, F. G. Rays of positive electricity and their application to chemical analyses. By Sir J. J. Thomson, O. M. F.R.S. Second edition. Pp. x + 237. London: Longmans, Green and Co., 1921. Price 16s. *J. Soc. Chem. Ind.* **42,** 861–861 (1923).
2. Thomson, J. J. Bakerian Lecture: Rays of Positive Electricity. *Proc. R. Soc. Math. Phys. Eng. Sci.* **89,** 1–20 (1913).
3. Fenn, J. B. Electrospray Wings for Molecular Elephants (Nobel Lecture). *Angew. Chem. Int. Ed.* **42,** 3871–3894 (2003).
4. Tanaka, K. *et al.* Protein and polymer analyses up tom/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2,** 151–153 (1988).
5. Chang, H.-C. Ultrahigh-Mass Mass Spectrometry of Single Biomolecules and Bioparticles. *Annu. Rev. Anal. Chem.* **2,** 169–185 (2009).
6. Kebarle, P. & Verkerk, U. H. in *Electrospray and MALDI Mass Spectrometry* (ed. Cole, R. B.) 1–48 (John Wiley & Sons, Inc., 2012).
7. Banerjee, S. & Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012,** 1–40 (2012).
8. Taylor, G. I. & McEwan, A. D. The stability of a horizontal fluid interface in a vertical electric field. *J. Fluid Mech.* **22,** 1 (1965).
9. Rayleigh, Lord. XX. *On the equilibrium of liquid conducting masses charged with electricity*. *Philos. Mag. Ser. 5* **14,** 184–186 (1882).
10. Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. An introduction to quadrupole-time-of-flight mass spectrometry. *J. Mass Spectrom. JMS* **36,** 849–865 (2001).

11. Kelly, R. T., Tolmachev, A. V., Page, J. S., Tang, K. & Smith, R. D. The ion funnel: theory, implementations, and applications. *Mass Spectrom. Rev.* **29,** 294–312 (2010).

12. Giles, K. *et al.* Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom. RCM* **18,** 2401–2414 (2004).

13. Lanucara, F., Holman, S. W., Gray, C. J. & Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nat. Chem.* **6,** 281–294 (2014).

14. Stauber, J. *et al.* On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry. *J. Am. Soc. Mass Spectrom.* **21,** 338–347 (2010).

15. Shliaha, P. V., Bond, N. J., Gatto, L. & Lilley, K. S. Effects of Traveling Wave Ion Mobility Separation on Data Independent Acquisition in Proteomics Studies. *J. Proteome Res.* **12,** 2323–2339 (2013).

16. Saba, J., Bonneil, E., Pomiès, C., Eng, K. & Thibault, P. Enhanced sensitivity in proteomics experiments using FAIMS coupled with a hybrid linear ion trap/Orbitrap mass spectrometer. *J. Proteome Res.* **8,** 3355–3366 (2009).

17. McDaniel, E. W., Martin, D. W. & Barnes, W. S. Drift Tube-Mass Spectrometer for Studies of Low-Energy Ion-Molecule Reactions. *Rev. Sci. Instrum.* **33,** 2 (1962).

18. Mason, E. A. & Schamp, H. W. Mobility of gaseous lons in weak electric fields. *Ann. Phys.* **4,** 233–270 (1958).

19. Ruotolo, B. T. *et al.* Evidence for macromolecular protein rings in the absence of bulk water. *Science* **310,** 1658–1661 (2005).

20. Uetrecht, C. *et al.* Stability and shape of hepatitis B virus capsids in vacuo. *Angew. Chem. Int. Ed Engl.* **47,** 6247–6251 (2008).

21. Kolakowski, B. M. & Mester, Z. Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). *The Analyst* **132,** 842–864 (2007).

22. Purves, R. W., Barnett, D. A., Ells, B. & Guevremont, R. Elongated conformers of charge states +11 to +15 of bovine ubiquitin studied using ESI-FAIMS-MS. *J. Am. Soc. Mass Spectrom.* **12,** 894–901 (2001).

23. Purves, R. W., Barnett, D. A., Ells, B. & Guevremont, R. Gas-phase conformers of the [M?+?2H]2+ ion of bradykinin investigated by combining high-field asymmetric waveform ion mobility spectrometry, hydrogen/deuterium exchange, and energy-loss measurements. *Rapid Commun. Mass Spectrom.* **15,** 1453–1456 (2001).

24. Shvartsburg, A. A. *et al.* Field Asymmetric Waveform Ion Mobility Spectrometry Studies of Proteins: Dipole Alignment in Ion Mobility Spectrometry? *J. Phys. Chem. B* **110,** 21966–21980 (2006).

25. Shvartsburg, A. A. & Smith, R. D. Separation of Protein Conformers by Differential Ion Mobility in Hydrogen-Rich Gases. *Anal. Chem.* **85,** 6967–6973 (2013).

26. Schneider, B. B., Covey, T. R., Coy, S. L., Krylov, E. V. & Nazarov, E. G. Planar differential mobility spectrometer as a pre-filter for atmospheric pressure ionization mass spectrometry. *Int. J. Mass Spectrom.* **298,** 45–54 (2010).

27. Schneider, B. B., Covey, T. R., Coy, S. L., Krylov, E. V. & Nazarov, E. G. Control of chemical effects in the separation process of a differential mobility mass spectrometer system. *Eur. J. Mass Spectrom. Chichester Engl.* **16,** 57–71 (2010).

28. Ross, S. K., McDonald, G. & Marchant, S. The use of dopants in high field asymmetric waveform spectrometry. *The Analyst* **133,** 602–607 (2008).

29. Shvartsburg, A. A., Noskov, S. Y., Purves, R. W. & Smith, R. D. Pendular proteins in gases and new avenues for characterization of macromolecules by ion mobility spectrometry. *Proc. Natl. Acad. Sci.* **106,** 6495–6500 (2009).

30. Shvartsburg, A. A. Ultrahigh-Resolution Differential Ion Mobility Separations of Conformers for Proteins above 10 kDa: Onset of Dipole Alignment? *Anal. Chem.* **86,** 10608–10615 (2014).

31. Schneider, B. B., Covey, T. R., Coy, S. L., Krylov, E. V. & Nazarov, E. G. Chemical Effects in the Separation Process of a Differential Mobility/Mass Spectrometer System. *Anal. Chem.* **82,** 1867–1880 (2010).

32. Schneider, B. B., Covey, T. R. & Nazarov, E. G. DMS-MS separations with different transport gas modifiers. *Int. J. Ion Mobil. Spectrom.* **16,** 207–216 (2013).

33. Rand, K. D., Pringle, S. D., Morris, M. & Brown, J. M. Site-Specific Analysis of Gas-Phase Hydrogen/Deuterium Exchange of Peptides and Proteins by Electron Transfer Dissociation. *Anal. Chem.* **84,** 1931–1940 (2012).

34. Purves, R. W., Ells, B. & Guevremont, R. Combining H-D exchange and ESI-FAIMS-MS for detecting gas-phase conformers of equine cytochrome c. *Can J Chem* **83,** 1961–1968 (2005).

35. François, I., Sandra, K. & Sandra, P. Comprehensive liquid chromatography: fundamental aspects and practical considerations--a review. *Anal. Chim. Acta* **641,** 14–31 (2009).

36. Bellot, J. C. & Condoret, J. S. Liquid chromatography modelling: a review. *Process Biochem.* **26,** 363–376 (1991).

37. Cooper, W. T. in *Encyclopedia of Analytical Chemistry* (ed. Meyers, R. A.) (John Wiley & Sons, Ltd, 2006).

38. Younes, A. A., Mangelings, D. & Vander Heyden, Y. Chiral separations in normal phase liquid chromatography: Enantioselectivity of recently commercialized polysaccharide-based selectors. Part I: Enantioselectivity under generic screening conditions. *J. Pharm. Biomed. Anal.* **55,** 414–423 (2011).

39. Dorsey, J. G. & Dill, K. A. The molecular mechanism of retention in reversed-phase liquid chromatography. *Chem. Rev.* **89,** 331–346 (1989).

40. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. & Gygi, S. P. Evaluation of Multidimensional Chromatography Coupled with Tandem Mass Spectrometry (LC/LC−MS/MS) for Large-Scale Protein Analysis: The Yeast Proteome. *J. Proteome Res.* **2,** 43–50 (2003).

41. Wen, J., Arakawa, T. & Philo, J. S. Size-Exclusion Chromatography with On-Line Light-Scattering, Absorbance, and Refractive Index Detectors for Studying Proteins and Their Interactions. *Anal. Biochem.* **240,** 155–166 (1996).
42. Walton, H. F. Ion exchange and liquid column chromatography. *Anal. Chem.* **48,** 52–66 (1976).
43. Barghorn, S., Biernat, J. & Mandelkow, E. Purification of recombinant tau protein and preparation of Alzheimer-paired helical filaments in vitro. *Methods Mol. Biol. Clifton NJ* **299,** 35–51 (2005).
44. Pfaunmiller, E. L., Paulemond, M. L., Dupper, C. M. & Hage, D. S. Affinity monolith chromatography: a review of principles and recent analytical applications. *Anal. Bioanal. Chem.* **405,** 2133–2145 (2013).
45. Glish, G. L. & Vachet, R. W. The basics of mass spectrometry in the twenty-first century. *Nat. Rev. Drug Discov.* **2,** 140–150 (2003).
46. Gingras, A.-C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* **8,** 645–654 (2007).
47. Wenk, M. R. The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **4,** 594–610 (2005).
48. Zaia, J. Mass spectrometry of oligosaccharides. *Mass Spectrom. Rev.* **23,** 161–227 (2004).
49. Wishart, D. S. *et al.* HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **35,** D521–D526 (2007).
50. Hvidt, A. & Linderstrøm-Lang, K. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim. Biophys. Acta* **14,** 574–575 (1954).
51. Englander, S. W., Mayne, L., Bai, Y. & Sosnick, T. R. Hydrogen exchange: The modern legacy of Linderstrøm-Lang. *Protein Sci.* **6,** 1101–1109 (1997).
52. Zhang, Z. & Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci.* **2,** 522–531 (2008).
53. 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.* **133,** 399–416 (1979).

54. Englander, J. J., Rogero, J. R. & Englander, S. W. Protein hydrogen exchange studied by the fragment separation method. *Anal. Biochem.* **147,** 234–244 (1985).
55. Morgan, C. R. & Engen, J. R. in *Current Protocols in Protein Science* (eds. Coligan, J. E., Dunn, B. M., Speicher, D. W. & Wingfield, P. T.) 17.6.1-17.6.17 (John Wiley & Sons, Inc., 2009).
56. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins* **17,** 75–86 (1993).
57. Smith, D. L., Deng, Y. & Zhang, Z. Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J. Mass Spectrom. JMS* **32,** 135–146 (1997).
58. Perrin, C. L. Proton exchange in amides: Surprises from simple systems. *Acc. Chem. Res.* **22,** 268–275 (1989).
59. *Hydrogen exchange mass spectrometry of proteins: fundamentals, methods, and applications*. (John Wiley & Sons, Inc, 2016).
60. Englander, S. W. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **16,** 521–655 (1983).
61. Skinner, J. J., Lim, W. K., Bédard, S., Black, B. E. & Englander, S. W. Protein dynamics viewed by hydrogen exchange. *Protein Sci. Publ. Protein Soc.* **21,** 996–1005 (2012).
62. Hvidt, A. & Nielsen, S. O. Hydrogen exchange in proteins. *Adv. Protein Chem.* **21,** 287–386 (1966).
63. 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.* **43,** 1021–1036 (2008).
64. Englander, J. J. *et al.* Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7057–7062 (2003).
65. 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 (Mosc.)* **50,** 8722–8732 (2011).

66. 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.* **17,** 1498–1509 (2006).

67. Oldfield, C. J. & Dunker, A. K. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annu. Rev. Biochem.* **83,** 553–584 (2014).

68. Connors, K. A. *Chemical kinetics: the study of reaction rates in solution*. (Wiley-VCH, 1990).

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

70. Zhu, S. *et al.* Hyperphosphorylation of intrinsically disordered tau protein induces an amyloidogenic shift in its conformational ensemble. *PloS One* **10,** e0120416 (2015).

71. Rumi-Masante, J. *et al.* Structural basis for activation of calcineurin by calmodulin. *J. Mol. Biol.* **415,** 307–317 (2012).

72. Rob, T. *et al.* Measuring dynamics in weakly structured regions of proteins using microfluidics-enabled subsecond H/D exchange mass spectrometry. *Anal. Chem.* **84,** 3771–3779 (2012).

73. Connelly, G. P., Bai, Y., Jeng, M. F. & Englander, S. W. Isotope effects in peptide group hydrogen exchange. *Proteins* **17,** 87–92 (1993).

74. Anthis, N. J. & Clore, G. M. Visualizing transient dark states by NMR spectroscopy. *Q. Rev. Biophys.* **48,** 35–116 (2015).

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

76. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450,** 964–972 (2007).

77. Goswami, D. *et al.* Time window expansion for HDX analysis of an intrinsically disordered protein. *J. Am. Soc. Mass Spectrom.* **24,** 1584–1592 (2013).

78. Udgaonkar, J. B. & Baldwin, R. L. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature* **335,** 694–699 (1988).

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

80. Dharmasiri, K. & Smith, D. L. Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins. *Anal. Chem.* **68,** 2340–2344 (1996).

81. Simmons, D. A., Dunn, S. D. & Konermann, L. Conformational dynamics of partially denatured myoglobin studied by time-resolved electrospray mass spectrometry with online hydrogen-deuterium exchange. *Biochemistry (Mosc.)* **42,** 5896–5905 (2003).

82. Liuni, P., Rob, T. & Wilson, D. J. A microfluidic reactor for rapid, low-pressure proteolysis with on-chip electrospray ionization. *Rapid Commun. Mass Spectrom. RCM* **24,** 315–320 (2010).

83. Pan, J., Wilson, D. J. & Konermann, L. Pulsed hydrogen exchange and electrospray charge-state distribution as complementary probes of protein structure in kinetic experiments: implications for ubiquitin folding. *Biochemistry (Mosc.)* **44,** 8627–8633 (2005).

84. Pan, J., Han, J., Borchers, C. H. & Konermann, L. Characterizing short-lived protein folding intermediates by top-down hydrogen exchange mass spectrometry. *Anal. Chem.* **82,** 8591–8597 (2010).

85. Yang, H. & Smith, D. L. Kinetics of Cytochrome *c* Folding Examined by Hydrogen Exchange and Mass Spectrometry †. *Biochemistry (Mosc.)* **36,** 14992–14999 (1997).

86. Khanal, A., Pan, Y., Brown, L. S. & Konermann, L. Pulsed hydrogen/deuterium exchange mass spectrometry for time-resolved membrane protein folding studies. *J. Mass Spectrom. JMS* **47,** 1620–1626 (2012).

87. Liuni, P., Jeganathan, A. & Wilson, D. J. Conformer selection and intensified dynamics during catalytic turnover in chymotrypsin. *Angew. Chem. Int. Ed Engl.* **51,** 9666–9669 (2012).

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

89. Mitchell, J. L. *et al.* Functional Characterization and Conformational Analysis of the Herpesvirus saimiri Tip-C484 Protein. *J. Mol. Biol.* **366,** 1282–1293 (2007).

90. Mandelkow, E.-M. & Mandelkow, E. Biochemistry and cell biology of tau protein in neurofibrillary degeneration. *Cold Spring Harb. Perspect. Med.* **2,** a006247 (2012).

91. Keppel, T. R. & Weis, D. D. Mapping residual structure in intrinsically disordered proteins at residue resolution using millisecond hydrogen/deuterium exchange and residue averaging. *J. Am. Soc. Mass Spectrom.* **26,** 547–554 (2015).

92. 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. Publ. Protein Soc.* **20,** 2023–2034 (2011).

93. Mittag, T., Kay, L. E. & Forman-Kay, J. D. Protein dynamics and conformational disorder in molecular recognition. *J. Mol. Recognit. JMR* **23,** 105–116 (2010).

94. Forman-Kay, J. D. & Mittag, T. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Struct. Lond. Engl. 1993* **21,** 1492–1499 (2013).

95. Redfern, O. C., Dessailly, B. & Orengo, C. A. Exploring the structure and function paradigm. *Curr. Opin. Struct. Biol.* **18,** 394–402 (2008).

96. Baldwin, A. J. & Kay, L. E. NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5,** 808–814 (2009).

97. Osawa, M., Takeuchi, K., Ueda, T., Nishida, N. & Shimada, I. Functional dynamics of proteins revealed by solution NMR. *Curr. Opin. Struct. Biol.* **22,** 660–669 (2012).

98. Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J. & Pavletich, N. P. Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature* **382,** 325–331 (1996).

99. Breidenbach, M. A. & Brunger, A. T. Substrate recognition strategy for botulinum neurotoxin serotype A. *Nature* **432,** 925–929 (2004).

100. Fuxreiter, M. *et al.* Malleable machines take shape in eukaryotic transcriptional regulation. *Nat. Chem. Biol.* **4,** 728–737 (2008).

101. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. & Uversky, V. N. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **272,** 5129–5148 (2005).

102. 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.* **337,** 635–645 (2004).

103. Lobley, A., Swindells, M. B., Orengo, C. A. & Jones, D. T. Inferring function using patterns of native disorder in proteins. *PLoS Comput. Biol.* **3,** e162 (2007).

104. Brown, C. J. *et al.* Evolutionary rate heterogeneity in proteins with long disordered regions. *J. Mol. Evol.* **55,** 104–110 (2002).

105. Gunasekaran, K., Tsai, C.-J., Kumar, S., Zanuy, D. & Nussinov, R. Extended disordered proteins: targeting function with less scaffold. *Trends Biochem. Sci.* **28,** 81–85 (2003).

106. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are ‘natively unfolded’ proteins unstructured under physiologic conditions? *Proteins* **41,** 415–427 (2000).

107. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475,** 324–332 (2011).

108. Hilser, V. J. & Thompson, E. B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl. Acad. Sci.* **104,** 8311–8315 (2007).

109. Ferreon, A. C. M., Ferreon, J. C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498,** 390–394 (2013).

110. Kim, P. M., Sboner, A., Xia, Y. & Gerstein, M. The role of disorder in interaction networks: a structural analysis. *Mol. Syst. Biol.* **4,** (2008).

111. Patil, A. & Nakamura, H. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* **580,** 2041–2045 (2006).

112. Bailey, R. W., Dunker, A. K., Brown, C. J., Garner, E. C. & Griswold, M. D. Clusterin, a binding protein with a molten globule-like region. *Biochemistry (Mosc.)* **40,** 11828–11840 (2001).

113. Fink, A. L. Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **3,** R9–R23 (1998).

114. Jones, R. B., Gordus, A., Krall, J. A. & MacBeath, G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439,** 168–174 (2006).

115. Ernst, A. *et al.* Coevolution of PDZ domain-ligand interactions analyzed by high-throughput phage display and deep sequencing. *Mol. Biosyst.* **6,** 1782–1790 (2010).

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

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

118. Zilfou, J. T. & Lowe, S. W. Tumor Suppressive Functions of p53. *Cold Spring Harb. Perspect. Biol.* **1,** a001883–a001883 (2009).

119. Turoverov, K. K., Kuznetsova, I. M. & Uversky, V. N. The protein kingdom extended: Ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog. Biophys. Mol. Biol.* **102,** 73–84 (2010).

120. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* **37,** 215–246 (2008).

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

122. Uversky, V. N., Roman, A., Oldfield, C. J. & Dunker, A. K. Protein intrinsic disorder and human papillomaviruses: increased amount of disorder in E6 and E7 oncoproteins from high risk HPVs. *J. Proteome Res.* **5,** 1829–1842 (2006).

123. Black, D. L. Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* **72,** 291–336 (2003).

124. Romero, P. R. *et al.* Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8390–8395 (2006).

125. Roberts, R. *et al.* Altered phosphorylation and intracellular distribution of a (CUG)n triplet repeat RNA-binding protein in patients with myotonic dystrophy and in myotonin protein kinase knockout mice. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 13221–13226 (1997).

126. Skotheim, R. I. & Nees, M. Alternative splicing in cancer: noise, functional, or systematic? *Int. J. Biochem. Cell Biol.* **39,** 1432–1449 (2007).

127. Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **273,** 729–739 (1997).

128. Jenkins, L. M. M., Durell, S. R., Mazur, S. J. & Appella, E. p53 N-terminal phosphorylation: a defining layer of complex regulation. *Carcinogenesis* **33,** 1441–1449 (2012).

129. Joerger, A. C. & Fersht, A. R. The p53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. *Annu. Rev. Biochem.* **85,** 375–404 (2016).

130. Meek, D. W. Multisite phosphorylation and the integration of stress signals at p53. *Cell. Signal.* **10,** 159–166 (1998).

131. Ashcroft, M., Kubbutat, M. H. & Vousden, K. H. Regulation of p53 function and stability by phosphorylation. *Mol. Cell. Biol.* **19,** 1751–1758 (1999).

132. Lee, C. W., Ferreon, J. C., Ferreon, A. C. M., Arai, M. & Wright, P. E. Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 19290–19295 (2010).

133. Teufel, D. P., Bycroft, M. & Fersht, A. R. Regulation by phosphorylation of the relative affinities of the N-terminal transactivation domains of p53 for p300 domains and Mdm2. *Oncogene* **28,** 2112–2118 (2009).

134. Di Lello, P. *et al.* Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. *Mol. Cell* **22,** 731–740 (2006).

135. Chang, J., Kim, D. H., Lee, S. W., Choi, K. Y. & Sung, Y. C. Transactivation ability of p53 transcriptional activation domain is directly related to the binding affinity to TATA-binding protein. *J. Biol. Chem.* **270,** 25014–25019 (1995).

136. Van Orden, K., Giebler, H. A., Lemasson, I., Gonzales, M. & Nyborg, J. K. Binding of p53 to the KIX domain of CREB binding protein. A potential link to human T-cell leukemia virus, type I-associated leukemogenesis. *J. Biol. Chem.* **274,** 26321–26328 (1999).

137. MacPherson, D. *et al.* Defective apoptosis and B-cell lymphomas in mice with p53 point mutation at Ser 23. *EMBO J.* **23,** 3689–3699 (2004).

138. Friedhoff, P., von Bergen, M., Mandelkow, E. M. & Mandelkow, E. Structure of tau protein and assembly into paired helical filaments. *Biochim. Biophys. Acta* **1502,** 122–132 (2000).

139. Goedert, M. & Jakes, R. Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization. *EMBO J.* **9,** 4225–4230 (1990).

140. Chen, J., Kanai, Y., Cowan, N. J. & Hirokawa, N. Projection domains of MAP2 and tau determine spacings between microtubules in dendrites and axons. *Nature* **360,** 674–677 (1992).

141. Jeganathan, S., von Bergen, M., Brutlach, H., Steinhoff, H.-J. & Mandelkow, E. Global hairpin folding of tau in solution. *Biochemistry (Mosc.)* **45,** 2283–2293 (2006).

142. Gong, C.-X. & Iqbal, K. Hyperphosphorylation of microtubule-associated protein tau: a promising therapeutic target for Alzheimer disease. *Curr. Med. Chem.* **15,** 2321–2328 (2008).

143. Hutton, M. *et al.* Association of missense and 5’-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393,** 702–705 (1998).

144. Hardy, J. & Singleton, A. The HapMap: Charting a Course for Genetic Discovery in Neurological Diseases. *Arch. Neurol.* **65,** (2008).

145. Bloom, G. S. Amyloid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol.* **71,** 505–508 (2014).

146. Roberson, E. D. *et al.* Reducing endogenous tau ameliorates amyloid beta-induced deficits in an Alzheimer’s disease mouse model. *Science* **316,** 750–754 (2007).

147. Biernat, J., Gustke, N., Drewes, G., Mandelkow, E. M. & Mandelkow, E. Phosphorylation of Ser262 strongly reduces binding of tau to microtubules: distinction between PHF-like immunoreactivity and microtubule binding. *Neuron* **11,** 153–163 (1993).

148. Köpke, E. *et al.* Microtubule-associated protein tau. Abnormal phosphorylation of a non-paired helical filament pool in Alzheimer disease. *J. Biol. Chem.* **268,** 24374–24384 (1993).

149. Hanger, D. P. *et al.* Novel Phosphorylation Sites in Tau from Alzheimer Brain Support a Role for Casein Kinase 1 in Disease Pathogenesis. *J. Biol. Chem.* **282,** 23645–23654 (2007).

150. Drewes, G., Ebneth, A., Preuss, U., Mandelkow, E. M. & Mandelkow, E. MARK, a novel family of protein kinases that phosphorylate microtubule-associated proteins and trigger microtubule disruption. *Cell* **89,** 297–308 (1997).

151. Biernat, J. *et al.* The switch of tau protein to an Alzheimer-like state includes the phosphorylation of two serine-proline motifs upstream of the microtubule binding region. *EMBO J.* **11,** 1593–1597 (1992).

152. Goedert, M. *et al.* Epitope mapping of monoclonal antibodies to the paired helical filaments of Alzheimer's disease: identification of phosphorylation sites in tau protein. *Biochem. J.* **301,** 871–877 (1994).

153. Shammas, S. L. *et al.* A mechanistic model of tau amyloid aggregation based on direct observation of oligomers. *Nat. Commun.* **6,** 7025 (2015).

154. von Bergen, M. *et al.* Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif (306VQIVYK311) forming beta structure. *Proc. Natl. Acad. Sci.* **97,** 5129–5134 (2000).

155. Margittai, M. & Langen, R. Template-assisted filament growth by parallel stacking of tau. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 10278–10283 (2004).

156. Wegmann, S., Medalsy, I. D., Mandelkow, E. & Müller, D. J. The fuzzy coat of pathological human Tau fibrils is a two-layered polyelectrolyte brush. *Proc. Natl. Acad. Sci.* **110,** E313–E321 (2013).

157. Gamblin, T. C. *et al.* Caspase cleavage of tau: Linking amyloid and neurofibrillary tangles in Alzheimer's disease. *Proc. Natl. Acad. Sci.* **100,** 10032–10037 (2003).

158. Li, X., Lu, F., Wang, J.-Z. & Gong, C.-X. Concurrent alterations of O-GlcNAcylation and phosphorylation of tau in mouse brains during fasting. *Eur. J. Neurosci.* **23,** 2078–2086 (2006).

159. Cohen, T. J. *et al.* The acetylation of tau inhibits its function and promotes pathological tau aggregation. *Nat. Commun.* **2,** 252 (2011).

160. Citron, M. Alzheimer's disease: strategies for disease modification. *Nat. Rev. Drug Discov.* **9,** 387–398 (2010).

161. Goh, C.-S., Milburn, D. & Gerstein, M. Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.* **14,** 104–109 (2004).

162. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. Transient protein-protein interactions: structural, functional, and network properties. *Struct. Lond. Engl. 1993* **18,** 1233–1243 (2010).

163. Jones, S. & Thornton, J. M. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13–20 (1996).

164. Neduva, V. & Russell, R. B. Linear motifs: evolutionary interaction switches. *FEBS Lett.* **579,** 3342–3345 (2005).

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

166. Carroll, M. J. *et al.* Evidence for dynamics in proteins as a mechanism for ligand dissociation. *Nat. Chem. Biol.* **8,** 246–252 (2012).

167. Horn, J. R. *et al.* The role of protein dynamics in increasing binding affinity for an engineered protein-protein interaction established by H/D exchange mass spectrometry. *Biochemistry (Mosc.)* **45,** 8488–8498 (2006).

168. Blau, J. *et al.* Three functional classes of transcriptional activation domain. *Mol. Cell. Biol.* **16,** 2044–2055 (1996).

169. Shieh, S. Y., Ahn, J., Tamai, K., Taya, Y. & Prives, C. The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. *Genes Dev.* **14,** 289–300 (2000).

170. Moll, U. M. & Petrenko, O. The MDM2-p53 interaction. *Mol. Cancer Res. MCR* **1,** 1001–1008 (2003).

171. Loughery, J. & Meek, D. Switching on p53: an essential role for protein phosphorylation? *Biodiscovery* (2013). doi:10.7750/BioDiscovery.2013.8.1

172. Daughdrill, G. W. *et al.* Understanding the structural ensembles of a highly extended disordered protein. *Mol. Biosyst.* **8,** 308–319 (2012).

173. Lee, H. *et al.* Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *J. Biol. Chem.* **275,** 29426–29432 (2000).

174. Homaeian, L., Kurgan, L. A., Ruan, J., Cios, K. J. & Chen, K. Prediction of protein secondary structure content for the twilight zone sequences. *Proteins* **69,** 486–498 (2007).

175. Li, P., Fan, Y., Wang, Y., Lu, Y. & Yin, Z. Characterization of plasma protein binding dissociation with online SPE-HPLC. *Sci. Rep.* **5,** 14866 (2015).

176. Borg, M. *et al.* Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 9650–9655 (2007).

177. Buée, L., Bussière, T., Buée-Scherrer, V., Delacourte, A. & Hof, P. R. Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. *Brain Res. Brain Res. Rev.* **33,** 95–130 (2000).

178. Drubin, D. G. & Kirschner, M. W. Tau protein function in living cells. *J. Cell Biol.* **103,** 2739–2746 (1986).

179. Spillantini, M. G. & Goedert, M. Tau pathology and neurodegeneration. *Lancet Neurol.* **12,** 609–622 (2013).

180. Hernandez, F., Lucas, J. J. & Avila, J. GSK3 and tau: two convergence points in Alzheimer's disease. *J. Alzheimers Dis. JAD* **33 Suppl 1,** S141-144 (2013).

181. Paul, M. M. & LeVine, I. H. Alzheimer's Disease and the Amyloid-β Peptide. *J. Alzheimers Dis.* 311–323 (2010). doi:10.3233/JAD-2010-1221

182. Engel, T., Hernández, F., Avila, J. & Lucas, J. J. Full reversal of Alzheimer's disease-like phenotype in a mouse model with conditional overexpression of glycogen synthase kinase-3. *J. Neurosci. Off. J. Soc. Neurosci.* **26,** 5083–5090 (2006).

183. Ishiguro, K. *et al.* Glycogen synthase kinase 3 beta is identical to tau protein kinase I generating several epitopes of paired helical filaments. *FEBS Lett.* **325,** 167–172 (1993).

184. Sperber, B. R., Leight, S., Goedert, M. & Lee, V. M. Glycogen synthase kinase-3 beta phosphorylates tau protein at multiple sites in intact cells. *Neurosci. Lett.* **197,** 149–153 (1995).

185. Mondragón-Rodríguez, S., Perry, G., Luna-Muñoz, J., Acevedo-Aquino, M. C. & Williams, S. Phosphorylation of tau protein at sites Ser(396-404) is one of the earliest events in Alzheimer's disease and Down syndrome. *Neuropathol. Appl. Neurobiol.* **40,** 121–135 (2014).

186. Lee, H.-G. *et al.* Tau phosphorylation in Alzheimer's disease: pathogen or protector? *Trends Mol. Med.* **11,** 164–169 (2005).

187. Kragelj, J., Ozenne, V., Blackledge, M. & Jensen, M. R. Conformational propensities of intrinsically disordered proteins from NMR chemical shifts. *Chemphyschem Eur. J. Chem. Phys. Phys. Chem.* **14,** 3034–3045 (2013).

188. Mukrasch, M. D. *et al.* Structural polymorphism of 441-residue tau at single residue resolution. *PLoS Biol.* **7,** e34 (2009).

189. Ozenne, V. *et al.* Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *J. Am. Chem. Soc.* **134,** 15138–15148 (2012).

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

191. Mylonas, E. *et al.* Domain conformation of tau protein studied by solution small-angle X-ray scattering. *Biochemistry (Mosc.)* **47,** 10345–10353 (2008).

192. Nath, A. *et al.* The conformational ensembles of α-synuclein and tau: combining single-molecule FRET and simulations. *Biophys. J.* **103,** 1940–1949 (2012).

193. Dunn, B. M. Overview of pepsin-like aspartic peptidases. *Curr. Protoc. Protein Sci.* **Chapter 21,** Unit 21.3 (2001).

194. Daebel, V. *et al.* β-Sheet core of tau paired helical filaments revealed by solid-state NMR. *J. Am. Chem. Soc.* **134,** 13982–13989 (2012).

195. Steinhilb, M. L., Dias-Santagata, D., Fulga, T. A., Felch, D. L. & Feany, M. B. Tau phosphorylation sites work in concert to promote neurotoxicity in vivo. *Mol. Biol. Cell* **18,** 5060–5068 (2007).

196. Augustinack, J. C., Schneider, A., Mandelkow, E.-M. & Hyman, B. T. Specific tau phosphorylation sites correlate with severity of neuronal cytopathology in Alzheimer's disease. *Acta Neuropathol. (Berl.)* **103,** 26–35 (2002).

197. Hagestedt, T., Lichtenberg, B., Wille, H., Mandelkow, E. M. & Mandelkow, E. Tau protein becomes long and stiff upon phosphorylation: correlation between paracrystalline structure and degree of phosphorylation. *J. Cell Biol.* **109,** 1643–1651 (1989).

198. Elbaum-Garfinkle, S. & Rhoades, E. Identification of an aggregation-prone structure of tau. *J. Am. Chem. Soc.* **134,** 16607–16613 (2012).

199. Farrell, D. W., Speranskiy, K. & Thorpe, M. F. Generating stereochemically acceptable protein pathways. *Proteins* **78,** 2908–2921 (2010).

200. Willard, L. *et al.* VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* **31,** 3316–3319 (2003).

201. Wang, Q. M., Fiol, C. J., DePaoli-Roach, A. A. & Roach, P. J. Glycogen synthase kinase-3 beta is a dual specificity kinase differentially regulated by tyrosine and serine/threonine phosphorylation. *J. Biol. Chem.* **269,** 14566–14574 (1994).

202. Rankin, C. A., Sun, Q. & Gamblin, T. C. Tau phosphorylation by GSK-3beta promotes tangle-like filament morphology. *Mol. Neurodegener.* **2,** 12 (2007).

203. Scheltens, P. *et al.* Alzheimer's disease. *Lancet Lond. Engl.* **388,** 505–517 (2016).

204. Mangialasche, F., Solomon, A., Winblad, B., Mecocci, P. & Kivipelto, M. Alzheimer's disease: clinical trials and drug development. *Lancet Neurol.* **9,** 702–716 (2010).

205. Birks, J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst. Rev.* CD005593 (2006). doi:10.1002/14651858.CD005593

206. Winblad, B. *et al.* Donepezil in patients with severe Alzheimer's disease: double-blind, parallel-group, placebo-controlled study. *Lancet Lond. Engl.* **367,** 1057–1065 (2006).

207. Reisberg, B. *et al.* Memantine in moderate-to-severe Alzheimer's disease. *N. Engl. J. Med.* **348,** 1333–1341 (2003).

208. Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297,** 353–356 (2002).

209. Nisbet, R. M., Polanco, J.-C., Ittner, L. M. & Götz, J. Tau aggregation and its interplay with amyloid-β. *Acta Neuropathol. (Berl.)* **129,** 207–220 (2015).

210. Adolfsson, O. *et al.* An effector-reduced anti-β-amyloid (Aβ) antibody with unique aβ binding properties promotes neuroprotection and glial engulfment of Aβ. *J. Neurosci. Off. J. Soc. Neurosci.* **32,** 9677–9689 (2012).

211. Salloway, S. *et al.* Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. *N. Engl. J. Med.* **370,** 322–333 (2014).

212. Seubert, P. *et al.* Antibody capture of soluble Abeta does not reduce cortical Abeta amyloidosis in the PDAPP mouse. *Neurodegener. Dis.* **5,** 65–71 (2008).

213. Costa, P. *The Nature and Origin of Amyloid Fibrils.* (John Wiley & Sons, 1996).

214. Li, J., Dangott, L. J. & Fitzpatrick, P. F. Regulation of phenylalanine hydroxylase: conformational changes upon phenylalanine binding detected by hydrogen/deuterium exchange and mass spectrometry. *Biochemistry (Mosc.)* **49,** 3327–3335 (2010).

215. Liu, Z. & Huang, Y. Advantages of proteins being disordered: Advantages of IDPs. *Protein Sci.* **23,** 539–550 (2014).

216. Bohrer, B. C., Merenbloom, S. I., Koeniger, S. L., Hilderbrand, A. E. & Clemmer, D. E. Biomolecule analysis by ion mobility spectrometry. *Annu. Rev. Anal. Chem. Palo Alto Calif* **1,** 293–327 (2008).

217. Bennett, R. V., Gamage, C. M., Galhena, A. S. & Fernández, F. M. Contrast-enhanced differential mobility-desorption electrospray ionization-mass spectrometry imaging of biological tissues. *Anal. Chem.* **86,** 3756–3763 (2014).

218. Guevremont, R., Purves, R. . & Barnett, D. . FAIMS apparatus and method using carrier gas of mixed compositions. *US Pat. No 6* **774,** (2004).

219. Guevremont, R., Purves, R. W., Barnett, D. . & Ells, B. FAIMS apparatus and method using carrier gases that contain a trace amount of dopant species. *US Pat. No 7* **26,** (2006).

220. Kafle, A. *et al.* Understanding gas phase modifier interactions in rapid analysis by differential mobility-tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **25,** 1098–1113 (2014).

221. Shvartsburg, A. A., Li, F., Tang, K. & Smith, R. D. Characterizing the structures and folding of free proteins using 2-D gas-phase separations: observation of multiple unfolded conformers. *Anal. Chem.* **78,** 3304–3315 (2006).

222. Evans, S. E., Lueck, N. & Marzluff, E. M. Gas phase hydrogen/deuterium exchange of proteins in an ion trap mass spectrometer. *Int. J. Mass Spectrom.* **222,** 175–187 (2003).

223. Geller, O. & Lifshitz, C. A fast flow tube study of gas phase H/D exchange of multiply protonated ubiquitin. *J. Phys. Chem. A* **109,** 2217–2222 (2005).

224. Cox, H. A., Julian, R. R., Lee, S.-W. & Beauchamp, J. L. Gas-phase H/D exchange of sodiated glycine oligomers with ND3: exchange kinetics do not reflect parent ion structures. *J. Am. Chem. Soc.* **126,** 6485–6490 (2004).

225. Wood, T. D. *et al.* Gas-phase folding and unfolding of cytochrome c cations. *Proc. Natl. Acad. Sci.* **92,** 2451–2454 (1995).

226. Rand, K. D. *et al.* Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Anal. Chem.* **81,** 10019–10028 (2009).

227. Robinson, E. W. & Williams, E. R. Multidimensional separations of ubiquitin conformers in the gas phase: Relating ion cross sections to H/D exchange measurements. *J. Am. Soc. Mass Spectrom.* **16,** 1427–1437 (2005).

228. Zhu, S., Campbell, J. L., Chernushevich, I., Le Blanc, J. C. Y. & Wilson, D. J. Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution. *J. Am. Soc. Mass Spectrom.* **27,** 991–999 (2016).

229. Parson, W. B. *et al.* Rapid analysis of isomeric exogenous metabolites by differential mobility spectrometry-mass spectrometry. *Rapid Commun. Mass Spectrom. RCM* **25,** 3382–3386 (2011).

230. Campbell, J. L., Le Blanc, J. C. Y. & Schneider, B. B. Probing Electrospray Ionization Dynamics Using Differential Mobility Spectrometry: The Curious Case of 4-Aminobenzoic Acid. *Anal. Chem.* **84,** 7857–7864 (2012).

231. Wright, P., Zhang, J. & Douglas, D. Conformations of Gas-Phase Ions of Ubiquitin, Cytochrome c, Apomyoglobin, and β-Lactoglobulin Produced from Two Different Solution Conformations. *J. Am. Soc. Mass Spectrom.* **19,** 1906–1913 (2008).

232. Andrews, G. L., Simons, B. L., Young, J. B., Hawkridge, A. M. & Muddiman, D. C. Performance Characteristics of a New Hybrid Quadrupole Time-of-Flight Tandem Mass Spectrometer (TripleTOF 5600). *Anal. Chem.* **83,** 5442–5446 (2011).

233. Guttman, M. *et al.* Tuning a High Transmission Ion Guide to Prevent Gas-Phase Proton Exchange During H/D Exchange MS Analysis. *J. Am. Soc. Mass Spectrom.* **27,** 662–668 (2016).

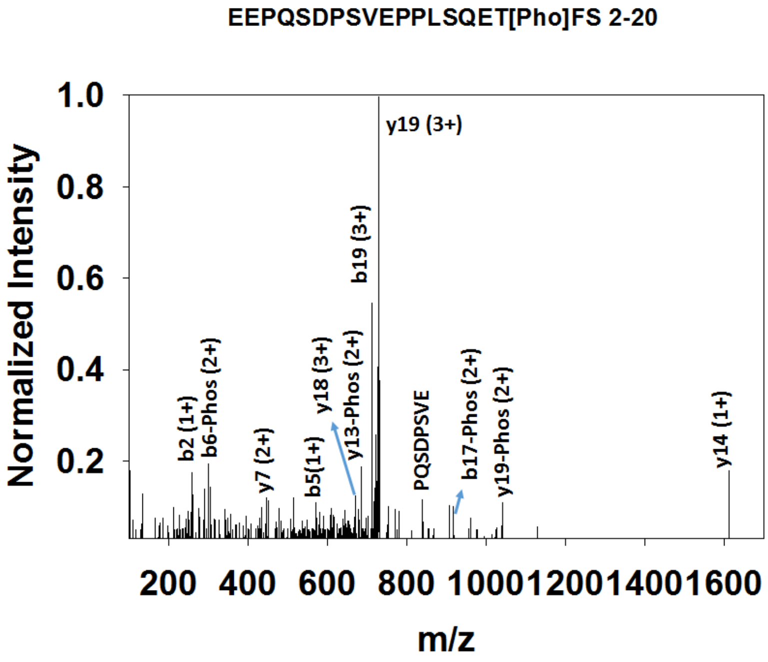
234. Wilson, D. J. & Konermann, L. Ultrarapid Desalting of Protein Solutions for Electrospray Mass Spectrometry in a Microchannel Laminar Flow Device. *Anal. Chem.* **77,** 6887–6894 (2005).

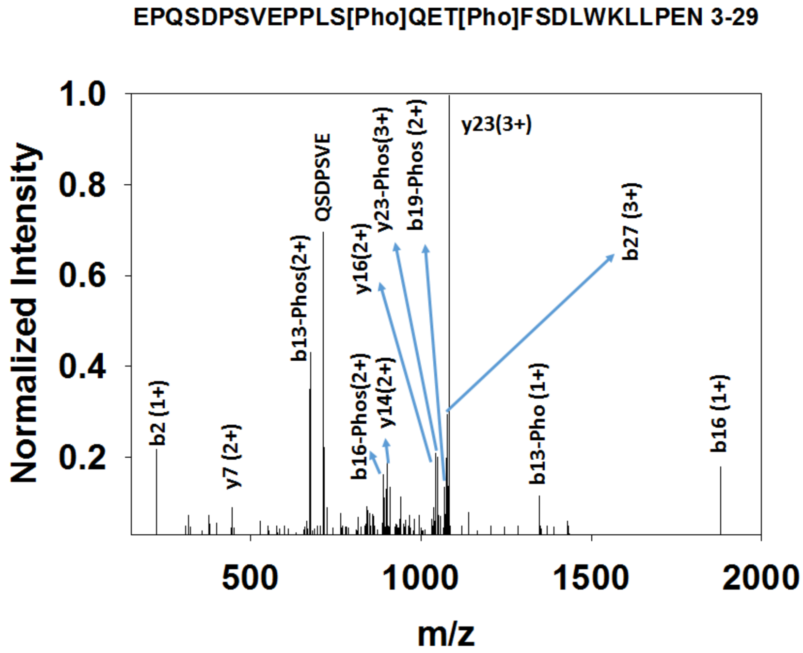
235. Meier, S. *et al.* Pathological Tau Promotes Neuronal Damage by Impairing Ribosomal Function and Decreasing Protein Synthesis. *J. Neurosci. Off. J. Soc. Neurosci.* **36,** 1001–1007 (2016).

236. Bullock, A. N. *et al.* Thermodynamic stability of wild-type and mutant p53 core domain. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 14338–14342 (1997).

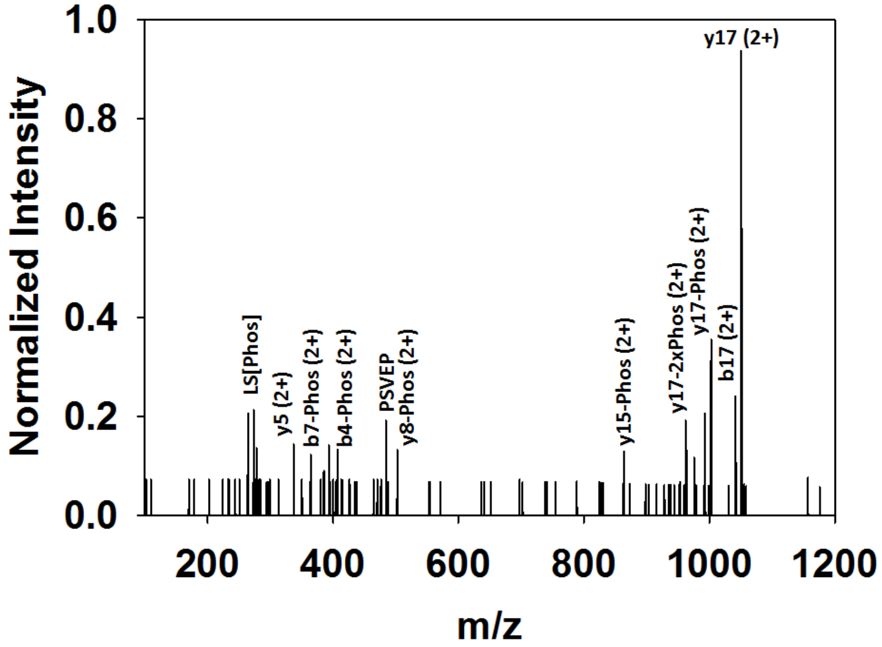
Appendices

Appendix A: Supplementary informations for chapter 2

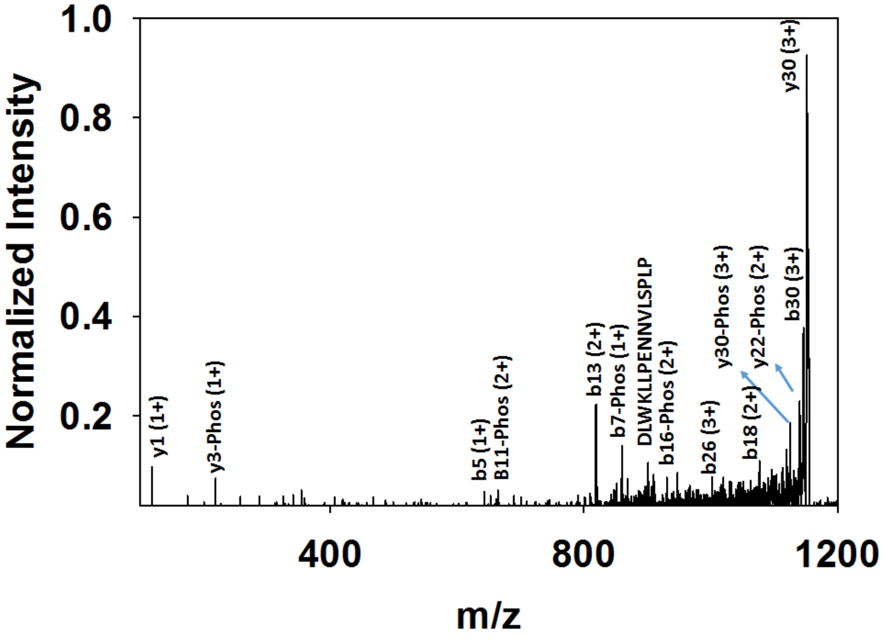

EEPQSDPSVEPPLSQET[Pho]FS 2-20

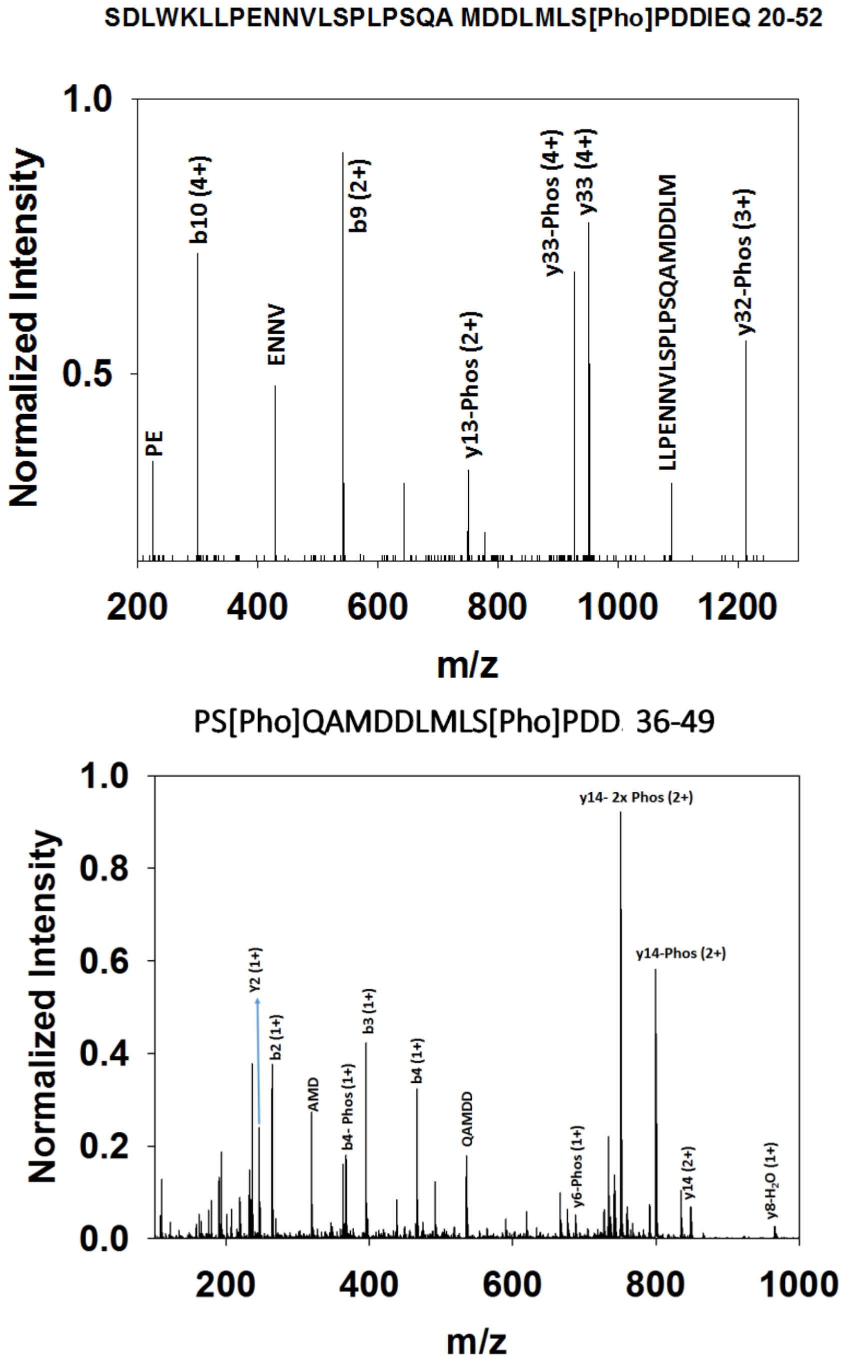

EPQSDPSVEPPLS[Pho]QET[Pho]FSDLWKLLPEN 3-29

QS[Pho]DPSVEPPLS[Pho]QET[Pho]FSD 5-21



TFS[Pho]DLWKLLPENNVLSPLPSQAMDDLMLSP 18-47




SDLWKLLPENNVLSPLPSQA MDDLMLS[Pho]PDDIEQ 20-52

PS[Pho]QAMDDLMLS[Pho]PDD 36-49

Figure A1. MS/MS spectra for phosphopeptides encompassing phosphorylation at residues Ser6, Ser15, Thr18, Ser20, Ser37, and S46

Appendix B: Supplementary informations for chapter 3

Table B1. Proteolytic peptides resulting from native and hyperphosphorylated tau along with their associated PF. Adapted from Zhu *et al*[70].

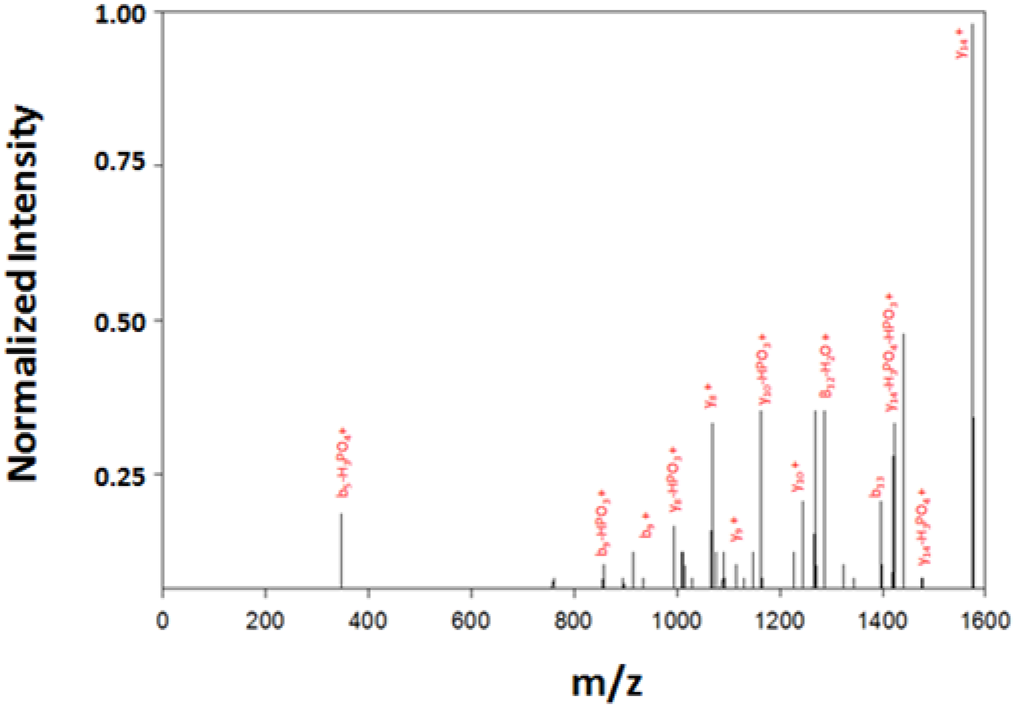
Native		**Hyperphosphorylated***	
Peptide	**PF**	**Peptide**	**PF**
(A)/EPRQEF/(E)3-8	6.4±2.9	(A)/EPRQE/(F) 3-7	1.2±0.3
(R)KDQGGYTMHQDQ/(E) 24-35	4.9±1.2	(Q)/EFEVME/(D) 7-12	2.6±0.6
(R)KDQGGY/(T) 24-29	70.0±7	(L)/GDRKD(Q) 21-25	1.3±0.12
(M)HQDQEGDT(D) 32-39	3.6±0.9	(T)MHQDQE/(G) 31-36	1.4±0.13
(E)/GDTDAGLKESPLQT(P) 37-50	2.7±0.5	(Q)/DQEGDTDAGLKE/(S) 34-45	2.7±0.6
(D)AGLKESP(L) 41-47	1.7±0.2	(D)AGLK(E) 41-44	1.6±0.15
(L)/QTPTE/(D)49-53	69.0±7	(E)/SPLQ/(T) 46-49	1.2±0.2
(E)/AAALIGDEPLENYLDTE/(Y) 56-72	5.4±2.3	(G)SEEPGSE/(T) 56-62	1.4±0.07
(S)EEPGSETSDAK(S)57-67	3.0±0.8	(E)/TSDAKS(T) 63-68	3.2±0.4
(E)/TSDA/(K) 63-66	1.9±0.3	(A)/KSTpPTAEDVTAPL/(V) 67-79 PHOS	2.7±0.05
(T)AEDVTAP(L) 72-78	1.3±0.13	(P)TAED(V) 71-74	1.0+0.02
(A)/PLVDEGAPGKQAAAQPHTE/(I)78-96	7.0±2.33	(T)AEDVTAP(L) 72-78	1.0±0.07
(L)/VDEGAP(G)80-85	10.0±1.6	(Q)/AAAQ/(P) 89-92	1.1±0.05
(V)DEGAPG(K) 81-86	7.4±0.25	(A)/QPHTEIPEGTTAEEA/(G) 92-106	1.1±0.05
(K)QAAAQPHTEIPEGT(T)88-101	1.5±0.2	(H)TEIP(E) 95-98	1.1±0.1
(K)QAAAQP(H) 88-93	1.2±0.16	(E)/IPEGTTAEE/(A) 97-105	1.6±0.5
(Q)/AAAQ/(P) 89-92	1.0±0.1	(G)TTAEEA/(G) 101-106	4.2±1.4
(Q)/PHTEIP(E)93-98	3.2±0.21	(E)/AGIGD(T) 106-110	2.1±0.2
(H)TEIP(E) 95-98	3.0±0.47	(E)/DEAA/(G) 116-119	2.5±0.01
(E)/IPEGTTAEEA/(G)97-106	66.1±6.6	(A)/GHVTQ/(A) 120-124	7.9±0.6
(A)/EEAGIGDTPSL/(E) 104-114	23.3±5.6	(V)SKSKD(G) 129-133	2.3±0.8
(L)/EDEA/(A)115-118	1.3±0.18	(S)KSKDGT(G) 130-135	3.4±0.37
(A)/GHVTQ/(A) 120-124	2.5±0.6	(G)SDDK(K) 137-140	2.3±0.5
(G)SDDK(K)137-140	1.6±0.04	(K)KAKGADG(K) 141-147	1.4±0.15
(D)DKKAKG(A)139-144	3.3+0.9	(K)KAKGADGKT(K) 141-149	1.0±0.12
(D)KKAK(G)140-143	1.9±0.26	(A)/DGKT(K) 146-149	2.4±0.18
(I)ATPRGAAP(P) 152-159	11.7±0.8	(A)/DGKTpKIATPRGAA/(P) 146-158 PHOS	3.0±0.3
(A)/TPRGA/(A) 153-157	2.3±0.45	(A)/TpPRGAA/(P) 153-158 PHOS	2.4±0.18
(A)/APPGQK(G)158-163	2.8±0.5	(T)RIPAKTp?PPAPKTp?P(P) 170-182 PHOS	2.2±0.25

(A)/APPGQKGQ/(A)158-165	7.8±0.3	(A)/KTp?PPAPKTp?PPSSpGEPPKSGD RSp?GYSp?SpPGSp?(P) 174-202 (PHOS 4X)	1.9±0.25
(P)PGQKGQA/(N) 160-166	10.1+1	(P)APKTPPSS(G) 178-185	3.1±0.16
/KGQA/(N) 163-166	1.1±0.1	(G)YSSP(G) 197-200	4.1±0.1
(Q)/KGQANATRIPA/(K)163-173	8.8±1.8	(P)GSpPG(T) 201-204 PHOS	3.0±0.23
(P)APKTPPSS(G)178-185	8.8±0.42	(S)PGTP(G) 203-206	2.4±0.07
(P)SSGEPPKSGDRSGY/(S) 184-197	5.4±1.5	(S)PGTPG(S) 203-207	1.3±0.16
(E)/PPKSGDRSGYSSPG(S) 188-201	4.4±0.9	(T)PGSRSRTPSL/(P) 206-215	1.7±0.29
(Y)/SSPGSPGTPGSRSRT(P)198-212	6.4±2	(R)TPSpL/(P) 212-215 PHOS	4.0±0.04
(G)TPGSRSRTPSLPTP(P)205-218	3.8±0.8	(P)SpLPT(P) 214-217 PHOS	3.6±0.8
(R)EPKKV(A)222-226	2.1+0.5	(P)TPPT(R) 217-220	2.1±0.4
(K)KVAV(V) 225-228	2.0±0.4	(P)TpPPTp?REPKKVAVVRTpPPKSp?P Sp?SAKSp?R(L) 217-242 (PHOS 3X)	1.3±0.16
(K)VAVVRTPPKSPSS(A)226-238	3.7±1.2	(T)PPKSP(S) 232-236	3.0±0.15
(V)AVVRTPP(K) 227-233	2.3±0.02	(S)PSSAKSRLQTA/(P) 236-246	3.0±0.13
(V)AVVRT(P) 227-231	7.0+0.03	(A)/KSRL/(Q) 240-243	2.7±0.1
(A)/KSRL/(Q) 240-243	4.9±1	(Q)/PGGGK(V) 270-274	2.5±0.3
(A)/KSRLQ/(T)240-244	4.4±0.02	(I)INKK(L) 278-281	6.3±1.5
(K)NVKSKIG(S) 255-261	2.4±0.1	(K)HVPGGGS(V) 299-305	1.3±0.18
(K)SKIG(S)258-261	1.8±0.7	(K)HVPGGGSpV(Q) 299-306 PHOS	1.0±0.07
(K)HQPGGGKVQII(N)268-278	7.8±3	(K)PVDLSKVTSKCG(S) 312-323	2.3±0.19
(G)GKVQ/(I) 273-276	3.4±0.38	(L)/GNIHHKPGGGQ/(V) 326-336	4.2±0.75
(S)NVQSKCGSKD(N) 286-295	35.2±3.5	(L)/DFK(D) 345-347	3.6±1.6
(K)HVPGGGSVQIV(Y)299-309	6.5±0.8	(D)RVQSp(K) 349-352 PHOS	11.0±2.3
(P)VDLSKVTS(K)313-320	4.9±0.03	(I)GSLDNI(T) 355-360	1.1±0.04
(P)VDLSK(V) 313-317	5.8±0.03	(V)PGGGN(K) 364-368	4.7±1.8
(P)VDLSKV(T)313-318	5.0+0.2	(E)/TpHKL/(T) 373-376 PHOS	1.4±0.02
(L)/GNIHHKPGGGQVEVKSE/(K)326-342	2.0+0.24	(L)/TFRE/(N) 377-380	3.0±0.37
(Q)/VEVKS(E)337-341	1.1+0.1	(T)FRENA/(K) 378-382	2.9±0.45
(E)/VKSEK(L) 339-343	1.6+0.2	(E)/NAKA/(K) 381-384	5.0±0.25
(K)IGSLDN(I) 354-359	17.5±0.2	(T)DHGAEIV(Y) 387-393	2.8±0.3
(S)LDNITHV(P)357-363	23.3±0.2	(Y)/KSp?PVVSp?(G) 395-400 PHOS	1.4±0.09
(N)ITH(V)360-362	1.7±0.3	(S)GDTSpPR(H) 401-406 PHOS	1.7±0.28
(I)THVP(G) 361-364	1.7±0.3	(S)STGSI(D) 413-417	1.7±0.22
(G)GGNKKIETH(K)366-374	1.8±0.3	(G)SIDMV(D) 416-420	2.0±0.12
(G)NKKI(E) 368-371	1.1±0.1	(S)IDMVD(S) 417-421	2.6±0.04
(N)KKIET(H) 369-373	2.0±0.4	(Q)/LATLADEVSASL/(A) 425-436	3.4±0.28
(E)/THKLTF/(R) 373-378	1.9±0.3	(L)/ATLA/(D) 426-429	1.0±0.02
(V)YKSP(V)394-397	1.5±0.2	(L)/ATLADEVSA/(S) 426-434	1.5±0.23
(P)VVSGDTSPRHLSNV(S) 398-411	2.3±0.25	(L)/ADEVSASLAKQGL 429-441	1.8±0.06
(P)VVSGDTSPRHLSNVSSTGSI D(M)398-418	4.6±2	(A)/DEVSA/(S) 430-434	1.6±0.13

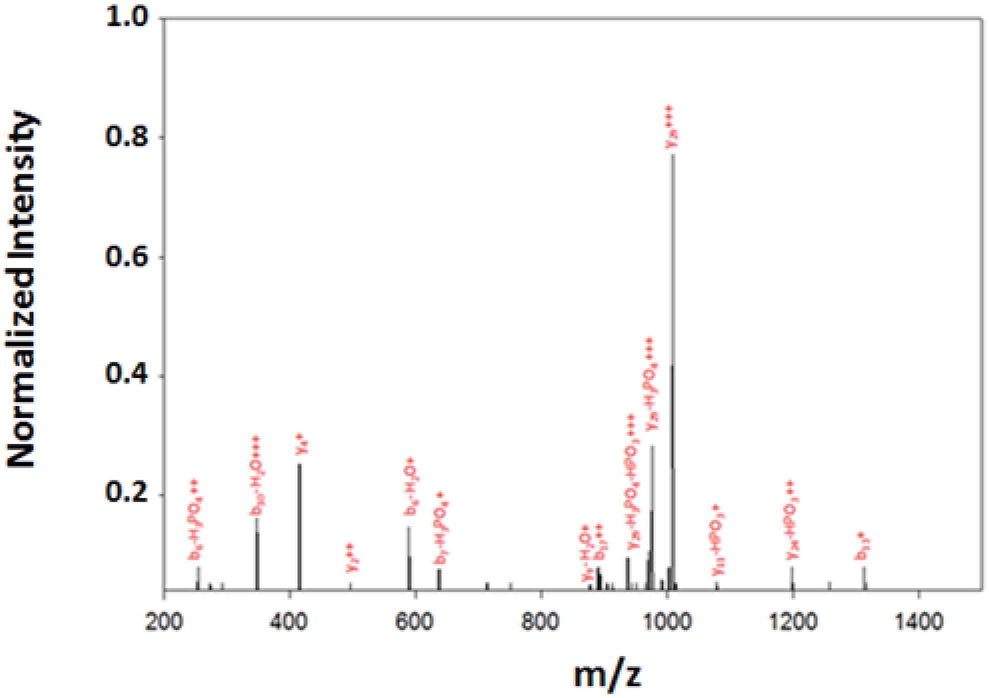
(H)LSNVSSTGSIDM(V) 408-419	2.8±0.17	(E)/VSASL/(A) 432-436	1.8±0.12
(L)/SNVSSTGS(I) 409-416	5.4±0.32	(E)/VSASLA/(K) 432-437	4.7±1.4
(D)SPQLATLADE/(V) 422-431	1.5±0.3	(L)/AKQGL 437-441	2.5±0.03
(T)LADEVSASLAKQGL 428-441	17.5±0.9		
(L)/ADEVSASLAKQGL 429-441	9.4±0.5		
(E)/VSASLA/(K) 432-437	10.1±1		
(A)/SLAKQGL 435-441	11.7±2		

* Ambiguous phosphorylation sites are marked (X)p? that could not be conclusively identified by MS/MS. Multiple phosphorylation sites are labeled as (PHOS *n*X) where *n* represents the number of phosphorylations on the observed peptide.

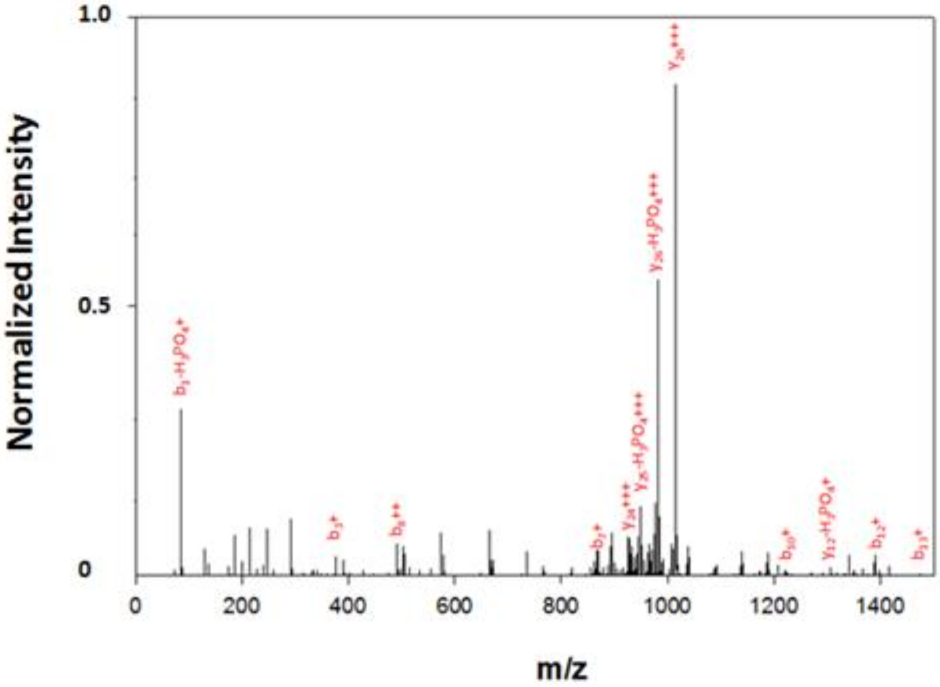
AKGADGKTpKIATpPR (142-155)



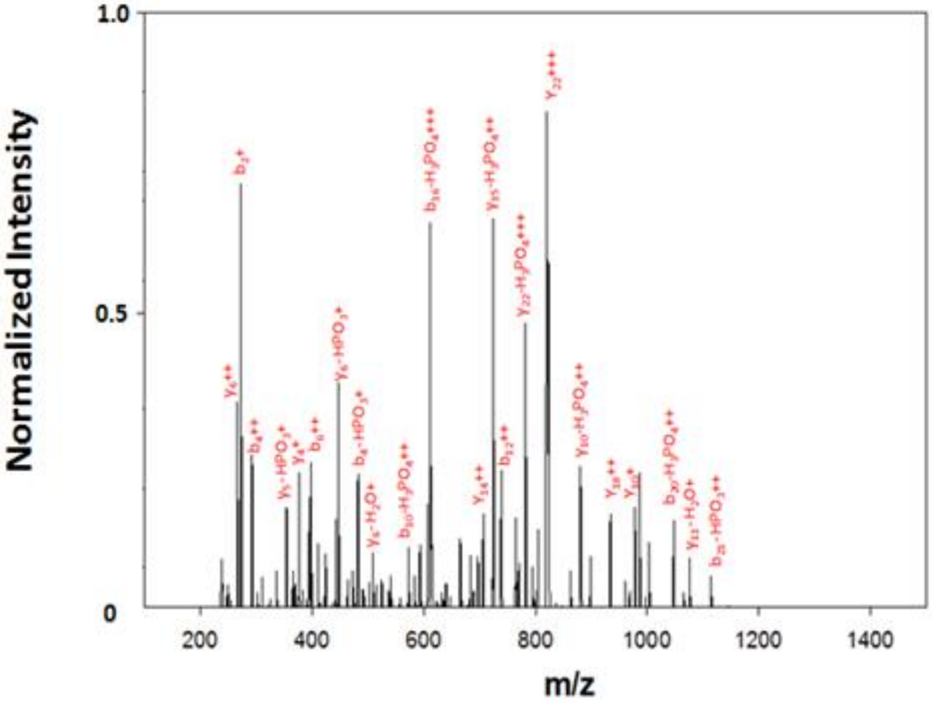
TPPSpSGEPPKSGDRSGYSSpPGSpPGTPGSR (181-209)

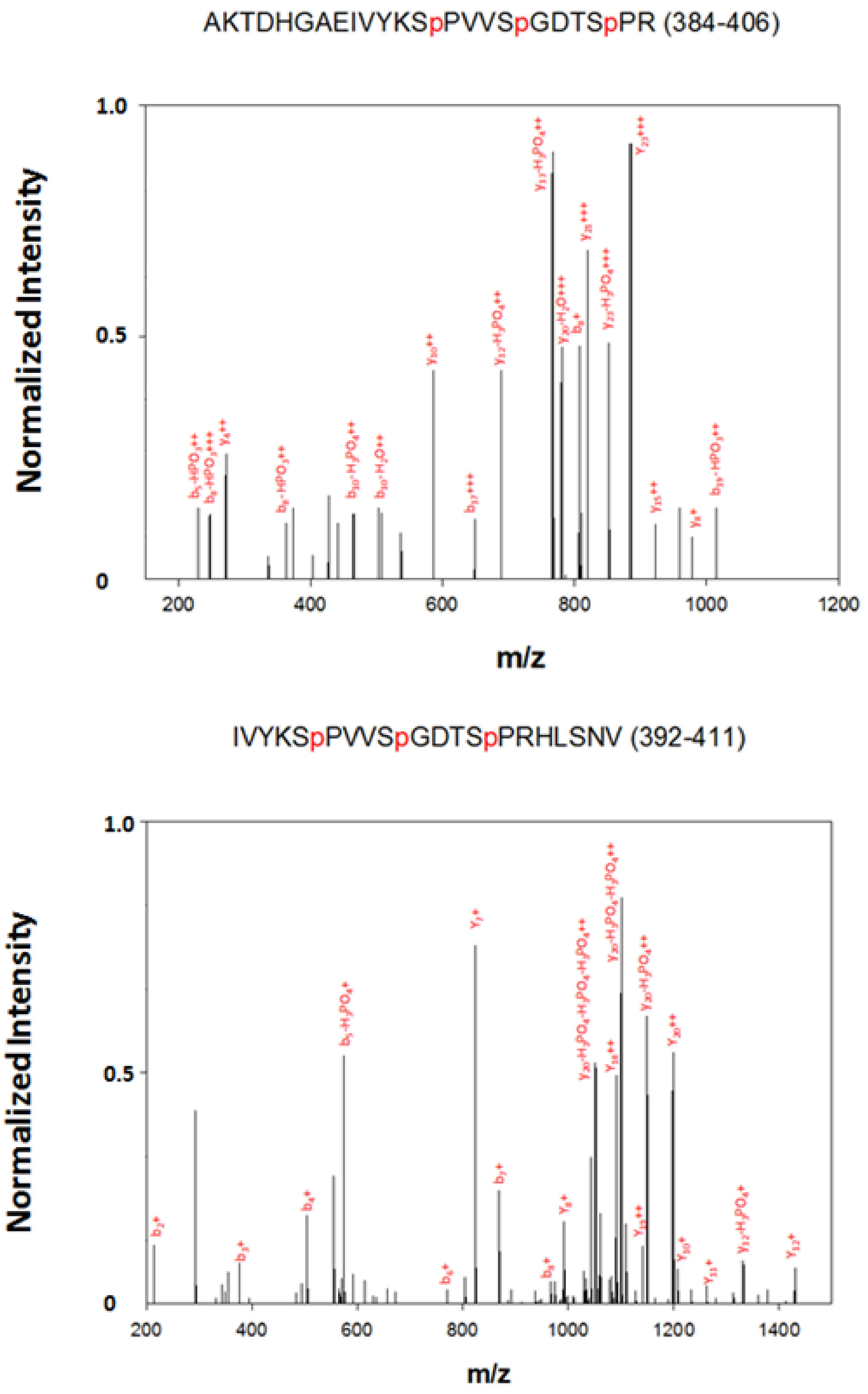


TpPPTp?REPKKVAVVRTpPPKSp?PSp?SAKSp?R (217-242)



DRVQSpKIGSpLDNITHVPGGGNK (348-369)




AKTDHGAEIVYKSpPVVSpGDTSpPR (384-406)

IVYKSpPVVSpGDTSpPRHLSNV (392-411)

Figure B1. MS/MS spectra for phosphorylated tau peptides. Adapted from Zhu *et al*[70].

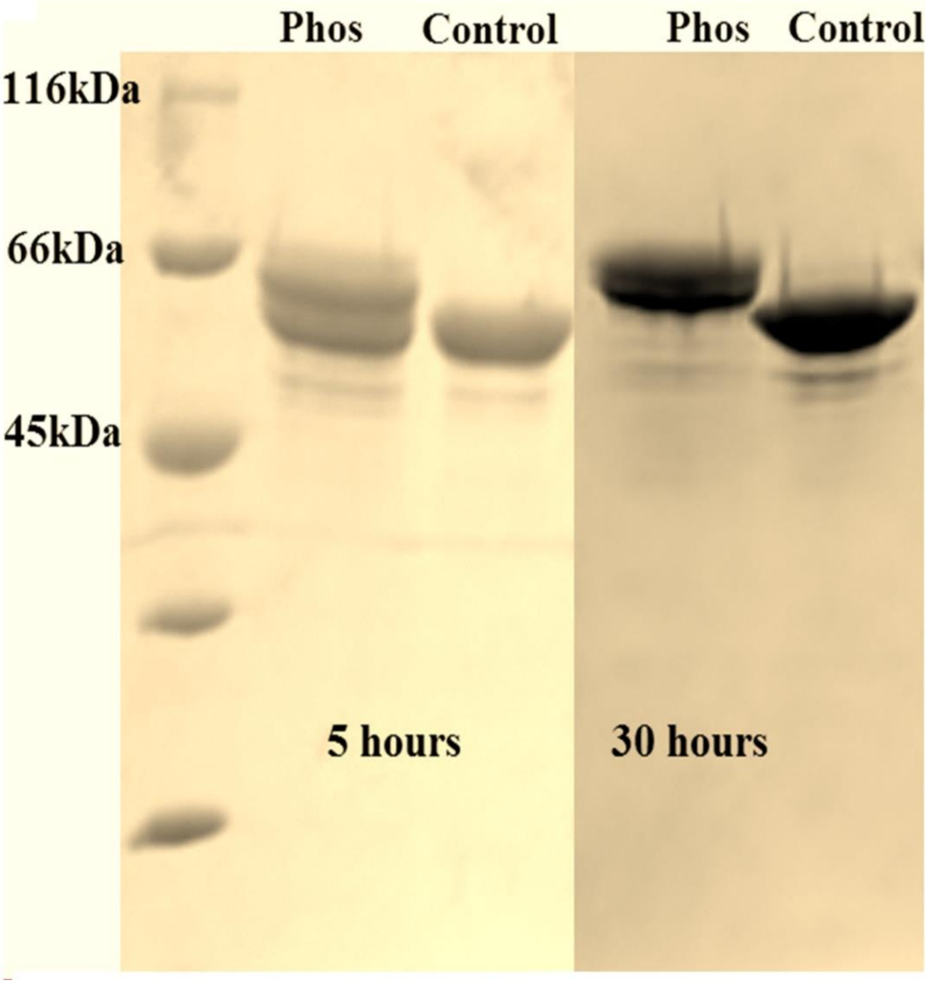


Figure B2. Denaturing SDS gel showing shift of phosphorylated tau (5 and 30 hours incubations) to higher mass. Adapted from Zhu *et al*[70].

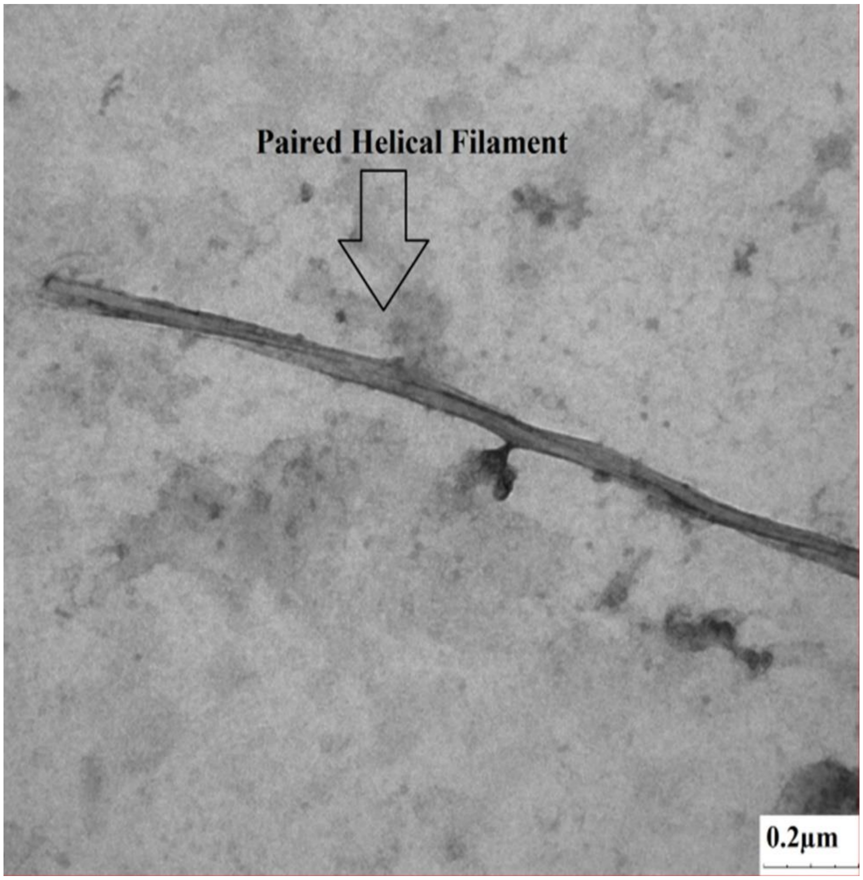
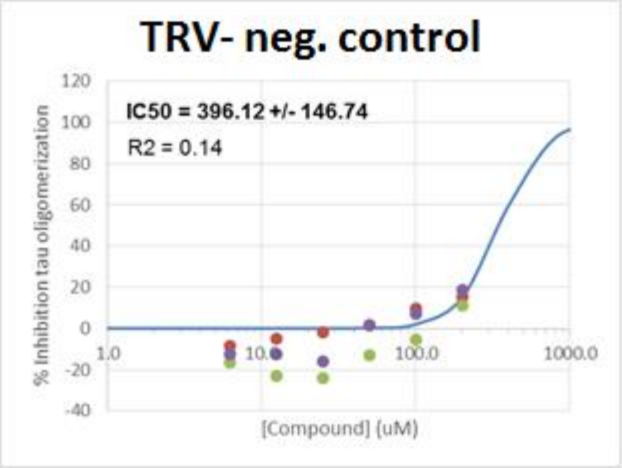
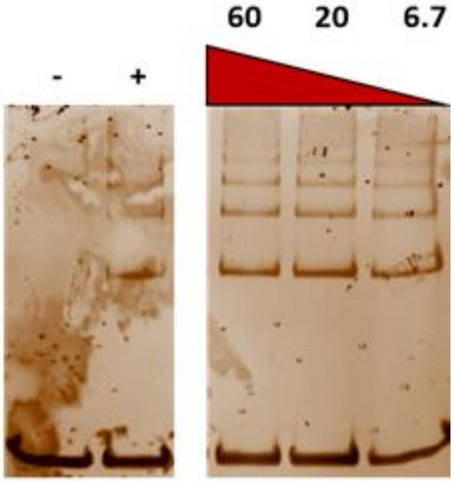
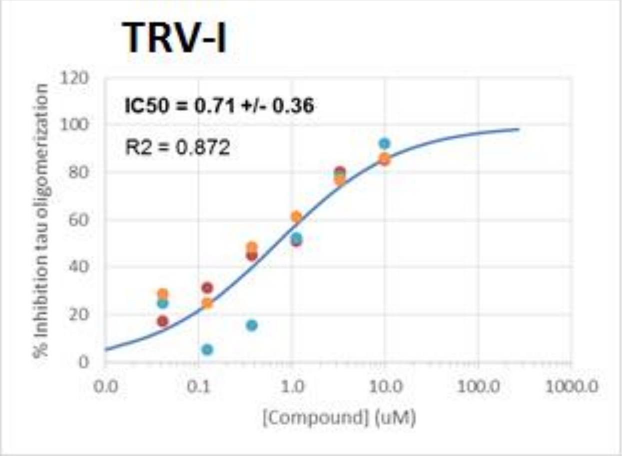


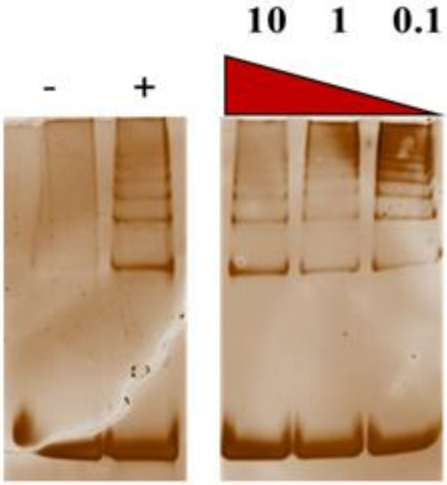
Figure B3. TEM image of PHF at 56 hours after tau hyperphosphorylation. Adapted from Zhu *et al*[70].

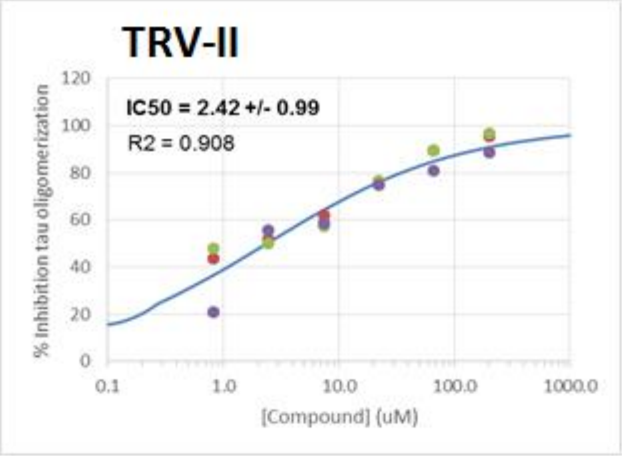
Appendix C: Supplementary informations for chapter 4

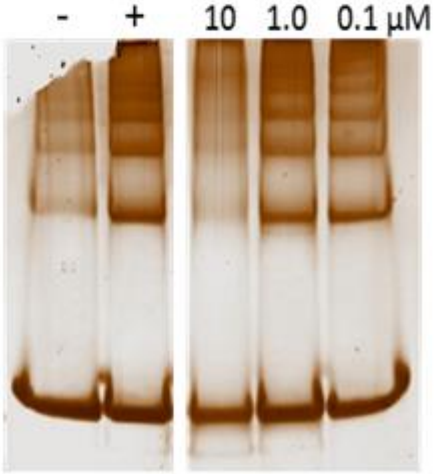


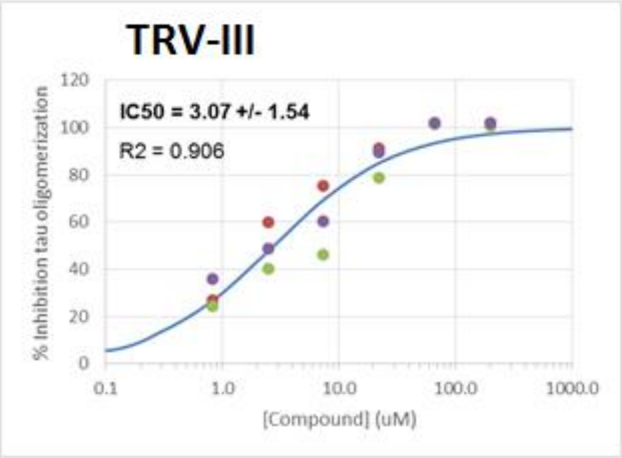


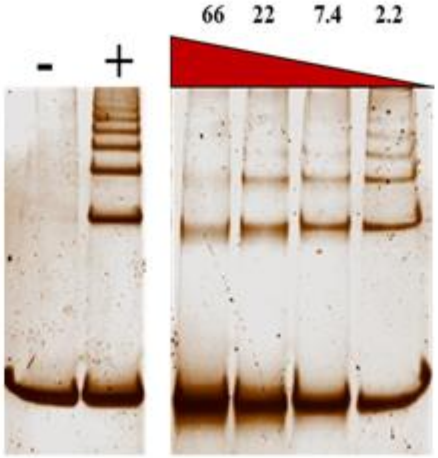


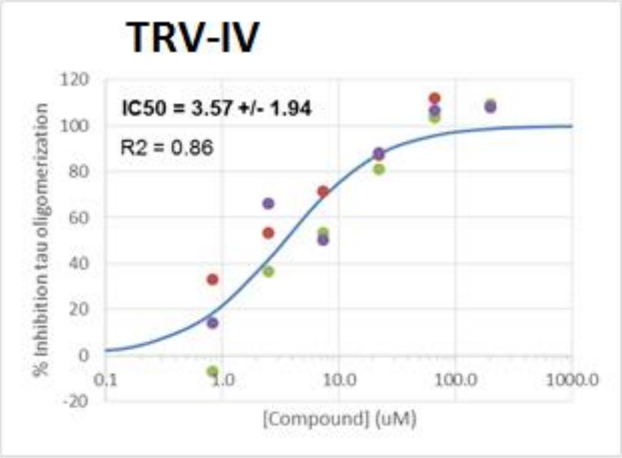


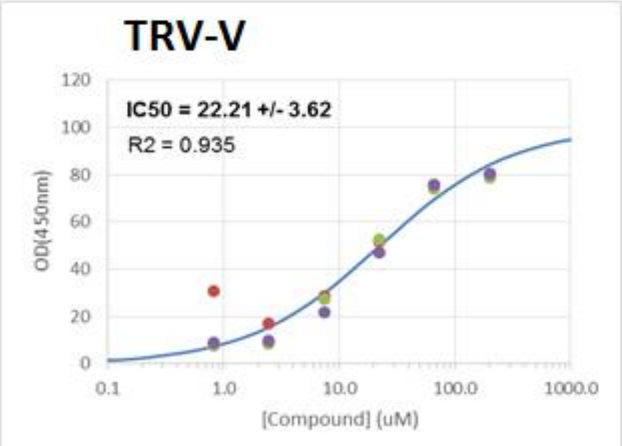


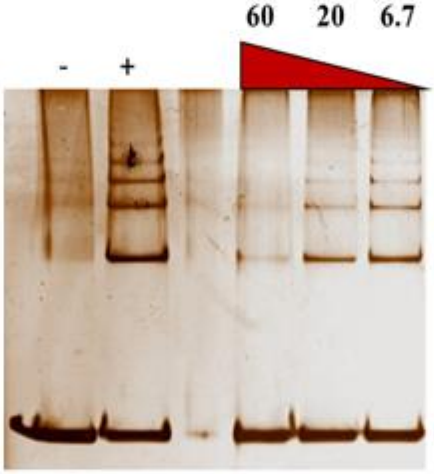



TRV-III
IC50 = 3.07 +/- 1.54
R2 = 0.906
% Inhibition tau oligomerization
120
100
80
60
40
20
0
0.1
1.0
10.0
100.0
1000.0
[Compound] (uM)


66
22
7.4
2.2
-
+


TRV-IV
IC50 = 3.57 +/- 1.94
R2 = 0.86
% Inhibition tau oligomerization
120
100
80
60
40
20
0
-20
0.1
1.0
10.0
100.0
1000.0
[Compound] (uM)


TRV-V
IC50 = 22.21 +/- 3.62
R2 = 0.935
OD(450nm)
120
100
80
60
40
20
0
0.1
1.0
10.0
100.0
1000.0
[Compound] (uM)


60
20
6.7
-
+

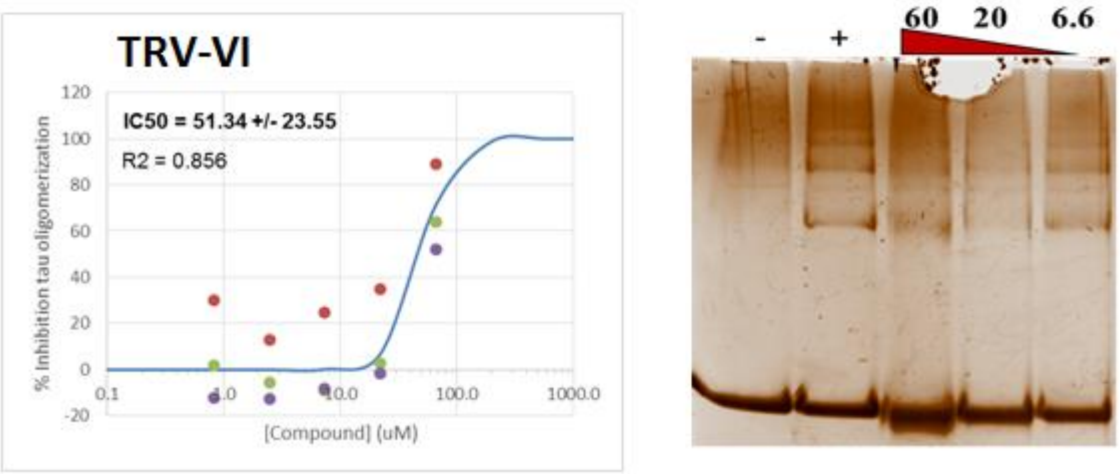


Figure C1. (Left) IC$_{50}$ value deduced from triplicate ELISA runs (each colour represents a single run). (Right) Silver staining gel. Negative and positive signs represent negative and positive control lanes respectively and adjacent lanes show the protein/aggregate migration lanes with their corresponding spiked drug concentrations in µM. The data was provided by Treventis.

Appendix D: Supplementary informations for chapter 5

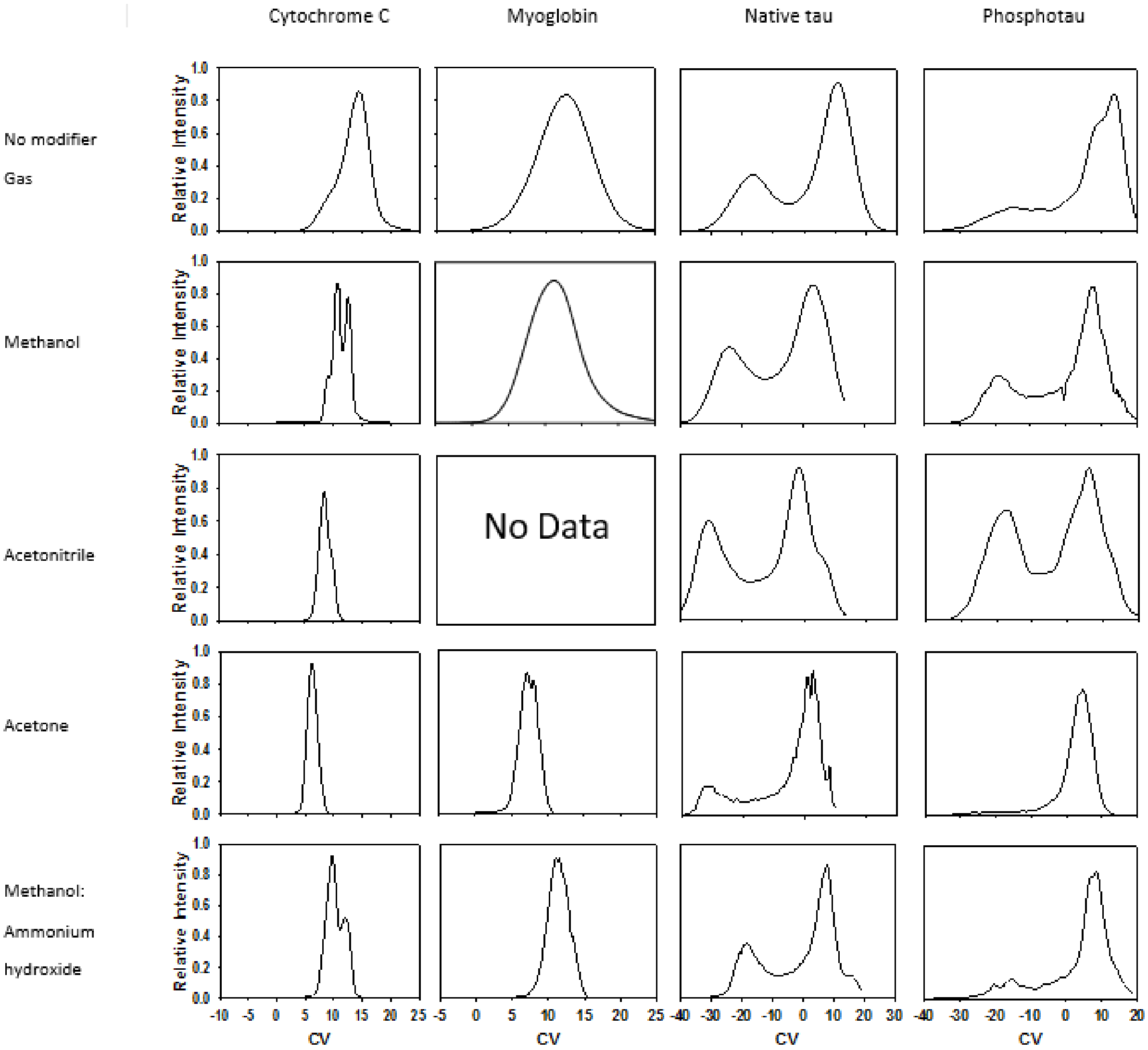


Figure D1: DMS ionograms of cytochrome c, myoglobin, native tau, and phosphotau under influences of different modifier gas conditions. Adapted from *Zhu et al*[228].

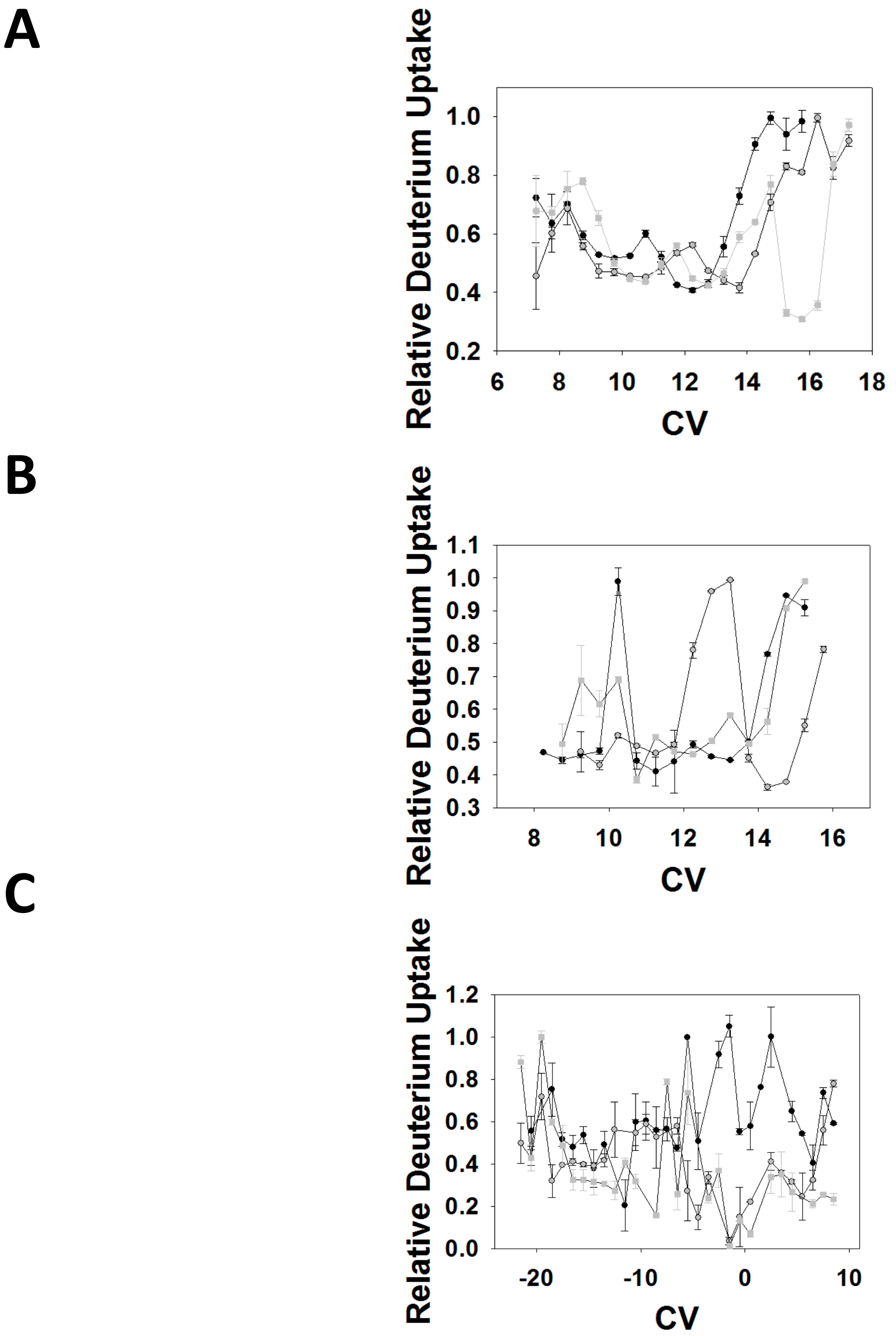


Figure D2: DMS-HDX profiles using labeling through resolution gas only. In panel A, native cytochrome c is labeled with D_2O. Charge states, +9 (filled circle), +10 (grey circle), and +11 (filled grey square) are illustrated. In panel B, unfolded cytochrome c is labeled with D_2O. Charge states +8 (filled circle), +9 (grey circle) and +10 (filled grey square) are illustrated. In panel C, phosphotau is labeled with ND_4OD. Charge states +36 (filled circle), +38 (grey circle), and +40 (filled grey square) are shown. Adapted from *Zhu et al*[228].

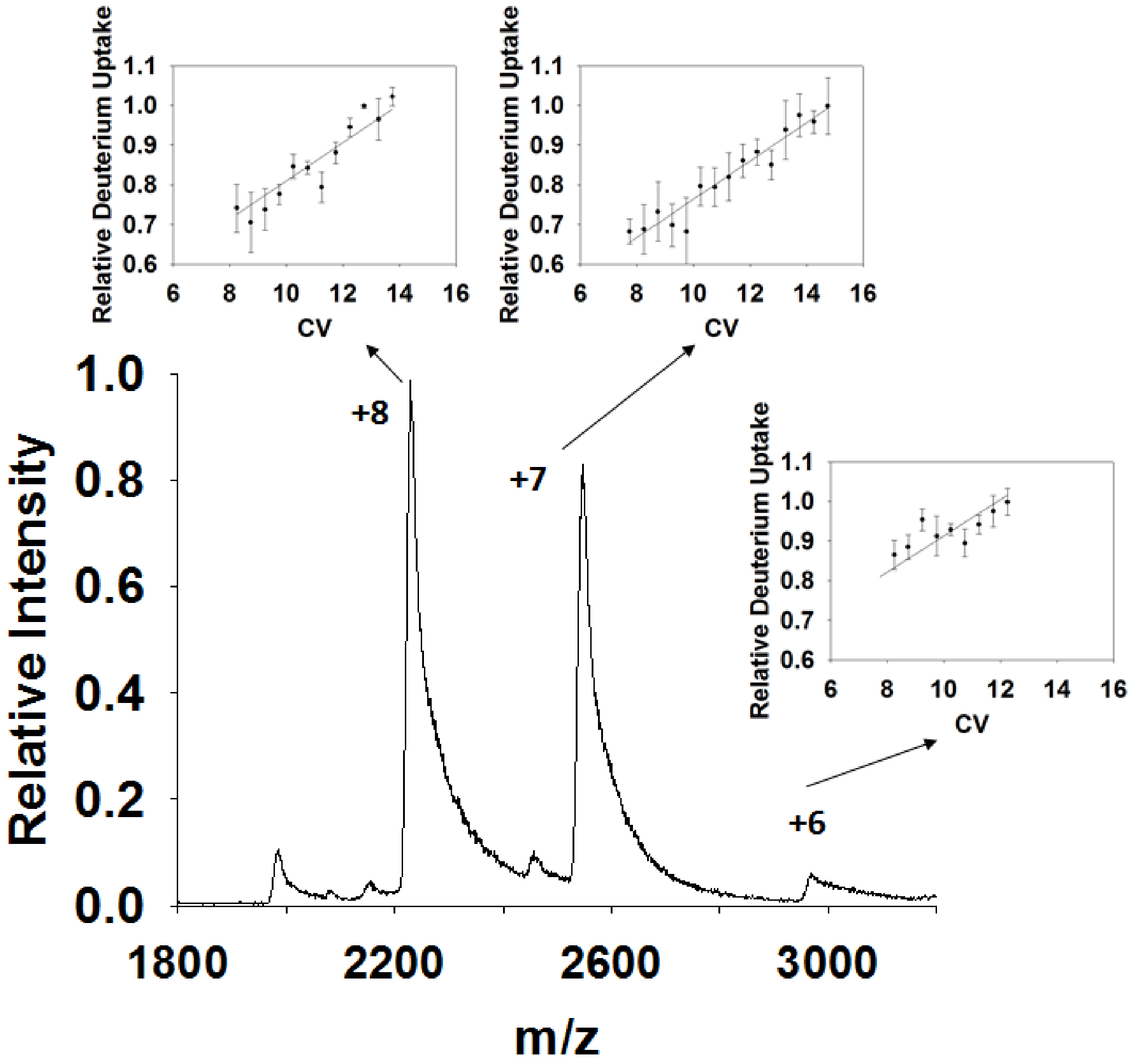


Figure D3: Mass spectrum of native myoglobin with insets showing DMS-HDX profiles of respective charge states labeled using modifier gas ND_4OD:MeOD and throttle gas ND_4OD. Adapted from *Zhu et al*[228].