Elucidating the Conformational Dynamics of YTH Domain Containing Family of Proteins

Iqra Tahir

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

GRADUATE PROGRAM IN CHEMISTRY
YORK UNIVERSITY
TORONTO, ONTARIO

December 2025

Abstract

Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) has emerged as a powerful and time-efficient technique for capturing 'breathing motions' and intricate conformational dynamics of proteins. Here, we conduct HDX-MS to elucidate the conformational dynamics of the YTHDF proteins. The function of the three YTHDF proteins has been debated in the scientific community. According to the prevailing model, the three YTHDF proteins (YTHDF1, YTHDF2, and YTHDF3) have different functions, while the unified model proposes shared functions. We aimed to detect whether the observed conformational dynamics between the three YTHDFs are redundant or different for the fate of the m^6A-modified RNA ligand. We observed that all three YTHDFs showed significant stabilization in the bound state relative to the unbound state. However, the conformational dynamics of YTHDF2 were most different, particularly in the α2–β3 region. HDX-MS was useful in detecting unique conformational dynamics in areas influenced by allostery, such as the α2–β3 region.

Acknowledgements

I acknowledge Professor Derek Wilson for his supervision and for the opportunity to study protein conformational dynamics using HDX-MS. I acknowledge Professor Ryan Hili for his collaboration and for his brilliant idea of a very exciting project for me. I also acknowledge Professor Phillip Johnson for bringing his support into my advisory committee as well as Professor Satinder Brar for serving as an internal external. I acknowledge the past members – Esther and Shalong for their training videos on DynamX, which helped me with my learning curve; as well as the present members of the Wilson Group – Alex for offering insights into improving natives of my protein, Alice for her friendship during the Trent conference and for new insights about using PLGS and DynamX, Ayesha for teaching me to automate CFIT, Joe for answering my EAD related questions for my MS course project, Cristina for trainings on cyclic instrument, and Yasaman from Hili lab for her friendship and for teaching me about protein work. I am grateful for my time at YorkU, which helped shaped me into a stronger and more resilient researcher. I am thankful to my family and friends for their unwavering support.

Table of Contents

List of Figures

List of Tables

List of Symbols, Abbreviations, or Nomenclature

σ	standard deviation propagated error
CCS	collisional cross section
CEM	chain ejection model
CID	collision-induced dissociation
CRM	charge residue model
DC	direct current
DMA	differential mobility analyzers
DMSO	dimethyl sulfoxide
DT-IMS	drift time ion mobility
DTT	dithiothreitol
EI	electron impact
eIFs	translation initiation factors
ESI	electrospray ionization
FAB	fast atom bombardment
FMRP	Fragile X Mental Retardation Protein
FT-ICR	fourier transform ion cyclotron resonance
FXS	Fragile X Syndrome
GST	glutathione-S-transferase
HDX	hydrogen deuterium exchange
HTS	high-throughput screening
IDRs	intrinsically disordered regions
IEM	ion evaporation model

IMS	ion mobility spectrometry
k_{ch}	rate of chemical exchange
K_D	dissociation constant
k_{HX}	rate of HDX
k_{int}	intrinsic rate
k_{on}/ k_{off}	ligand association/dissociation rates
k_{open}/ k_{close}	rate of opening/closing
m/z	mass-to-charge ratio
m^6A	*N6*-methyladenine
MALDI	matrix-assisted laser desorption
MCP	multiple-channel plate
MD	molecular dynamics
MS	mass spectrometry
MW	molecular weight
NMR	nuclear magnatic resonance
P-bodies	processing bodies
PDLA	peptide poly-DL-alanine
PLGS	Protein Lynx Global Server
Q-TOF	hybrid quadrupole and TOF mass analyzer
RF	radiofrequency current
RMSF	root mean square fluctuations
SAC	Salvianolic Acid C
SDS-PAGE	sodium dodecyl sulfate polyacrylamide

SEC	size exclusion chromatography
SGs	stress granules
ToF	time-of-flight mass analyzer
TW-IMS	travelling wave ion mobility
YTH	YT521-B homology (YTH) domain
YTH1	YTH domain of YTHDF1
YTH2	YTH domain of YTHDF2
YTH3	YTH domain of YTHDF3
YTHDF	YTH domain-containing family of proteins

Chapter 1 | Introduction

1.0 Introduction

Hydrogen-Deuterium Exchange (HDX) has emerged as a powerful and time-efficient technique for elucidating protein dynamics and interactions[1–3]. Methods such as x-ray crystallography can provide atomic-level resolution of proteins; however, it is time-intensive and not well-suited for proteins containing intrinsically disordered regions (IDRs) – which are present in ~1/3 of all proteins and play numerous roles in driving biological activities, such as participating in protein-protein interactions[4]. In addition, crystals only capture a 'snapshot' of proteins[1–3]. Proteins demonstrate 'breathing motions' and undergo intricate conformational dynamics that can be captured with HDX in the protein's native environment – i.e. catching proteins in action, making HDX especially relevant in biophysical analysis. HDX was initially complemented with nuclear magnetic resonance (NMR); however, it posed limitations on the size of the protein that could be studied. The implementation of HDX with mass spectrometry (MS) has revolutionized this field; MS and HDX are briefly introduced below.

1.1 Mass Spectrometry

The history of mass spectrometry dates to the 1870s, when Eugen Goldstein observed that negative particles in a ray moved towards the anode (+) and away from the cathode (-)[5,6]. Later, to deflect these rays, Professor J.J. Thomson's work involved the use of electric fields[5,6]. Francis William Aston developed the first mass spectrometer – this design included both the electric and magnetic fields, referred to as a sector, and final masses were plotted on a photographic plate[5,6]. At this time, the applications of MS were limited to the identification of isotopes of stable elements. Since then, the use of mass spectrometry has expanded to various fields (e.g.,

atmospheric chemistry, biochemistry, and applications extending to space) due to its *sensitivity* and *selectivity* – different species can be *detected* and *distinguished*, respectively[5,6].

1.1.1 Ionization Techniques

To generate ions, hard or soft ionization techniques may be used[6]. In **hard ionization** techniques, ions are generated through the transfer of high energy[6]. For example, electron impact (EI) is a form of hard ionization technique. In EI, high-energy electrons (i.e. > 60eV) from a hot filament are bombarded such that one electron is lost from the parent analyte (M) as follows: $M + e \rightarrow M^{\bullet+} + 2e$. This technique is highly effective in ionization. However, ionization usually only requires ~10eV of energy; the bombardment by high-energy electrons of more than 60eV results in fragmentation, often making it too difficult to identify parent ions[6]. Hard ionization techniques can be useful for identifying small molecules.

In comparison, **soft ionization** techniques, as their name suggests, are much gentler, making them more suitable for exploiting protein-ligand and protein-protein interactions[6]. Fast atom bombardment (FAB) is an example of a soft ionization technique where neutral gas, such as Xenon, is used (instead of electrons) for bombardment to generate positive ions[6]. One of the limitations of this technique is that the collision of the fast atoms often disrupts the matrix, resulting in matrix peaks and excessive noise at lower m/z values[6]. Matrix-assisted laser desorption (MALDI) has been a promising soft ionization technique. Aromatic acids are often selected for the matrix and mixed with the analyte, and ions are generated through laser desorption, using IR or N_2 lasers[6]. In MALDI, charge (z) = 1; due to this, the technique may not be ideal for large proteins,

as the m/z range in modern MS instruments is often limited[6,7]. MALDI has predominantly been replaced with electrospray ionization.

Electrospray ionization (ESI) was proposed in 1968 by Michael Dole and was first described in the context of mass spectrometry in 1984 by John Fenn in the paper, "Electrospray Ion Source – Another Variation on the Free Jet Theme?" (Nobel Prize in Chemistry)[6,8]. In ESI-MS, solution-phase analytes are transferred to gas-phase ions[7]. The analyte is first prepared or buffer-exchanged in solvents such as ammonium acetate, which is ideal for ESI-MS, as NH_3 is volatile and does not tend to form adducts (protein can gain a proton from NH_4^+)[7]. The sample is flowed through a high-voltage capillary (**Figure 1.1**). It is accepted that the sample solvent tends to be positively charged (i.e. partly due to the presence of NH_4^+). Surface tension and Coulombic repulsion lead to the formation of a Taylor cone at the tip of the capillary[6,7]. The applied voltage must be optimized to ensure the formation of a stable Taylor cone. From the Taylor cone, initial parent droplets are produced. The positive charge resides on the periphery of the droplets. The parent droplets continue to shrink, and as they do, the effects of surface tension and Coulombic repulsion further accelerate the shrinkage. Following solvent evaporation, Columbic repulsion builds until the Rayleigh limit is reached – at this point, a jet fission event occurs, and offspring droplets are generated from the parent droplet. The offspring droplets continue to shrink until a gas-phase ion is generated. The initial droplet size tends to be in the μm range, and the final droplet size is in the nm range. In Nano-ESI, the capillary is narrower, generating smaller initial droplets. The mechanism by which ionization occurs in ESI-MS has been debated. We quote from Professor Konermann et al. (2012): "Uncovering the intricacies of the ESI process has proven to be surprisingly challenging and remains an active area of research."[7] Currently, ionization with ESI

is proposed to occur through the ion evaporation model (IEM), charge residue model (CRM), or chain ejection model (CEM).

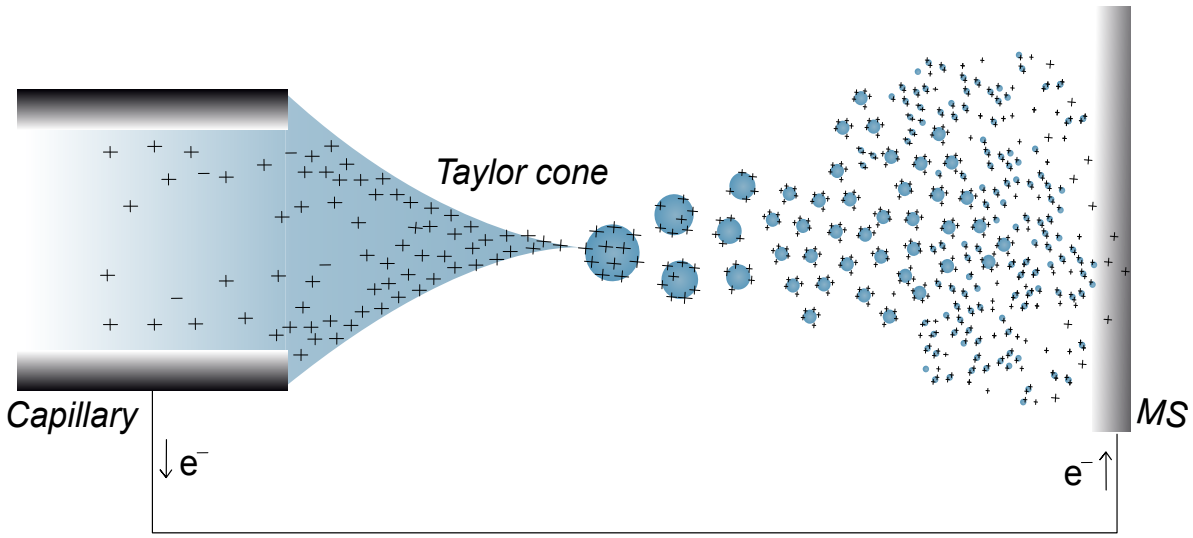


Figure 1.1 | Electrospray Ionization Process
The Taylor cone is shown at the end of the capillary. Solution is in blue. Droplets are shown to shrink until gas-phase ions reach the MS.

IEM, CRM, and CEM are nicely summarized in work by Konermann and coworkers[7]. Briefly, the **IEM** model postulates that the polar solvent surrounds the positively charged analyte (first step). Due to charge repulsion, the solvated molecule is then separated from the initial droplet, creating a 'solvent bridge' (second step). Finally, collision by background gases facilitates the final ejection (third step). Hydrophobic compounds reside close to the droplet periphery and generally have more efficient ejection trajectories. IEM is relevant for small molecules which acquire one charge; however, it fails to capture why proteins have multiple charges. **CRM**, which is the simplest model to comprehend, proposes that proteins acquire all the charge from the final droplet

(i.e. charge residue). Charges are acquired independently of the charge the protein initially carries. One limitation of CRM and ESI-MS is the artifact of observing non-specific binding (i.e. ligand and protein may be non-specifically bound in the same droplet as it shrinks). Therefore, interactions that occur in the gas phase may not always accurately represent solution-phase chemistries. Finally, the **CEM** model is relevant to disordered polymers; as the hydrophobic residues are exposed, the polymer is expelled from the droplet. CEM presents three observations in the spectrum, as summarized by Konermann et al. (2013): peak intensity tends to be higher (owing to ns range fast process which gives greater ion counts), the spectrum tends to be wider (likely due to differing sizes of initial droplets leading to different charges on the ejected polymer), and disordered proteins acquire more charges (due to proton migration). **Figure 1.2** summarizes the three models for ionization in ESI.

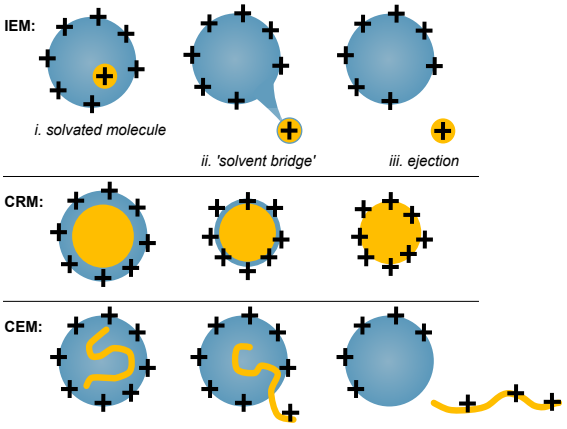


Figure 1.2 | Ionization Models in ESI.
IEM (top), CRM (middle), and CEM (bottom) are shown. The droplet is in blue, the analyte is in yellow. This figure is adapted with modifications from Konermann et al. (2012)[7].

1.1.2 Mass Analyzers

Mass analyzers are an integral component of an MS instrument[6]. Mass analyzers should enable high resolution (i.e., a narrow peak width), accuracy (measured from the peak center), sensitivity (to *detect* species), and selectivity (to *distinguish* between species). The first mass analyzer was a **sector** (i.e, it contained both an electric and magnetic field). In the electric field, ions are filtered based on their velocity, and in the magnetic field, they are focused based on their mass-to-charge ratio (m/z). In a **quadrupole** mass analyzer, alternating direct current (DC) and radiofrequency current (RF) are applied to the four rods. An ion must exhibit a stable trajectory through the quadrupole to be selected. Resolution in a quadrupole is inversely proportional to sensitivity. When both RF and DC are fixed, one ion m/z can be selected. In an ion guide, DC = 0, and only RF is used; this configuration is useful for spatial focusing and collisional cooling. In a triple quadrupole, ions are selected in Q1. Q2 is an RF-only collision cell filled with 'buffer gas'[9] such as nitrogen; it permits collision-induced dissociation (CID), making triple quads compatible with MS/MS. Q3 is used for scanning. In **Time-of-Flight** (TOF) mass analyzers, ions travel based on their m/z. Lighter ions travel faster, and heavier ions travel slower[6]. A 'W' geometry is incorporated with the Waters instruments such as Synapt G2-Si which helps with resolution due to the longer path length. A Fourier transform ion cyclotron resonance (FT-ICR) is best for resolution; ions undergo a high number of oscillations in the magnetic field. RF pulses cause ions to orbit in sync, allowing a detectable image current to be generated when the ions are close to the receiver plate. Cyclotron frequency can be used to calculate m/z. Orbitraps are similar to FT-ICR; however, orbitraps do not have the magnet, and rely on electric fields to oscillate ions. Most MS

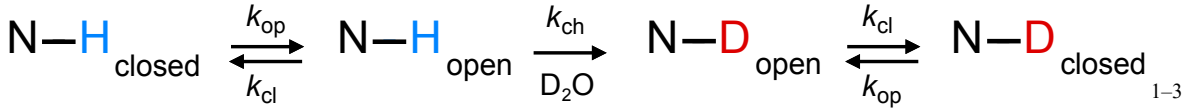
instruments are hybrid; for example, Waters Synapt G2-Si is equipped with both a quadruple and a TOF mass analyzer (Q-TOF).

1.1.3 Ion Mobility Spectrometry

In MS, ions are separated based on their m/z, and in **ion mobility spectrometry** (IMS), ions are separated by their mobility and collisional cross section (CCS)[6,9]. IMS is commonly paired with MS. CCS is impacted by size, charge, and shape, etc. In IMS, ions are propelled through an inert gas (i.e. helium or nitrogen) using electric field. The mobility of large or disordered polymers is hampered by background gases more, so they travel slower in comparison to smaller ions. In drift time ion mobility (DT-IMS), a uniformly and weak electric field is applied, and RF is often not required. DT-IMS allows for direct calculation of the CCS. Travelling wave ion mobility (TW-IMS) is included with Waters instruments such as the Synapt G2-Si. In TW-IMS, an oscillating electric field, including the use of RF, is applied through stacked rings to create 'wave' like motions for the ions to travel through. Because of the use of RF, ion diffusion is prevented, and high signal can be achieved (which is a downside of DT-IMS). In differential mobility analyzers (DMA), an electric field is applied in one direction (i.e. downwards), and gas such as nitrogen is flowed from the perpendicular direction (i.e. sideways). Only ions of a particular size can pass through the exit slit; hence, DMA can act as a filter. IMS is beneficial for application in separating different isomers as well for achieving optimized signal-to-noise. The pros and cons of different IMS methods are summarized in reference [9][9].

1.2 Hydrogen Deuterium Exchange Fundamentals

Hydrogen/Deuterium Exchange (HDX or H/D exchange), as the name suggests, measures the exchange between the two isotopes of hydrogen—protium (1Da) and deuterium (2 Da). Hydrogens (i.e. protium, ^{1}H) bound to heteroatoms – S, O, and N are labile, and can readily exchange with the deuterons (^{2}H) in the solvent. Hydrogens bound to carbons do not exchange. In 1993, the work by Zhang and Smith et al. demonstrated that side chains fully back-exchange at pH 2.5[10]. In a typical HDX experiment, only the exchange at the amide backbone of a polypeptide can be measured. Due to their even distribution and role in the secondary structure of proteins, deuterium uptake on the amide backbone can provide valuable insights into protein dynamics and structure.



[1–3]

The fundamentals of HDX are best summarized by the expression above [1–3]. All hydrogens on the amide backbone of the protein (except proline) can exchange with the solvent. According to the Linderstrøm-Lang model, each N-H will undergo open/close states, that is, the breaking and the forming of the hydrogen bonds (as shown in the expression above) where k_{op} is the rate of opening and k_{cl} is the rate of closing. When the N-H is open, in deuterium solvent, a hydrogen can be exchanged for a deuteron (1 Da heavier); k_{ch} is the rate of chemical exchange, which depends on various factors: temperature, pH, intrinsic properties of amino acids, hydrogen bonding, and solvent accessibility; these factors, as well as different exchange regimes, are discussed below.

Temperature and **pH** impact the rate of HDX[11]. First, increased collisions at higher temperatures, such as 20°C, accelerate the rate of HDX. The labelling step usually proceeds at 20°C. In the unstructured peptide poly-DL-alanine (PDLA), a reduction in temperature from 20°C to 0°C decreased the rate of HDX by 20-fold. The HDX reaction is quenched at 0°C. To prevent back-exchange, post-quench temperatures must be maintained at 0°C; however, temperatures are often elevated to 15°C during the digest step, which results in rapid back-exchange[1–3]. Second, a physiologically relevant pH of 7 is often used during labelling step, and a pH change from 7 to 2.5 effectively quenches the exchange reaction for the backbone amides; the rate of HDX is slowest at pH 2.5[11]. HDX can be base-, acid-, or water-catalyzed; acid and water catalysis made negligible contributions to the PDLA calibrant, and under physiological conditions, most contributions are owed to base-catalysis (**Figure 1.3**). pH can be adjusted differently for pulse versus continues labeling HDX[1–3]. In continuous labelling HDX experiments, which are ideal for studying protein dynamics and interactions, the HDX labelling reaction is conducted at pH ~ 7 through the base-catalyzed mechanism, and the quench step at pH ~2.5. In pulsed HDX reactions, which are useful for studying how proteins undergo folding events when switched from denaturing to native buffer, pulse labelling typically involves more basic pH for rapid labelling. In this work, we conduct continuous labelling of the HDX reaction. Experimental parameters, such as temperature and pH, are important considerations in HDX.

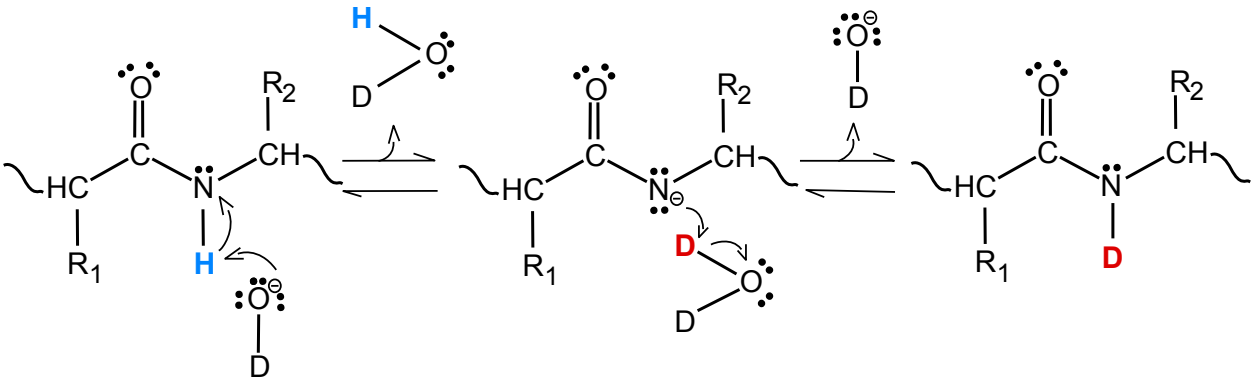


Figure 1.3 | Base-catalyzed HDX[11]
Figure adapted from reference [11][11].

Besides external parameters (i.e., temperature, pH), the **protein's local environment** can also influence the rate of HDX[11]. Bulkier side chains often decrease solvent accessibility and the rate of HDX of the amide in question[11]. The Englander lab has made significant contributions to calculating intrinsic rates in the 1990s[12,13]. Their original work, as well as recent adjustments to these theoretical calculators, can be found in references [12] and [13][12,13]. To calculate the theoretical rate, a fully unstructured peptide, such as PDLA, was used. However, some reports indicated that experimental rates were faster than the theoretical rates, which should not have been possible, as the rate of HDX is expected to be fastest for an unstructured PDLA calibrant[4]. In 2018, Englander's lab addressed this issue and noted that PDLA exhibited some structural propensity[12]. In 2025, Stofella et al. introduced a tri-alanine peptide (3-ALA) as a more suitable calibrant for intrinsic rate calculators[4]. In short, a protein's local environment influences HDX rate, and efforts have been made to calculate these intrinsic rates (k_{int}) accurately.

Both **solvent accessibility** and **hydrogen bonding** are known to influence the rate of HDX[11]. Loops and intrinsically disordered regions of the protein are often more readily accessible for exchange with deuterons in solvent compared to the hydrophobic buried cores[11]. Proteins must

display ‘breathing motions’ and undergo distortions of a few angstroms for exchange events to occur in buried regions. As illustrated by the Linderstrøm-Lang model, the breaking and forming of hydrogen bonds facilitate the open/close transient conformation that enables exchange events. In reference [1-3] by Konermann, it is suggested that even if a region of a protein is solvent-exposed, it will still exhibit reduced deuterium uptake if the region is also involved in hydrogen bonding[1–3]. Due to this, sometimes H/D exchange profiles are interpreted in terms of how ‘intact’ the hydrogen bond network is. While both hydrogen bonding and solvent accessibility play roles, the majority of HDX kinetics are owed to hydrogen bonding. We also note here that to explore protein dynamics, it is recommended to select a wide range of **time points.** Shorter times are ideal for loops or disordered regions, while longer incubation in deuterium is suited for structural regions. Even longer times may be necessary to distinguish two regions with structural propensity.

The chemical exchange rate (k_{ch}— which, as discussed above, depends on temperature, pH, side-chain amino acids, etc.) and the timescale of the opening and closing rates (k_{op} and k_{cl}) can be related to elucidate the exchange regime of amides[11]. **EX2** exchange regime occurs when $k_{cl} >> k_{ch}$ – rate of closing is substantially faster than the rate of chemical exchange. Proteins predominantly prefer to remain close/folded and only briefly visit the open/unfolded/excited state; making EX2 the most relevant regime in most cases. As the probability of an exchange event is low per open cycle, the binomial isotopic distribution of the peptide slowly drifts to the right over time (**Figure**). In contrast, **EX1**, although less common, can also be observed when $k_{cl} << k_{ch}$ – the rate of closing is slower than the rate of chemical exchange. This typically presents a bimodal peak because complete exchange occurs for amides in just one open cycle, while closed amides have not undergone any exchange at all (**Figure 1.4**).

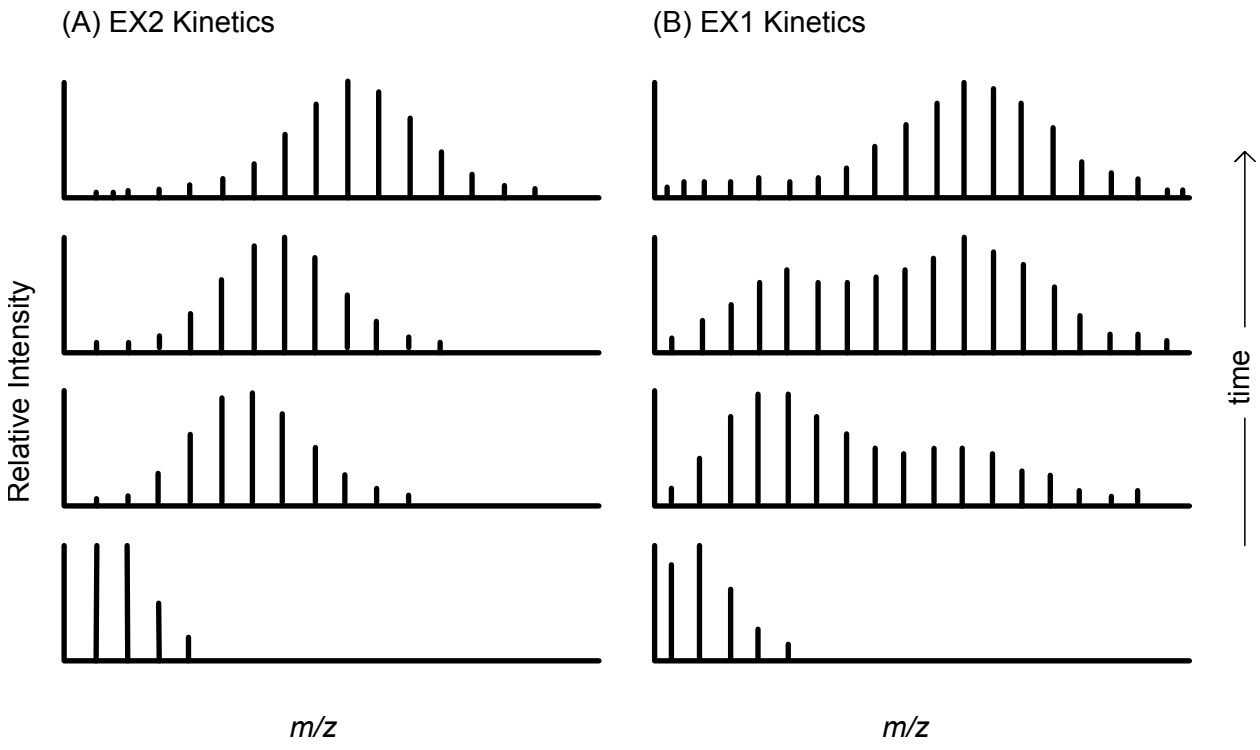


Figure 1.4 | EX1 and EX2 Kinetics[11]
Figure adapted from reference [11][11].

1.3 Types of HDX Experiments.

1.3.1 Top & Middle Down

Top-down proteomics with electron-activated dissociation offers enhanced spatial resolution[1–3]. Major back-exchange is mitigated because protein digestion and LC separation steps are omitted. Instead, intact proteins are fragmented. Initially, CID was incorporated with top-down proteomics. However, this presented various caveats: (1) unlike peptides (in bottom-up workflows), fragmentation of intact proteins did not follow predictable patterns; (2) CID is not compatible with post-translational modifications (PTMs) because in CID, the weakest bond between the amino acids breaks, and PTMs are even weaker and cannot be retained; and most importantly, (3) slow-heating methods such as CID lead to 'scrambling' which enables migration

of protons/deuterons and loss of information. To overcome these challenges, electron-activated dissociation (EAD) is complemented with top-down proteomics[1]. EAD is a non-ergodic fragmentation technique, and the N-Cα bond breaks to generate c and $z^{\bullet}$ ions[14]. Briefly, a multiple-charged precursor can capture an electron to generate a charge-reduced species and $c/z^{\bullet}$ ions; multiple electron capture events can occur (**Figure 1.5**)[14,15]. Two common types of EAD are: electron transfer dissociation (ETD) and electron capture dissociation (ECD). In ETD, an electron is transferred from a radical anionic reagent, and in ECD, electrons are directly captured (e.g. electrons emitted from beam). The Sciex ZenoTOF 7600 System is compatible with ECD for top-down proteomics[16].

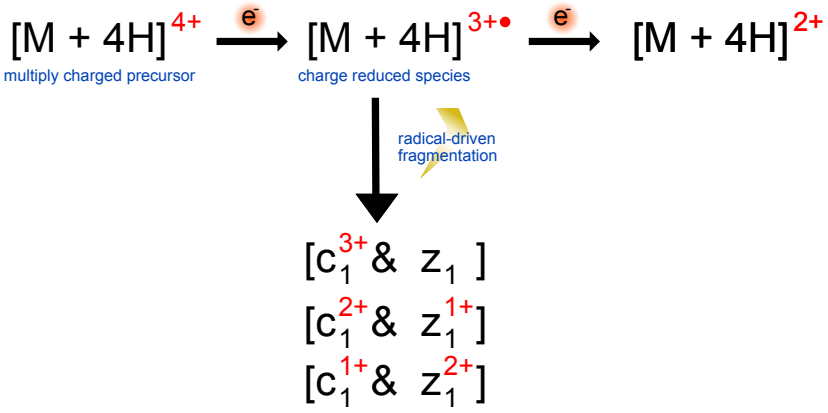


Figure 1.5 | Electron Capture Dissociation
Figure designed using ideas from personal communication with Dr. Joseph Anacleto[15].

The instrumental configuration of the Sciex ZenoTOF 7600 System is shown in **Figure 1.6**[16]. Precursors are selected in Q1. In the EAD cell, precursors can capture electrons emitted from the electron beam. Notably, this design enables the formation of an 'electron cloud' in the EAD cell, which significantly enhances capture efficiency. This facilitates effective EAD without reliance on expensive instruments such as FT-ICR (initially, EAD applications were

primarily limited to FT-ICR instruments)[14]. Sciex also offers a Q2 CAD cell orthogonal to the EAD cell[16]. EAD or CAD or both can be conducted with this design. The incorporation of both EAD and CAD has been shown to provide optimal coverage. Finally, generated fragments are separated with ToF analyzer, and detected with a multiple-channel plate (MCP) detector.

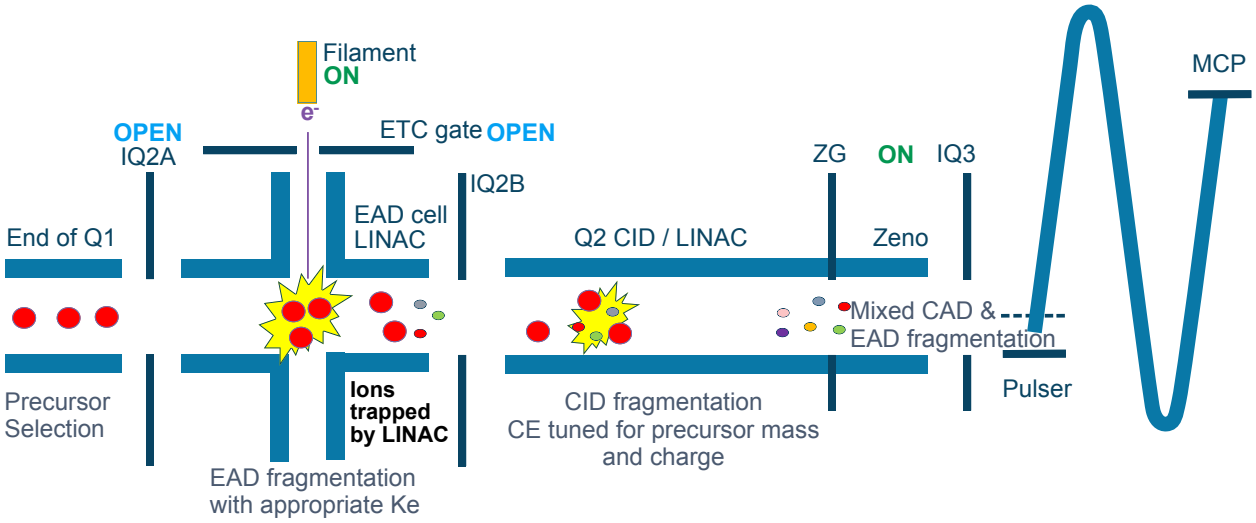


Figure 1.6 | Sciex ZenoTOF 7600 System[16]
Figure adapted with modification from Sciex[16].

Middle-down proteomics is often conducted to overcome the current challenges of top-down proteomics. Top-down EAD workflows are suitable for small proteins[17]. However, fragmentation efficiency is compromised with higher MW proteins. To study larger proteins, middle-down proteomics is often conducted – proteins are digested into larger polypeptides before fragmentation.

1.3.2 Bottom-up

A standard **bottom-up** HDX workflow includes a proteolytic digestion step; peptides are analyzed to decode the dynamics of the full protein (i.e. working bottom → up) [1–3,11]. In continuous

labelling setups, the protein is first incubated in deuterium for labelling (20°C, pH 7). Shorter incubation times are more suitable for disordered regions, while longer incubations are appropriate for structured regions. For enhanced signal-to-noise (S/N), 90% deuterium labelling is generally employed; however, for weaker interactions, 50% labelling may be used. Even 10% D_2O labelling is reportedly effective. We 'freeze' labelling in quench (0°C, pH 2.5); only labelling at amide backbones is expected to be retained, and side chains fully back-exchange at this step. For protein digestion, an immobilized pepsin column is commonly used. Pepsin is an enzyme from the stomach and remains biologically active at low pH. Pepsin digests peptides in a non-specific manner, often helping with redundancy. Generated peptides are separated with reverse-phase chromatography using a C18 column. This is followed by ESI-MS. **Figure 1.7** summarizes the workflow in bottom-up HDX. Collision-induced dissociation is used to generate b and y ions (the weakest bond between amino acids breaks) to identify peptide sequence in commonly used software, Protein Lynx Global Server (PLGS), and peptides are analyzed with DynamX.

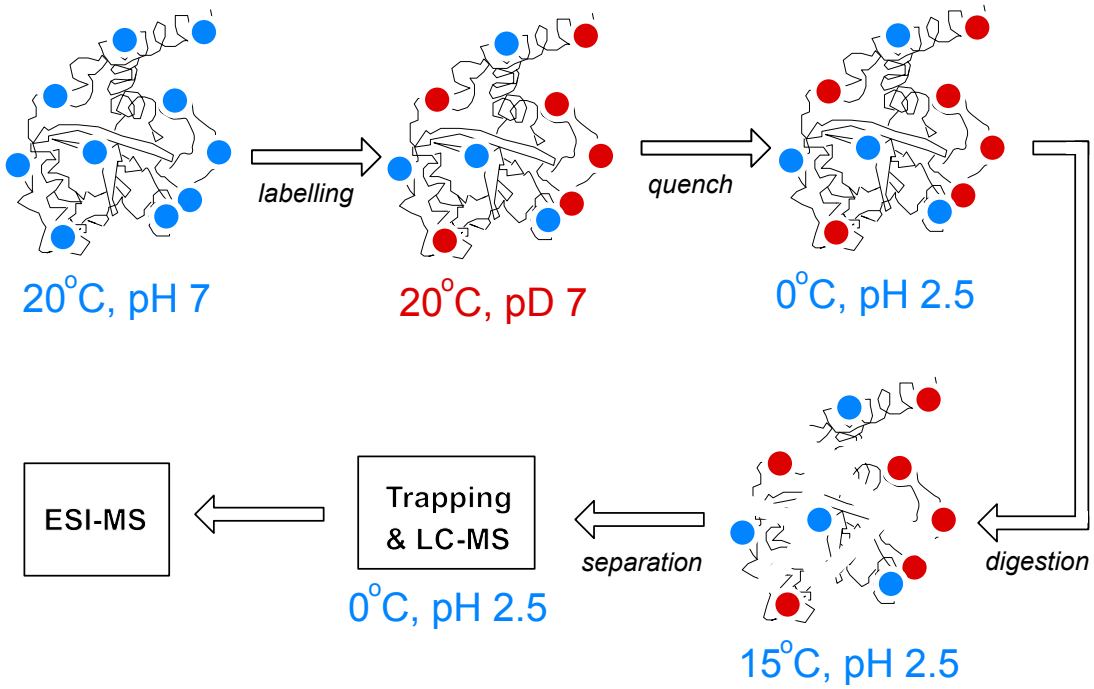


Figure 1.7 | Bottom-up proteomics
Bottom-up HDX workflow designed using commonly presented ideas in references [1-3, 11][1–3,11]. A rough sketch for the YTH protein (PDB: 4RCI) is shown.

Bottom-up proteomics has a fundamental limitation: back exchange. The impact of back exchange is described in work from the Konermann lab in [1-3][1–3]. As discussed earlier, work with PDLA has shown that the rate of HDX (k_{HX}) is slowest at 0°C and pH 2.5. Following labelling steps, attempts are made to maintain these conditions. However, in bottom-up proteomics, proteins are digested into peptides, and during this process, the temperature is raised to approximately 15°C, which causes significant back-exchange. Additionally, the conventional HDX protocol often involves reverse phase separation of peptides for 15-20 minutes using buffers that consist of H_2O in the mobile phase, and most back-exchange occurs during this stage. According to Dr. Konermann, one way to reduce back exchange is to lower the temperature during reverse phase separation to below 0°C. However, under highly optimized experimental conditions, some side chains may not fully back-exchange (which can complicate analysis later). In most setups, all side chains will fully back-exchange, and the first two positions in any peptide will also lose spatial resolution. The limitation of back-exchange can be addressed either through optimized experimental workflows or through back-exchange correction strategies. For the case of Δ HDX, back-exchange does not need to be considered.

1.3.2.1 Two-Valve LC System

In the conventional bottom-up HDX workflow, peptide digest and separation steps are experimentally accommodated with a 6-port 2-valve system (**Figure 1.8**); this is explained in reference [11][11]. The first valve (shown in blue in **Figure 1.8**) facilitates the digest step, and the

second valve (shown in green in **Figure 1.8**) accommodates the peptide separation step. The sample is first injected into port 2 – quenched protein is injected into valve 1, and digested peptides are injected into valve 2. In each case, port 3 leads to waste, port 6 is linked to buffers (pH 2.5), and port 5 connects to either the pepsin or the reverse phase column. During sample loading in valve 1, the sample is injected in port 2, then flows from port 1 to 4; the protein sample fills the sample loop (between ports 1 and 4), and excess is discarded from port 3 (solid arrow lines in valve 1). During the trapping mode, the direction of flow is reversed (dashed arrow in valve 1); and pepsin buffer (port 6) carries the samples from port 4 to 1 and into the pepsin column for digestion. In valve 2, the digested peptides are injected into port 2 → 1 → 4; port 4 is linked to a trap column. In the analytical phase, the direction of flow is reversed (dashed arrow in valve 2); and an increasing acetonitrile gradient is applied (from port 6) to elute peptides from port 4 to 1 and into the reverse phase column for peptide separation. The 6-port 2-valve system is visualized in **Figure 1.8**.

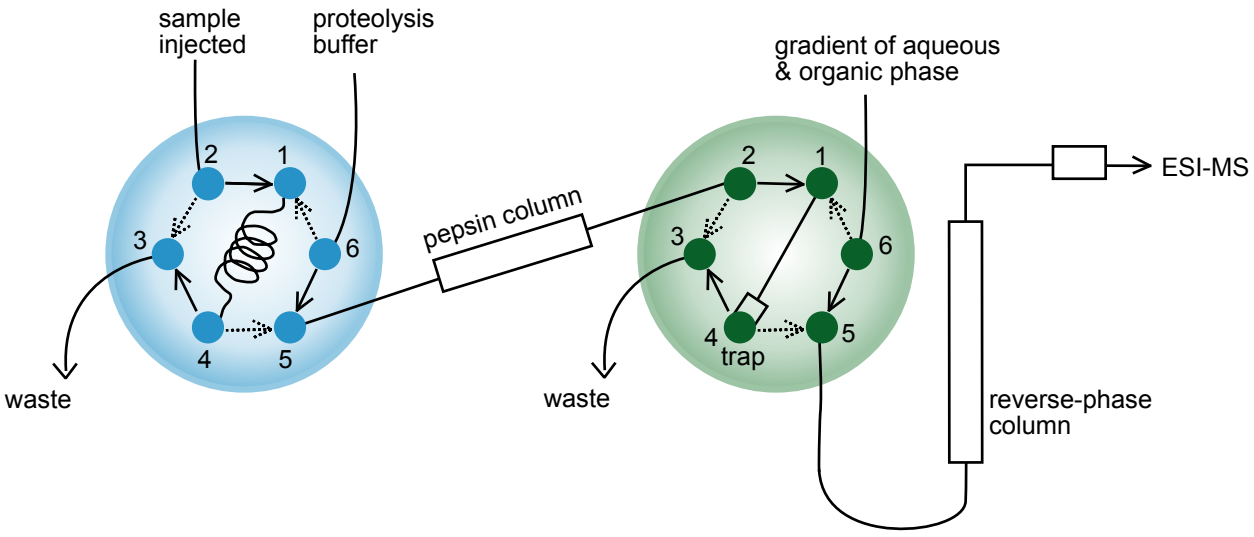


Figure 1.8 | Two-valve LC system
Figure adapted from reference [11][11].

1.3.2.1 Millisecond Timepoints & Online Labelling

Millisecond HDX and its history is summarized by Professor Wilson in reference [11][11]. HDX was initially complemented with NMR; however, amide resolution suffered at millisecond time points[18]. Nonetheless, millisecond HDX-NMR has been attempted, and the quench is flash-frozen to retain labelling information[18]. After the development of ESI-MS, millisecond HDX with MS gained interest[11]. Millisecond timepoints require *online* labelling (instead of offline labelling). Simmons et al. (2013) presented the first report of online labelling for top-down proteomics with EAD[19]. Bottom-up proteomics was still difficult at this time, as the configuration of a pepsin column in an online labelling setup had not been explored[11]. Wilson Group presented the first report of online labelling for bottom-down proteomics; in their setup, the protein sample was mixed with the protease enzyme in a microfluidic chip to incorporate the digest step[20]. Since then, Wilson et al. have made significant advancements to their millisecond workflows; their most recent apparatus (developed by Dr. Joseph Anacleto) is described below[21].

The apparatus for millisecond HDX is described by Dr. Joseph Anacleto and Wilson Group (**Figure 1.9**)[21]. For millisecond timepoints, online deuterium labelling is performed by adjusting the length of the protein line (inner capillary) relative to the deuterium line (outer capillary). The outer capillary can be adjusted to accommodate various sub-second times. This setup employs two Tri-union mixers. For the labelling step, protein and deuterium are mixed in the first tri-union (T-mixer 1), and the deuterated protein is quenched with acid in the second tri-union (T-mixer 2). This

is followed by the incorporation of an immobilized pepsin column (kept in a 15°C chamber) for digestion. The apparatus for online millisecond HDX is visualized in **Figure 1.9**.

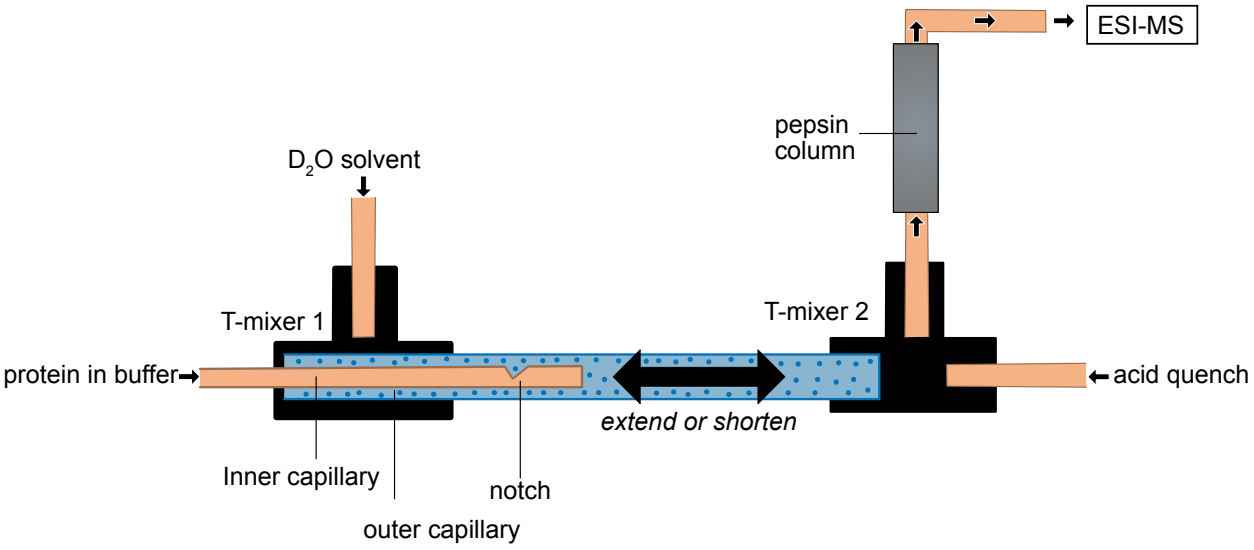


Figure 1.9 | Millisecond HDX workflow
Figure designed using ideas discussed in reference [21] [21].

1.3.3 Δ HDX & Data Analysis/Interpretation

Differential HDX (Δ HDX) is often employed with bottom-up approaches to study perturbations caused by ligands in protein-ligand interactions[1–3,11]. The binding of a ligand to a protein can reduce solvent access in that region (i.e. binding pocket); therefore, relative to the unbound state, the bound state is expected to present reduced deuterium uptake. As discussed earlier, while solvent access does impact H/D exchange, it is not considered the primary factor. Proteins always tend to exhibit 'breathing motion' – whether bound to a ligand or not. However, the thermal motions of the protein in the unbound state may be different from those in the bound state. According to Dr. Konermann, the binding of ligand often reduces the 'occupancy of the open state' of amides, resulting in reduced exchange[1–3]. Ligand binding generally decreases solvent

access to the binding pocket and strengthens the protein's hydrogen bond network. Δ HDX can provide valuable insights into protein-ligand interactions; however, there are at least two caveats: (1) the most significant decreases do not always correlate with the binding pocket, and (2) it is challenging to distinguish the binding pocket from allostery. Nonetheless, an extensive use of HDX has been summarized for the study of ligands, small molecules and drugs in the review by Masson et al. can be found in reference [22][22].

HDX data analysis is summarized by Sarpe and Schriemer in reference [11], and data interpretation by Konnerman in references [1-3] [1–3,11]. The isotopic envelope of peptides is expanded with deuteration, and the centroid (i.e. average mass of peptide) shifts to the right. In ΔHDX, centroid shifts are compared between unbound and bound states. The deuterium content (*D*) can be calculated with the following equation:

$$D = \{(m - m_0)/ (m_{100} - m_0)\} \times N$$

where 'm' is the peptide mass, m_0 and m_{100} is the masses of the un-deuterated and deuterated peptides (respectively) and *N* is the number of amides in the peptide. In most cases, ligand-binding reduces solvent accessibility and strengthens the hydrogen bond network in the binding pocket – this is described as a 'Type 1' scenario by Konermann[3]. Nonetheless, 'Type 0' and 'Type 2' scenarios can also be observed. 'Type 0' are regions of silent HDX exchange; when amides in the same peptides present different magnitudes and directions of HDX (i.e. accelerated and reduced exchange), the differences could 'cancel out', leading to no observed statistical significance. 'Type 2' scenarios are when an accelerated exchange profile is observed; protein undergoing large structural arrangements often exhibits this. In this thesis, we conduct ΔHDX, and the centroid

shifts in the isotopic distribution are compared between the ligand-bound and ligand-free state; Type 1 observations are expected in most cases.

1.4 YTH Domain Containing Family of Proteins

N6-methyladenine (m^6A) is one of the most abundant and conserved modifications on mRNA. m^6A marks (e.g. methyl on the 6^{th} position of adenine, **Figure 1.10**) are dynamically regulated by proteins classified into the three categories: 'readers', 'writers', and 'erasers' (**Figure 1.11**)[23]. Writers such as METTL3 and METTL14 incorporate methylations on RNA, erasers such as FTO remove these marks, and readers such as **YT521-B homology (YTH) domain-containing proteins** play roles in regulating the fate of m^6A-modified RNA. YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2 are the five YTH domain-containing proteins. YTHDC1-2 are nuclear proteins; YTHDC1 plays roles in RNA splicing[24]. YTHDF1-3 are cytoplasmic proteins that regulate translation and degradation of the m^6A RNA (**Figure 1.11**)[23].

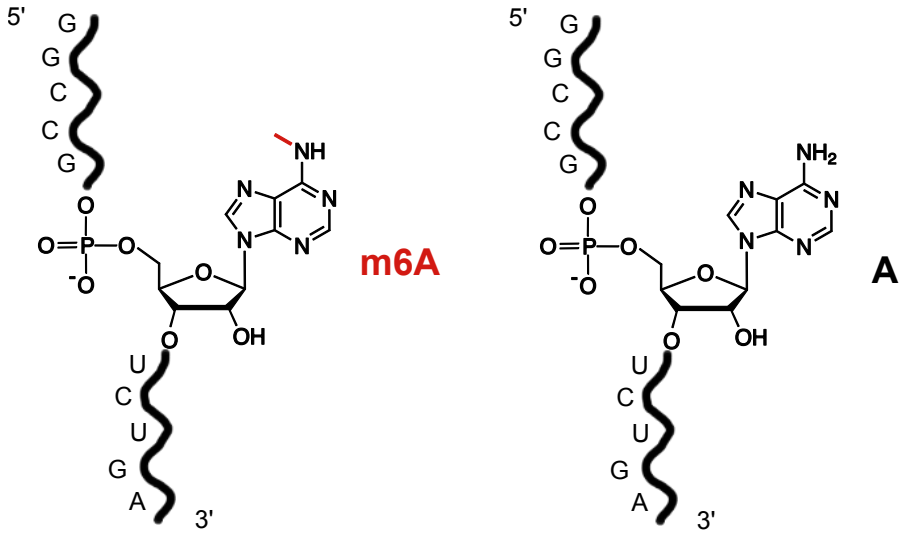


Figure 1.10 | m^6A versus A in RNA sequence
m6A (left) and A (right) are shown in RNA sequence. Figure designed using sequence identity retrieved from Arguello et al. (2019)[25].

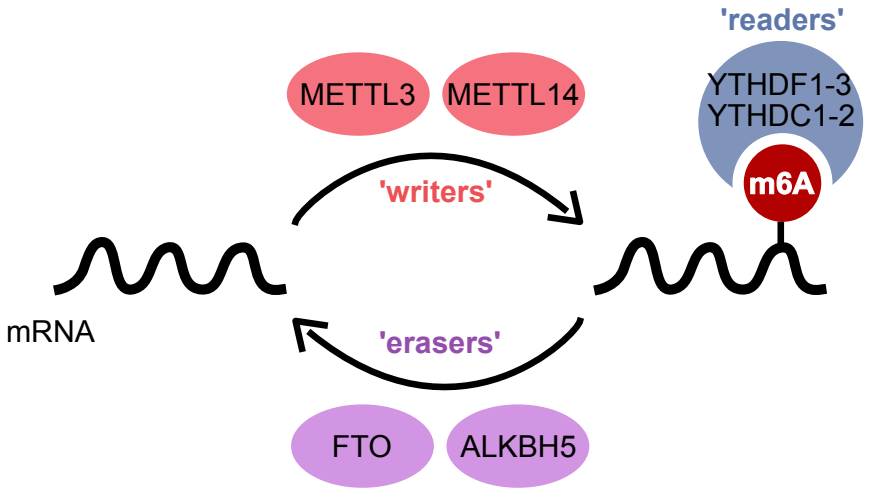


Figure 1.11 | m^{6}A readers, writers, and erasers
'Writers' such as METTL3 and METTL14 add the methyl group, 'erasers' such as FTO and ALKBH5 remove the methyl group, and 'readers' such as YTH domain containing family of proteins regulate downstream fate of the target mRNA (e.g. in splicing, translation, decay, etc.). Figure adapted with modifications from reference [23][23].

1.4.1. Biological Roles of the YTHDFs

YTHDF1 is primarily known for roles in initiating translation of m^{6}A RNA via recruiting eIF3A/B and promoting ribosomal occupancy[26]. According to Wang et al. (2015), mRNA in the cell can occupy three pools: (1) the non-ribosome pool, (2) the translatable mRNA pool and (3) the active translating pool. Translation does not occur in the first pool; the presence of ribosomes characterizes the latter two pools, facilitating translation. Wang et al. reported that YTHDF1 can impact the distribution of mRNA in the cell based on the following observations: (1) YTHDF1 enhances ribosomal occupancy of mRNA (i.e. it moves mRNA from the non-ribosome bound state to the translatable pool) and (2) YTHDF1 promotes the actively translating state (i.e. the transition from pool 2 to pool 3) – YTHDF1 can recruit translation initiation factors such as eIF3A and eIF3B to facilitate this. Overall, YTHDF1 interacts with downstream translation-associated

proteins and machineries to modulate the fate of m^{6}A-modified RNA in translation. At a molecular level, m^{6}A modifications (and hence YTHDF1 binding sites) are found near the stop codon of an mRNA. To influence translation-associated functions, YTHDF1 is brought in spatial proximity to the start site via bridging facilitated by various proteins involved in the initiation complex. YTHDF1 first binds the m^{6}A-modified site in an mRNA, then eIF3 is recruited, followed by ribosomes, and eventually translation follows **(Figure 1.12A)**[27]**.**

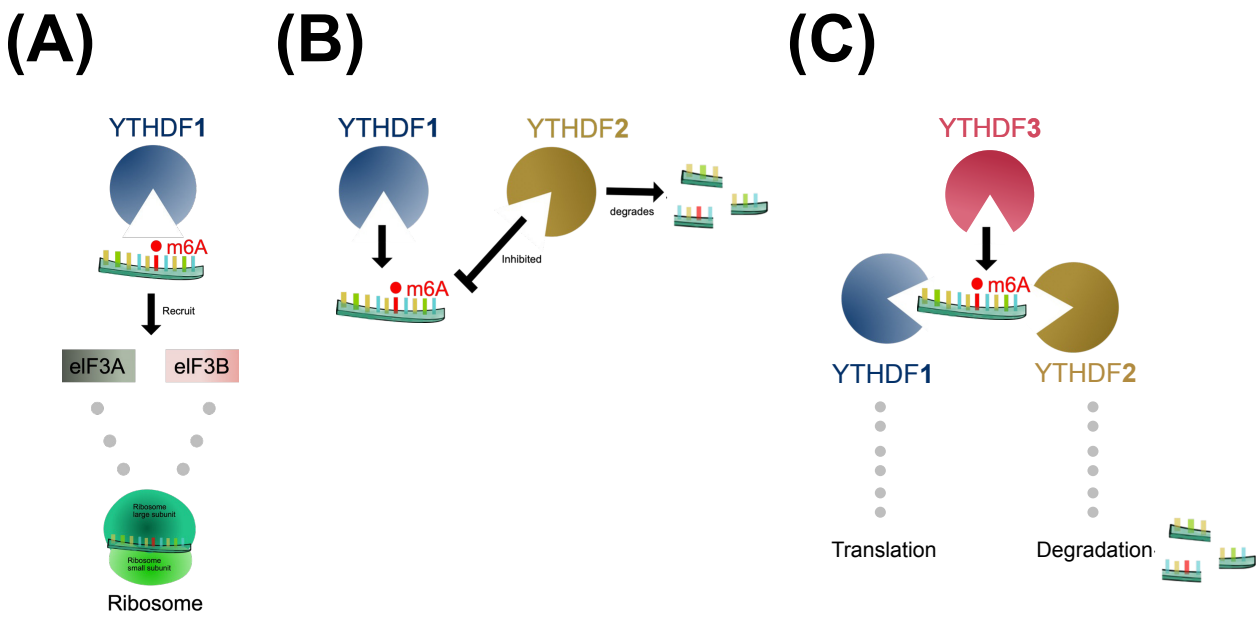


Figure 1.12 | Functions of the YTHDF1-3
Function of YTHDF1 is shown in translation and decay (A, B), function of YTHDF2 in decay (B), and coordinated function of YTHDF3 in cooperating with YTHDF1 and YTHDF2 in regulating the m^{6}A RNA in translation or decay (C). Figure adapted with modifications from reference [27][27].

YTHDF2 modulates the degradation of the m^{6}A RNA[28]. YTHDF2 can also alter the distribution of mRNA in the cell (between the three pools discussed above). Unlike YTHDF1, YTHDF2 shifted mRNA from the 'translatable pool' (i.e. the pool characterized by ribosomal occupancy) back to the non-ribosome state, suggesting that YTHDF2 binding prevents translation

of the mRNA. Furthermore, knockdown of YTHDF2 enhanced mRNA stability. YTHDF2 is also observed to recruit downstream degradation-associated proteins such as CNOTs and localizes mRNA in processing bodies (where mRNA is stored until decay). Based on various findings, the role of YTHDF2 was concluded in mRNA degradation in reference [28][28]. The binding of YTHDF1 to m^{6}A RNA can prevent degradation by YTHDF2 (**Figure 1.12B**)[27].

Finally, YTHDF3 reportedly facilitates both translation and decay of the m^{6}A RNA[29]. Like YTHDF1, YTHDF3 has also been shown to enhance the translation efficiency of mRNA in the cell and can also interact with ribosomal machinery (however, no interaction with eIF3 was detected). Similar to YTHDF2, YTHDF3 can also reduce mRNA stability. YTHDF3 likely permits both fates for the m^{6}A RNA, as it can 'shuffle' the mRNA between the YTHDFs – i.e. mRNA can be passed to YTHDF1 for translation or to YTHDF2 for degradation. YTHDF3 can also cooperatively bind to m^{6}A RNA with YTHDF1 or YTHDF2 to facilitate translation or decay (**Figure 1.12C**)[27].

1.4.2 The m^{6}A RNA Sequence

YTHDFs have specificity for their m^{6}A RNA sequence. To understand the sequence that YTHDFs prefer to bind to, Arguello et al (2019) generated a library with random RNA sequences (for example, 10^6 different possibilities for a 11-mer sequence)[25]. They conducted microscale thermophoresis experiments to identify the dissociation constant (K_D) of their chosen YTH proteins (YTH domains of YTHDF1, YTHDF2, and YTHDC1 – they did not study YTHDF3). Generally, m^{6}A readers prefer the sequence: DR<u>m6A</u>CH (where D can be A, G or U nucleotides, R can be A or G nucleotide, and H can be A, C, or U nucleotide). Arguello et al. (2019) showed

that YTH proteins can also bind to non-canonical sequences. The m^{6}A RNA sequence used in our study is derived from the Arguello et al (2019) study (referred in their study as probe 5 with the methyl group and probe 6 with adenosine). YTHDF1 had K_D of 0.51 μM for this sequence; the K_D drops to 12.4 μM when the m^{6}A mark is replaced by adenosine in the same sequence[25]. They note a K_D of 0.79 μM for YTHDF2 for m^{6}A RNA sequence, which drops to 33.1 μM in the absence of the methyl group – however, these K_D values were associated with a slightly different sequence identity (referred in their study as probe 7 with the methyl group and probe 8 with adenosine)[25]. In addition, we do not have any K_D values for the YTHDF3 protein from their study. While we do not know the exact K_D for each of the YTH domains for the specific sequence we have chosen, the affinities for the YTH proteins are generally considered to be similar for the same sequence; hence, we also assume similar K_D[30]. As noted by Arguello et al., it is essential to recognize that these are 'validated' sequences – and may not necessarily reflect *in vivo* scenarios[25]. In our study, we use the m^{6}A RNA sequence validated in the Arguello et al. study with a K_D of 0.51 μM for the YTH domain of YTHDF1, and a similar K_D is expected for the YTH domains of YTHDF2 and YTHDF3[25].

1.4.3 Sequence Alignment & Structure of the YTH Domains

The YTH domains share a conserved binding pocket. In our study, we focus on the YTH domain of YTHDF1, YTHDF2, and YTHDF3 (referred here as YTH-1, YTH-2, and YTH-3, respectively; or interchangeably as DF1, DF2, and DF3). The amino acid sequence alignment for the YTH domains is shown in **Figure 1.10**[31]. They share around ~86% similar amino acid sequences, notable differences are highlighted in **Figure 1.10**[31,32]. Ligand binding occurs on the YTH domain, which comprises of five alpha helices and six beta strands – they fold to form a beta

barrel structure[31]. The binding pocket is aromatically rich; W411, W465, and W470 (residue numbering of YTH-1) play key roles in ligand binding. These three tryptophan residues are conserved, also emphasized in **Figure 1.10**, and known to stabilize m6A binding[31]. It has also previously been noted that the methyl group occupies and fills the space between the tryptophan residues 470 and 465 (residue numbering of YTH-1)[33]. Therefore, it could also be of interest to study how the YTH proteins bind to the same sequence without the methyl group and what impact this would have on their conformational dynamics, particularly on these tryptophan residues. In addition, Y397, C412, and potentially D401 (residue numbering of YTH-1) can form H bonding with N3, N6, and N1 positions of m^{6}A, respectively[31]. These key amino acids are conserved between all YTH domains and play roles in similar interactions, as shown in **Figure 1.11** (note the residue numbering in YTH-1, YTH-2, and YTH-3 is different)[33].

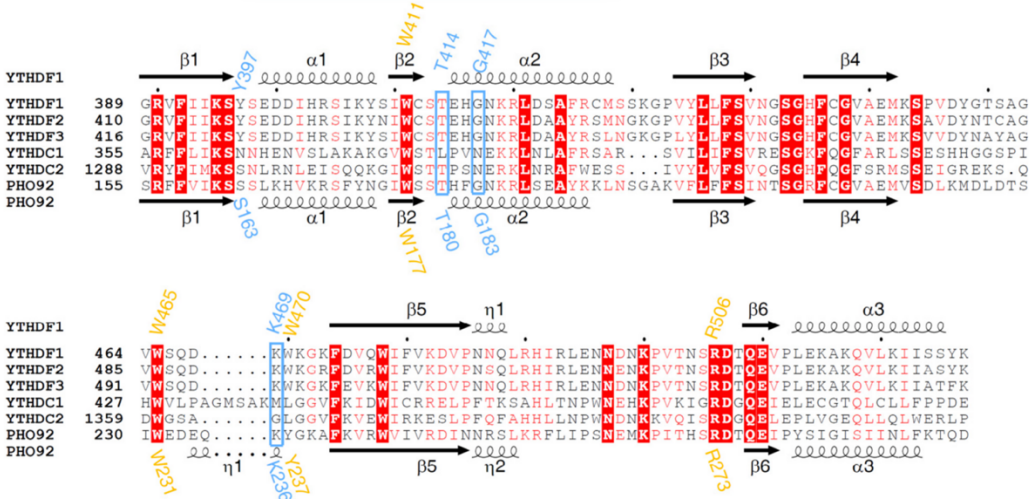


Figure 1.13 | YTH Sequence Alignment[31]
Sequence alignment of the five YTH domain-containing proteins (and PHO92, which shares structural similarity). This figure is from reference [31][31].

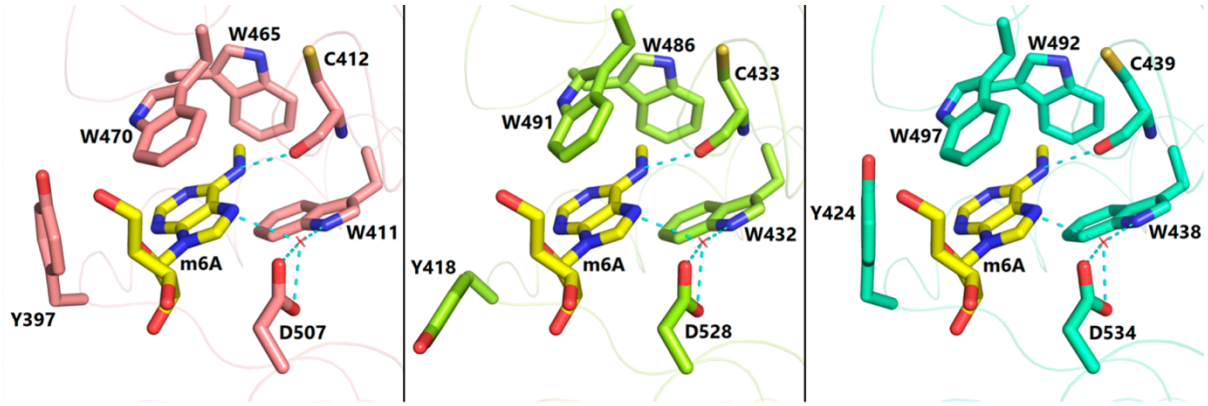


Figure 1.14 | YTH Binding Pocket Comparison[33]
Binding pocket residues of YTH domain of YTHDF1 (left), YTHDF2 (center) and YTHDF3 (right). Figure taken with permission from reference [33] © 2020 American Chemical Society [33].

We also summarize a few distinct features noted between the YTH domains. First, while the binding pocket is conserved, the m^{6}A recognition loop, located between β4 strand β5 strand (also known as the β4-β5 loop), is known to be important for binding and contains differences in the amino acid composition between the YTH domains **Figure 1.11** [31]. The differences in sequence identity, particularly in the β4-β5 loop, have been of interest in various molecular dynamics studies and are discussed further below. Second, in the halo structure of the three YTH domains (**Figure 1.11**), the tyrosine residue on bottom right of each panel appears to adapt a different positioning for the case of YTHDF2[33]. While we do not fully understand the reason for this difference, Li et al (2020) note that this is likely a byproduct of the differing length of the m6A oligonucleotide ligand used between the halo crystals[33]. Therefore, some differences in the orientation of the key amino acids and the dynamics of the m^{6}A recognition loop are of interest.

1.4.4. Plasticity of the β4-β5 between the YTH domains

Cazzanelli et al. 2024 provide insights into the plasticity of the m^{6}A recognition loop in the apo YTH domains of YTHDF1, YTHDF2, and YTHDF3[34]. The apo crystal structure of both YTH1

(PDB: 4RCI) and YTH3 (PDB: 6ZOT) has a disordered m6A recognition (β4-β5) loop and low electron density; whereas YTH2 (PDB: 4RDO) presents a structured m6A recognition loop (**Figure 1.12**). Despite the structural similarity between these domains, the crystal structures highlight notable differences, particularly in this β4-β5 loop. One of the key differences in the sequence of this loop is the presence of a glycine at position 459 of YTH-1, which is replaced by asparagine in YTH-2 and YTH-3 (**Figure 1.10**)[31]. G459 is expected to contribute to enhanced loop flexibility in YTH1 relative to YTH2 and YTH3 based on molecular dynamics simulations)[34]. However, when G459 was mutated, the m^{6}A recognition loop remained disordered; therefore, sequence differences likely do not explain the differences between the crystals; rather, they could be a byproduct of the crystallization process. Overall, crystals show differences in the electron density of the m^{6}A recognition loop between the apo YTH domains – this is likely not due to the sequence differences; the presence G459 in the loop is expected to contribute to increased flexibility in YTH1.

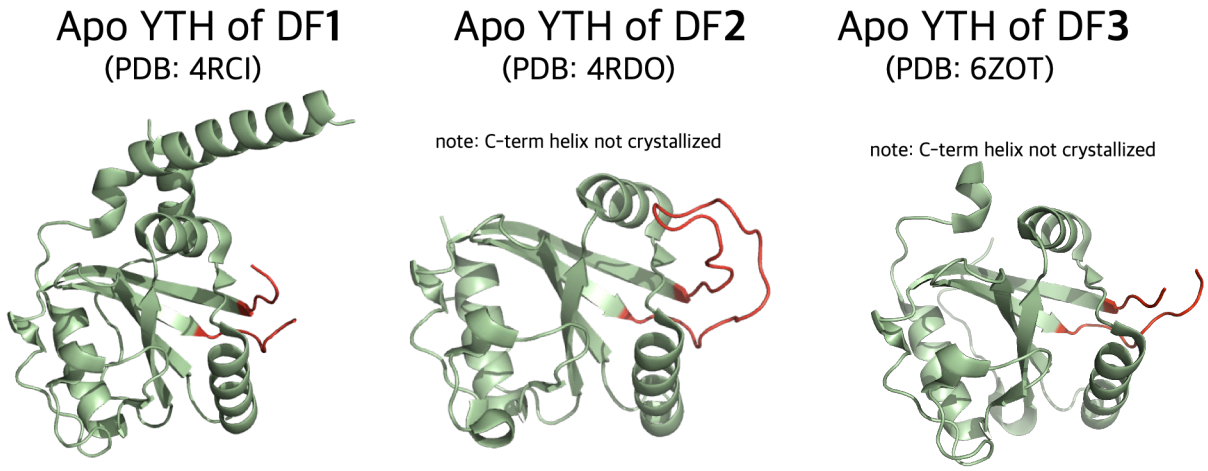


Figure 1.15 | YTH Apo Crystal Structure Comparison
Crystal structures for the apo YTH1-3 domains. The m6A recognition (β4-β5) loop is shown in red. The β4-β5 loop in YTH1 and YTH3 is missing in the crystal structure from reduced electron density. The full-length β4-β5 loop has been crystallized in the YTH2 apo structure. Figure designed using ideas discussed in reference [34][34].

Similar observations on the flexibility of the m^6A recognition loop were made by Li et al. (2020)[33]. Their molecular dynamics and root mean square fluctuations (RMSF) data indicate the highest loop flexibility for YTH1, followed by YTH3. However, other than the RMDF data, the authors do not discuss the dynamics of YTH2 in writing. Nonetheless, they do conclude that based on the similarity in dynamics of YTH1 and YTH3 proteins, it is likely that all three YTHDFs share similar functions (the function of these proteins has been debated in the literature, as discussed later in Chapter 2).

Chapter 2 | Hydrogen Deuterium Exchange Mass Spectrometry Perspective on the Function Debate of the three YTHDF Proteins

2.1 Abstract

The function of the three YTHDF proteins has been debated in the scientific community. According to the prevailing model, the three YTHDF proteins (YTHDF1, YTHDF2, and YTHDF3) have different functions, while the unified model proposes shared functions. Here, we conducted Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) to detect whether the observed conformational dynamics between the three YTHDFs are redundant or different for the fate of the m^6A-modified RNA ligand. We observed that all three YTHDFs showed significant stabilization in the bound state relative to the unbound state. However, the conformational dynamics of YTHDF2 were most different, particularly in the α2–β3 region. While previous studies and crystallography data suggest a similar mechanism of binding between the YTH domains, HDX-MS was useful in detecting unique conformational dynamics in areas influenced by allostery, such as the α2–β3 region. In addition, we observe that the conformational dynamics of YTH-1 and YTH-3 appear similar. Various of our observations align with those from the Chaun He Lab.

2.2 Introduction

The function of the three YTHDF proteins has been under scientific debate; the He lab has made some of the earliest contributions to support the differing functions of these proteins[26,28,29]. The YT521-B homology (YTH) Domain-containing Family of proteins (YTHDFs) are 'reader' proteins known to regulate the fate of RNA modified with *N6*-methyladenine (m^6A), one of the

most abundant modifications in RNA. In the Wang et al. 2014 and 2015 study (from the He Lab), the function of YTHDF1 and YTHDF2 was described for the first time in translation and decay, respectively[26,28]. Two of the relevant experiments from their study are the ribosomal profiling assay (to study translation by measuring how much m^6A RNA is occupied by ribosomes) and RNA lifetime assays (to study RNA stability/decay). Ribosomal occupancy is increased by YTHDF1, and RNA lifetime assays supported the role of YTHDF2 in decay (both observations were further supported by their tether reporter assays as well). Later in 2017, the function of YTHDF3 was described in both translation and decay in the Shi and Wang et al. 2017 study (also from the He lab)[29]. Their temporal-based assays further summarized the coordinated function of the YTHDFs as follows: YTHDF3 is the first protein to bind to m^6A RNA when it enters the cytoplasm and can also act as a 'buffering agent' to pass the m^6A RNA to YTHDF1 for translation or to YTHDF2 for decay, and YTHDF2 was shown to bind last in that order. Under the 'prevailing' model, YTHDF1, YTHDF2, and YTHDF3 are thought to play roles in translation or decay or both (respectively) – because these are 'opposite fates' for the m^6A RNA; the coordinated impact of the three YTHDFs is thought to 'balance protein production'[26].

Contrary to the above, the Zaccara et al. 2020 report described a 'unified model' (**Figure 2.1**) – the three YTHDF proteins were presented to have a 'redundant' role in degradation of the m^6A RNA[35]; contradicting with various observations from the He lab[26,28,29]. Zaccara et al. suggested (among other things) that based on Arguello et al. study and crystallography study by Xu et al.[25,31], the YTHDFs bind to m^6A RNA in similar ways and have similar preferences for the m^6A RNA motif, suggesting similar functions. This was of interest to us – because even if the conformational dynamics in the binding pocket are similar, the allosteric changes could potentially

be different between the YTHDFs, suggesting different functions. Zaccara et al. also noted that the localization patterns of all YTHDF paralogs are similar[35]; this contradicts with the He Lab where differing localization patterns, matching their differing functions, were presented[36]. Further, Zaccara et al. does not detect involvement of DF1 in translation (as was previously established in Wang et al.), and we quote, "after re-processing of the raw ribosomal profiling data from Wang et al. (2015), DF1 does not regulate translation of m^{6}A-RNA"[26,35]. According to the Zaccara et al. 2020 report, single knockdowns of YTHDFs paralogs (which had been done prior to this study) may not be effective in implying functions as the cell can compensate by overexpressing the other paralogs[35]. Zaccara et al. were the first to conduct a triple knockdown and support the view that the different YTHDFs have redundant functions in m^{6}A RNA degradation[35].

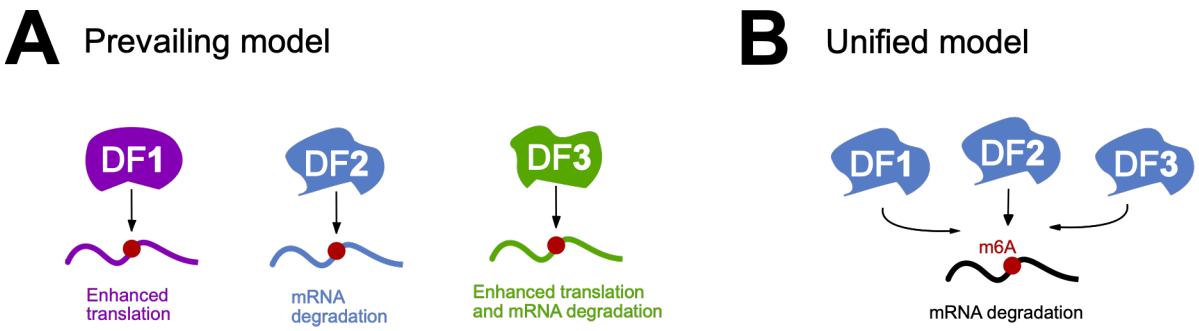


Figure 2.1 | Prevailing versus unified model of Function Debate[35]
The debate on the function of the YTHDF proteins. (A) Prevailing model versus (B) unified model. Figure adapted with minor modifications from Zaccara et al (2020)[35].

While Zho et al. 2023 (also from the He lab) has responded to the report from the Zaccara et al.[35] and continue to re-enforce the differing model view in reference [36][36]; other studies that support the unified model view via orthogonal techniques also exist. For example, Li et al. 2020 conducted molecular dynamics (MD) simulations and support the model with redundant function

as per their observations and we quote, “MD simulations show similar flexibility and correlated motion for the three YTH domains” (although some of their root mean square calculations did show some differences between the thermal motions) [33]. Overall, in our understanding, authors are debated on the ‘prevailing’ versus ‘unified’ view of the functions of the YTHDFs.

Here, we conduct hydrogen deuterium exchange mass spectrometry (HDX-MS) to study whether the observed conformational dynamics between the YTHDFs are redundant or different – we aim to provide an HDX perspective on the function debate. We compare the YTH domains of YTHDF1, YTHDF2, and YTHDF3 (referred here as **YTH-1**, **YTH-2**, and **YTH-3**, respectively – although this is not the commonly used convention, we have used it here for convenience). The three YTH domains have some notable differences in their amino acid sequences; a sequence alignment of the YTH domains can be found in the Xu et al. (2015) paper[31]. Differences in the side-chain amino acids can impact the rate of exchange and usually need to be accounted for as per the Englander lab’s work (there recent adjustments to these calculations in reference [12])[12]. However, we do not need to account for this in this study because we are *not* comparing the YTH 1-3 domains *between* each other. Rather, we study how the conformational dynamics for YTH-1 differs between the unbound state versus bound state. Then, we asked the same question for YTH-2 and YTH-3. Finally, we compare the three YTH domains to study whether the binding of m^6A RNA ligand induced any unique or redundant conformational dynamics between protein. In our study, we use the m^6A RNA sequence validated in the Arguello et al. study with K_D of 0.51 μM for YTH-1, and similar K_D is assumed for YTH-2 and YTH-3 (sequence shown in **Figure 2.2**)[25]. As noted by Arguello et al., it is important to realize that these are ‘validated’ sequences – and may not necessarily reflect *in vivo* scenarios[25]. Overall, the binding of the selected m^6A RNA is studied

with each of the YTH domains, and final observations are aligned between proteins to compare conformational dynamics.

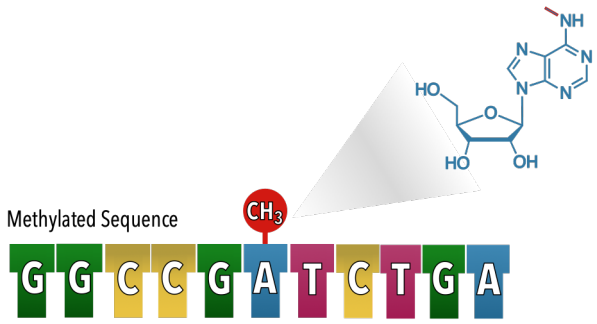


Figure 2.2 | The m^{6}A RNA sequence.
We have visualized our m^{6}A RNA ligand here; the sequence identity of the m^{6}A RNA sequence used in our study is retrieved from Arguello et al. study[25].

To conduct HDX-MS, we consider the recommendations in HDX Nature Method Paper as much as possible; for example, composition of various buffers and LC gradients is reported (refer to Methods section), multiple technical and biological replicates are acquired, data is made transparent (isotopic distributions and kinetics plots from key regions are included in this chapter), wide range of timepoints are selected, HDX runs between the unbound and bound state were randomized, coverage maps have been included, and software used for peptide identification and analysis are also reported (refer to Methods)[22]. To analyze our data, we consider data processing guidelines by Sarpe and Schriemer in reference [11][11] – we realize the limitations of data processing, and understand that in cases where the monoisotopic peak disappears, analyzing partial distributions across all timepoints has also been shown to be effective[11]. Fundamentals of HDX and guidelines on the interpretation of HDX data are described by Konermann in references [2] and [3][2,3]. As noted by Konermann et al., and we quote, "Unfortunately, the interpretation of HDX/MS data is not always straightforward, and novices may have some misinterpretations"[3]. We

have realized these misinterpretations; for example, proteins undergo open and close conformations, and it is important to realize that when regions present a reduction in deuterium uptake, it is not necessarily be because a particularly region has become more 'rigid'; rather, the protein occupies the 'open' state less upon ligand-binding[3]. We also understand that back exchange is a fundamental limitation of HDX-MS, we do not need to correct for back exchange when doing comparative HDX[2]. However, it is still important to note that the first two amino acids in any given peptide usually do not contribute to differences in the peptide (i.e. some spatial resolution is lost)[2]. Overall, we consider available recommendations, and guidelines for analysis as well as interpretation of HDX-MS work as much as possible.

Crystallography and molecular dynamic studies predict a conserved mechanism for ligand binding, and offer other insights that help us formulate our hypotheses[31,33,34]. As HDX is a complementary technique, we refer to crystal structure and molecular dynamics predictions to guide the interpretation of our HDX data. First, ligand binding occurs on the YTH domain, which comprises of five alpha helices and six beta strands – they fold to form a beta barrel structure[31]. The binding pocket is highly similar between the YTH domains and contains a conserved 'tryptophan cage' of residues W411, W465, and W470 (residue numbering of YTH-1) which play key roles in ligand binding. When these tryptophan residues were mutated to alanine, m^6A RNA binding was disturbed[31]. The m^6A recognition loop (also known as the β4-β5 loop) is also a recognized important region of the protein and contains two of these key tryptophan residues, W465 and W470 (residue numbering of YTH-1), and can participate in induced-fit to accommodate m^6A RNA binding[33,34]. In our HDX data, we expect to see reduced deuterium uptake in the hydrophobic aromatic binding pocket and the m^6A recognition loop region. Second, there

are residues which are not directly involved in ligand binding, but these residues are still important in stabilizing the interaction between YTH and m^{6}A RNA ligand. For example, according to the YTH-2 crystal structure by Zhu et al; R411, K416, R441, and R527 (residue numbering of YTH-2) may be important in YTH interaction with RNA backbone[37]. Xu et al. also notes the importance of additional residues for stabilizing the YTH – m^{6}A RNA interaction[31]. Therefore, we expected reduced deuterium uptake in *not* only the binding pocket, but also in peptides containing these other residues as they are thought to be important for stabilizing the RNA backbone. It is not unusual to observe that some proteins may present global (or less localized) changes in HDX-MS[1]. Third, based on molecular dynamics, the β4–β5 loop of YTH-1 is thought to be more flexible than YTH-2 or YTH-3, because YTH-1 loop contains a glycine which is replaced by asparagine in YTH-2 and YTH-3 – this observation is noted by both Li et al. (2020) and Cazzanelli et al. (2024)[33,34]. So we also expect that the m^{6}A recognition loop of YTH-1 will be more flexible in comparison[33,34]. Most importantly, we understand that the binding pocket mechanism between the YTH domains is thought to be conserved; however, if we could detect allosteric differences in conformational dynamics, that could suggest different functions for the YTH domains; and HDX is especially useful for detecting allostery.

2.3 Results and Discussion

2.3.1 Expression and Purification of the YTH domains of YTHDF1-3

YTH 1-3 were characterized on sodium dodecyl sulfate polyacrylamide (SDS-PAGE) and with native mass spectrometry (native MS); the protein sequence of the three N-terminally histidine tagged YTH domains can be found in **Table 2.1**. YTH-1 presents a single band at ~ 25

kDa on SDS-PAGE with estimated > 85% purity and matched the expected monomeric mass in unbound native MS of the protein (**Figure 2.3**). YTH-2 also matched with the expected monomeric mass of ~ 21 kDa in native MS and on SDS-PAGE (**Figure 2.3**). Although a minor dimeric species was seen; the band corresponding to the monomeric YTH-2 species appeared more dominant on gels (**Figure 2.3**). YTH-3 also matched the expected mass of ~ 24 kDa on gels and native MS; however, a dimeric species was seen on gels at equal intensity as the monomeric band. Attempts to reduce the dimeric species with 1-10 mM dithiothreitol (DTT) over 1.5 hours to as much as 24 hours at 4 ℃ were made (as shown in **Figure 2.4**); however, attempts were unsuccessful. In addition, we did not want to add too much reducing agent as pepsin (enzyme we needed for bottom-up proteomics) contains disulfide bonds itself (the trade-off of working with pepsin and disulfide bonded proteins in bottom-up proteomics is realized in the HDX community)[1]. We assumed that the dimeric species would not significantly interfere with our H/D exchange comparison between the YTH proteins. Because the dimer is present in both YTH-2 and YTH-3 (not just one protein exclusively), we assume YTH-2 and YTH-3 may still be comparable; and because YTH-3 and YTH-1 (which does *not* contain a dimer) still present comparable results (as discussed later), this gives us confidence that the dimer may not be impacting our results, at least not significantly. Besides this, the size of the YTH domains is different, primarily because YTH-1 domain contained the C-terminal helice which was either fully or partially truncated in YTH-2 and YTH-3, respectively. This difference could also be argued to generate data that may not be fully comparable and in some cases truncations of protein domains can potentially impact the conformational dynamics. We have made our attempts at making sure that all YTH domains are as comparable as possible; however, some limitations persist. And our HDX data is limited to the proteins we did have as transparent in **Figure 2.3**.

Table 2.1 | N-terminally Histidine Tagged YTH 1-3 Protein Sequences.

YTH-1 (top), YTH-2 (middle), and YTH-3 (bottom) sequences are shown. His-tag in the red; bolded sequence corresponds to YTH domain sequence, and the underlined sequence is the parts of the sequence that in included the PDB structure we used for coloring HDX data (4RCJ for YTH-1, 4RDO for YTH-2 and 6ZOT for YTH-3)

RYTMHHHHHHSSGRENLYFQG**SVESHPVLEKLKAAHSYNPKEFEWNLKSGRVFIIKSYSEDDIHRSIKYSIWCSTEHGNKRLDS AFRCMSSKGPVYLLFSVNGSGHFCGVAEMKSPVDYGTSAGVWSQDKWKGKFDVQWIFVKDVPNNQLRHIRLENNDNKPVT NSRDTQEVPLEKAKQVLKIISSYKHTTSIFDDFAHYEKRQEEEEVVRKERQSRNKQ**
MHHHHHHGGSG**SEPHPVLEKLRSINNYNPKDFDWNLKHGRVFIIKSYSEDDIHRSIKYNIWCSTEHGNKRLDAAYRSMNGKG PVYLLFSVNGSGHFCGVAEMKSAVDYNTCAGVWSQDKWKGRFDVRWIFVKDVPNSQLRHIRLENNENKPVTNSRDTQEVPL EKAKQVLKIIASYKHTTSI**
MHHHHHHGGSG**SVEVHPVLEKLKAINNYNPKDFDWNLKNGRVFIIKSYSEDDIHRSIKYSIWCSTEHGNKRLDAAYRSLNGK GPLYLLFSVNGSGHFCGVAEMKSVVDYNAYAGVWSQDKWKGKFEVKWIFVKDVPNNQLRHIRLENNDNKPVTNSRDTQEV PLEKAKQVLKIIATFKHTTSIFDDFAHYEKRQEEEEAMRRERNR**

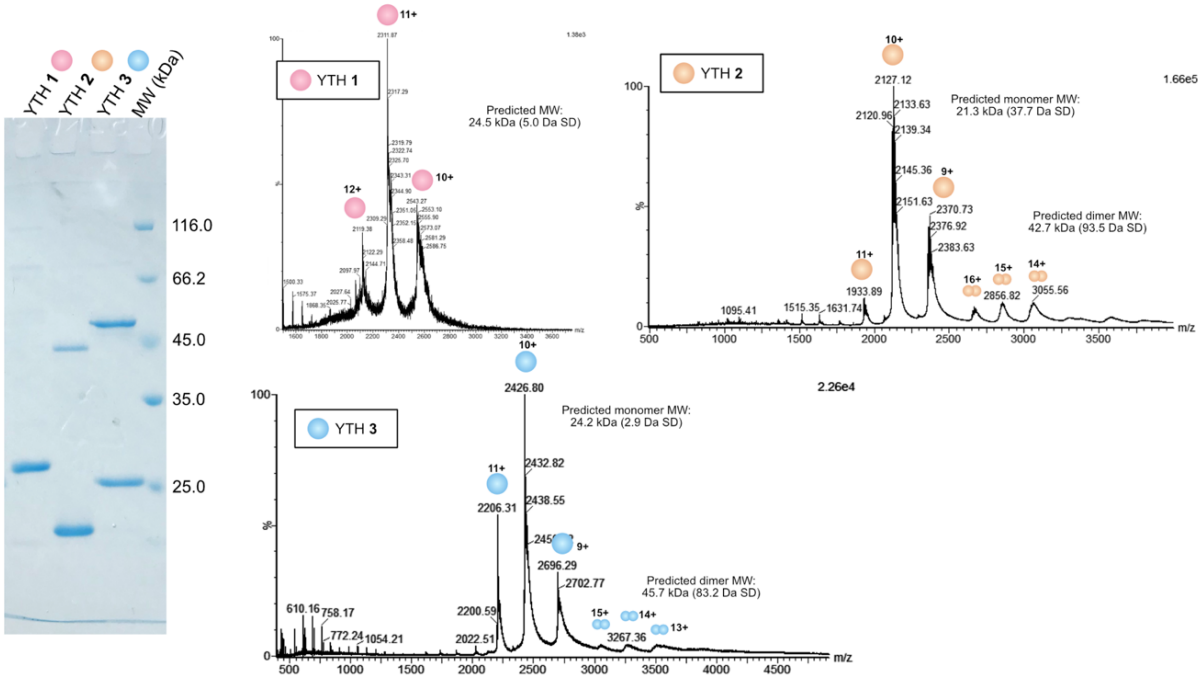


Figure 2.3 | YTH1-3 SDS-PAGE and Native MS
YTH-1 (pink circle), YTH-2 (orange circle), and YTH-3 (blue circle) are shown on SDS-PAGE and Native MS. A dimeric band was seen in gel for YTH-2 and YTH-3. For YTH-1 Native MS, trapping was used to improved signal to noise (but not for YTH-2 and YTH-3). All proteins matched with the expected monomeric mass on Native MS; for YTH-2 and YTH-3, a dimeric species was also seen on Native MS as well.

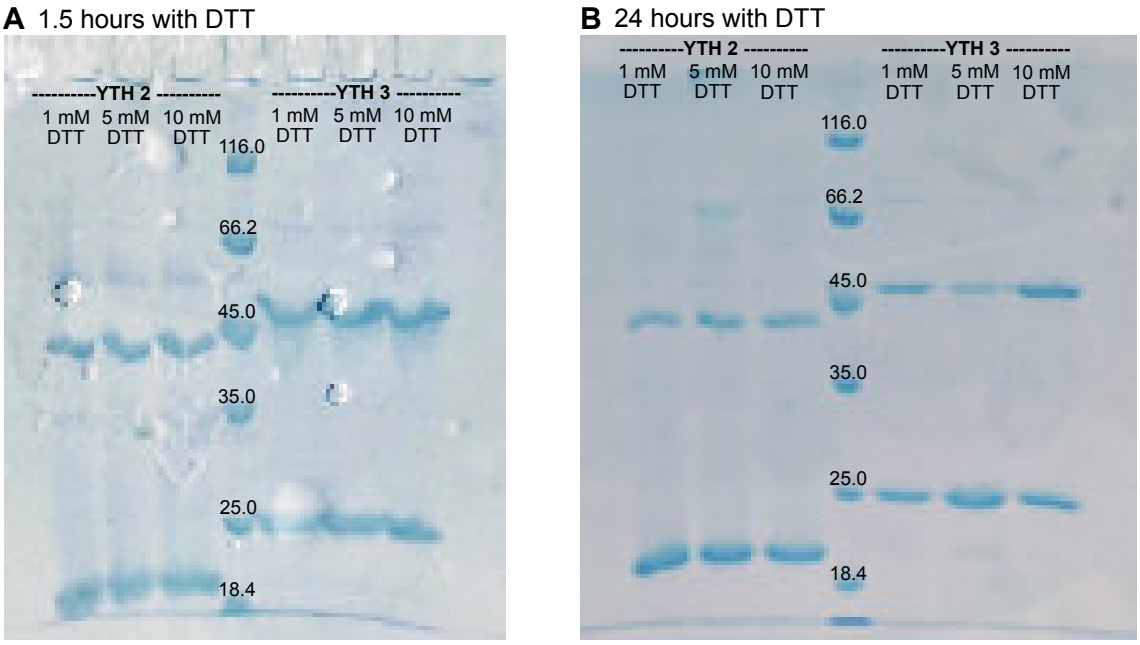


Figure 2.4 | Attempts to Reduce the Covalent Dimer with DTT
1-10 mM DTT was added in both YTH-2 (3 bands on left side) and YTH-3 (3 bands on right side) sample for 1.5 hours (A) and for up to 24 hours (B) at 4 ℃.

2.3.2 Reproducible Peptides and Coverage Across the YTH Domains

A *comprehensive* list of reproducible peptides is pooled and aligned across the YTH proteins in **Figure 2.5**, and *individual* difference plots for each YTH protein is included in **Figures 2.6-2.8**. **Figure 2.5** shows an alignment of peptides across the three YTH proteins; note that the residue numbering of the three proteins is different (for example, amino acid position 381 in YTH-1 corresponds to 402 in YTH-2, and 408 in YTH-3). A sequence alignment of the YTH domains can be found in the Xu et al (2015) paper[31]. In our study, each peptide was analyzed over 4 timepoints, with at least 2-3 technical and ≥ 2 biological replicates. Peptides matching our criteria are pooled in the *comprehensive* difference plot in **Figure 2.5**; we have also included *individual* difference plots with unfiltered peptides (i.e. includes peptides with less than 4 timepoints and/or lower replicates etc.) in **Figure 2.6-2.8** for YTH-1, YTH-2, and YTH-3, respectively. These *individual* plots include more peptides from key regions such as the m^6A recognition loop; peptides

from this region are mostly excluded from the *comprehensive* plot due to inability in getting reproducible coverage in this area. Our *individual* difference plots highlight the reproducibility of our data for the same YTH protein across different biological replicates, and the *comprehensive* plots highlight a peptide-by-peptide comparison across different YTH domains when bound to m^6A RNA.

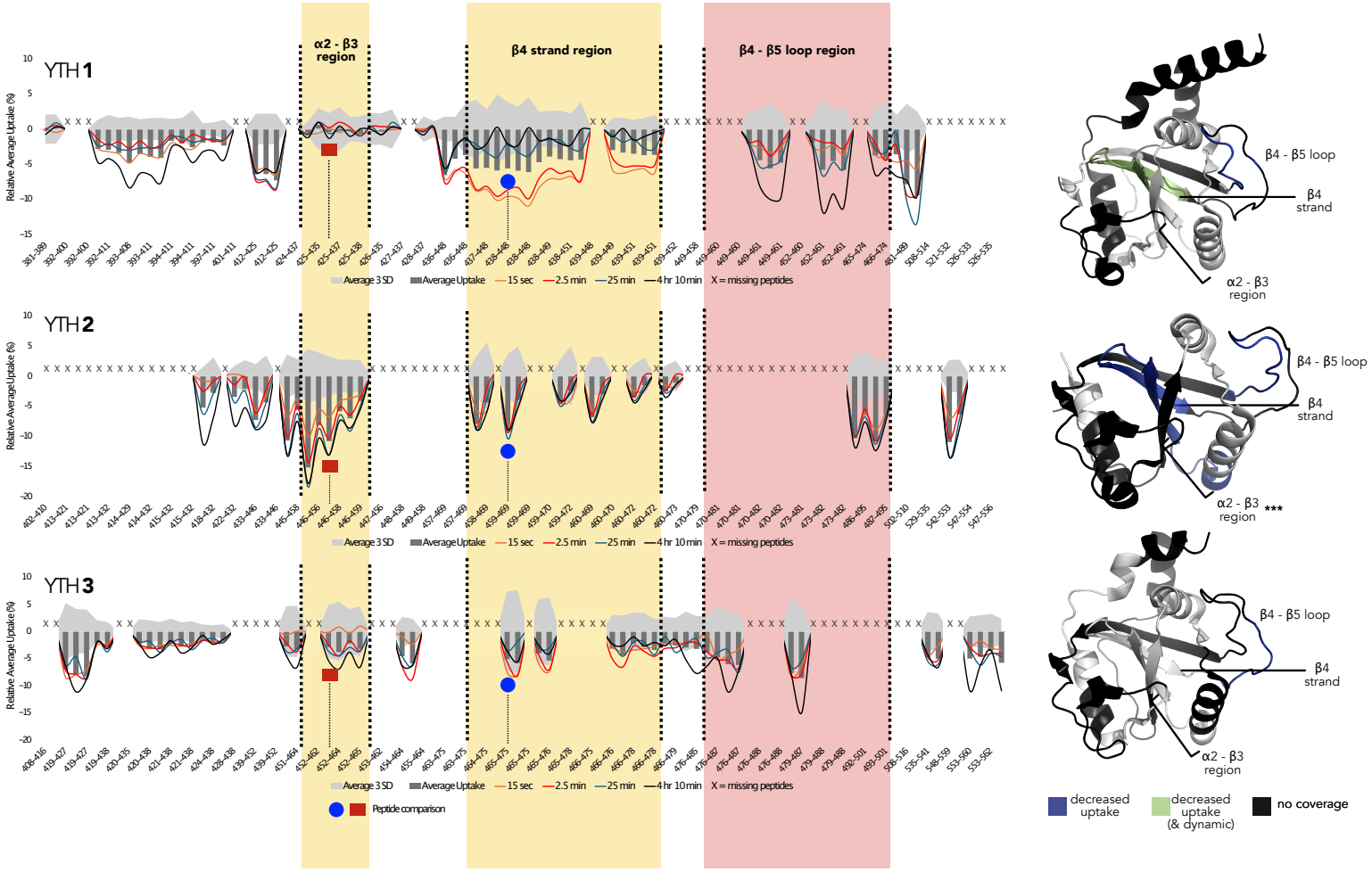


Figure 2.5 | YTH 1-3 Complied Difference Plots

The grey bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the m^{6}A bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints: 15 seconds (orange), 2.5 min (red), 25 min (blue), and 250 min (black). The gray shaded area indicates the error at 3 σ. The plots are aligned between YTH proteins, and a 'X' label is given to indicate missing peptides. Data is mapped on PDB structures: (4RCJ for YTH-1, 4RDO for YTH-2, and 6ZOT for YTH-3)[31,33,38]. Only the following regions are highlighted and mapped on structures: the α2 – β3 region, β4 strand, the m^{6}A recognition loop (β4-β5 loop). Blue color on crystals

represents reduced uptake, green color indicates potentially dynamic areas, and black for no coverage. The red square and blue circle on difference plot are used to flag examples of peptides that are comparable between the YTH domains (note that amino acid numbering is different between different YTH proteins).

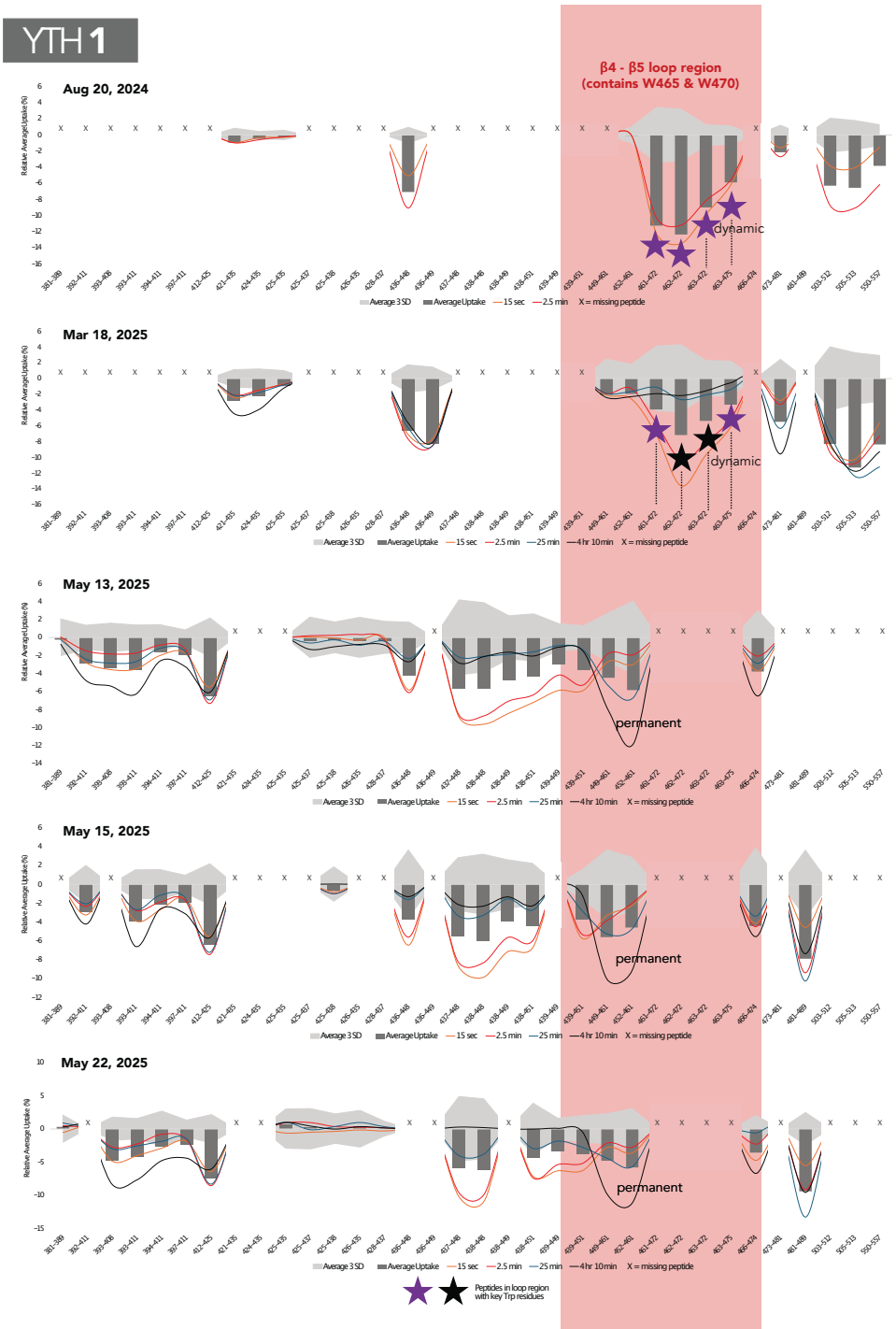


Figure 2.6 | YTH-1 Individual Biological Runs
The grey bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the m^{6}A bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints: 15 seconds (orange), 2.5 min (red), 25 min (blue), and 250 min (black). The gray shaded area indicates the error at 3 σ. The plots are aligned between YTH proteins, and a 'X' label is given to indicate missing peptides. Purple/black stars are used to flag peptides used in discussion.

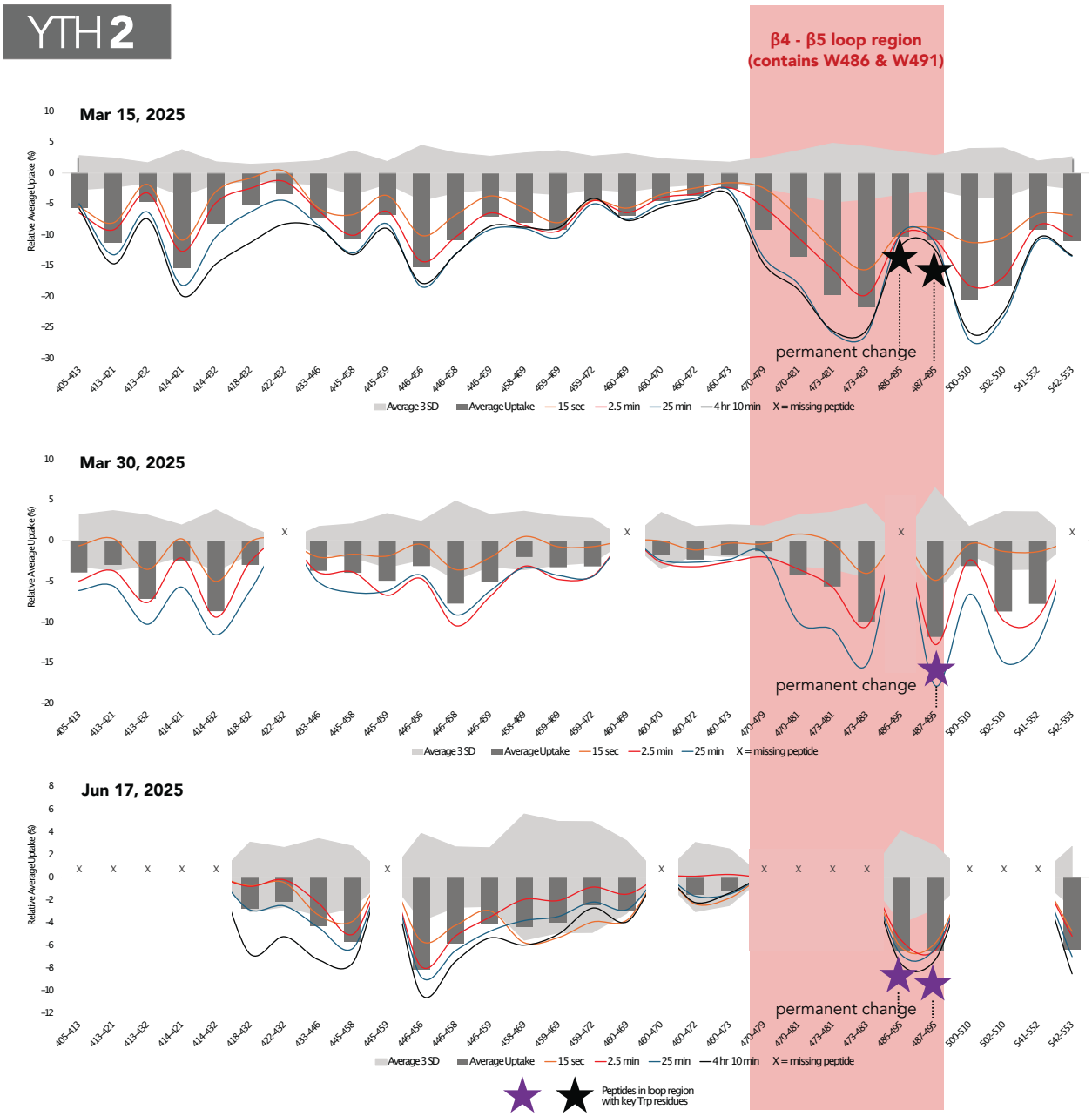


Figure 2.7 | YTH-2 Individual Biological Runs

The grey bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the m^{6}A bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints: 15 seconds (orange), 2.5 min (red), 25 min (blue), and 250 min (black). The gray shaded area indicates the error at 3 σ. The plots are aligned between YTH proteins, and a 'X' label is given to indicate missing peptides. Purple/black stars are used to flag peptides used in discussion.

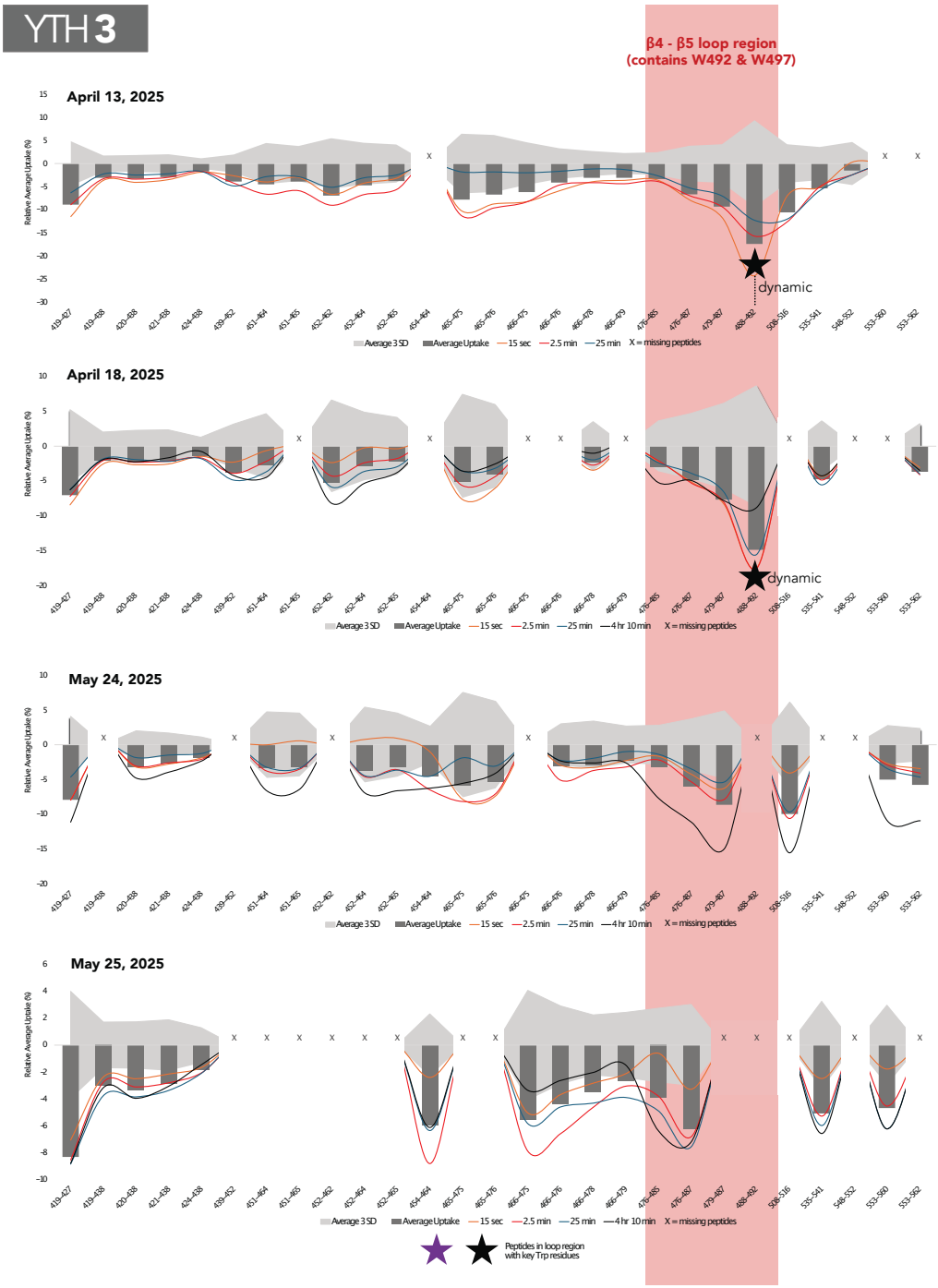


Figure 2.8 | YTH-3 Individual Biological Runs
The grey bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the m^{6}A bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints: 15 seconds (orange), 2.5 min (red), 25 min (blue), and 250 min (black). The gray shaded area indicates the error at 3 σ. The plots are aligned between YTH proteins, and a 'X' label is given to indicate missing peptides. Purple/black stars are used to flag peptides used in discussion.

For bottom-up proteomics, we aimed to get coverage in comparable peptides across all YTH domains to provide more accurate predictions on the conformational dynamics between similar regions, and experienced some difficulties as noted. First, we experienced various challenges in getting reproducible peptides. Commercial pepsin (expensive) was not used because RNA can non-specifically bind to pepsin and multiple injections involving RNA reduced its digestion capability. Therefore, for HDX-MS for protein-RNA complexes, pepsin was linked to NHS agarose and packed in columns in-house (so it can be cost-effectively re-packed as many times as needed). Because the pepsin column had to be re-packed between each run, it could have introduced variability and caused difficulty getting reproducible peptides between runs. Pepsin does not tend to give same peptides – this seems like a common observation in the HDX community[39]. Second, increased pressure of the digest column used for bottom proteomics are thought to provide enhanced digestion[1]. Commercial pepsin columns tend to reach pressures as high as approximately 8000 psi; however, our in-house pepsin column could not hold such high pressures (**Figure 2.9**). The column we used was originally a column designed for size exclusion chromatography (SEC). We recycled this column and converted it into an in-house pepsin column for RNA work. Unfortunately, this column only reaches pressures around ~2500 psi. A pressure comparison of the commercial pepsin column versus our in-house column is shown in **Figure 2.9**. In the future, using a column designed to hold higher pressure may optimize the digest. Third, we aimed to include at least three technical replicates for every time point; however, in cases of outlier results, such as that shown in **Figure 2.10,** the outlier technical replicate was removed (the removed peptide is shown in red in **Figure 2.10**). To make up for this, we tend to have more than two biological replicates in many cases, as shown in **Figures 2.6-2.8**. In future, we may lean towards collecting four technical replicates, so that we may have at least three even if one must be

removed. We also considered the recommendation of the attached paper as much as possible; the minimum recommendation in the HDX community is to have three technical replicates for at least one of the timepoints; we exceeded this expectation (with at least 2-3 technical replicates for each of the four timepoints and ≥ 2 biological replicates)[22]. And we were able to get redundant and comparable coverage across all YTH domains for at least one of the important regions for us, α2 – β3 region.

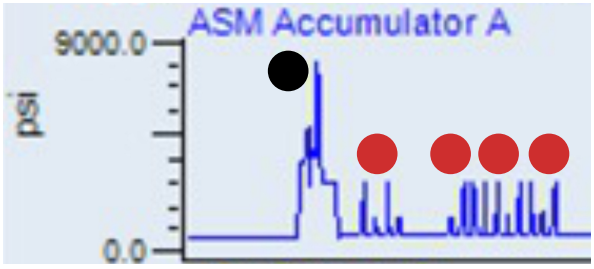


Figure 2.9 | Commercial Versus In-House Pepsin Trace Comparison
Commercial pepsin column pressures (indicated) by black circle reach as high as ~8000 psi. In-house pepsin column (red circle) pressures only reach to approximately ~2500 psi.

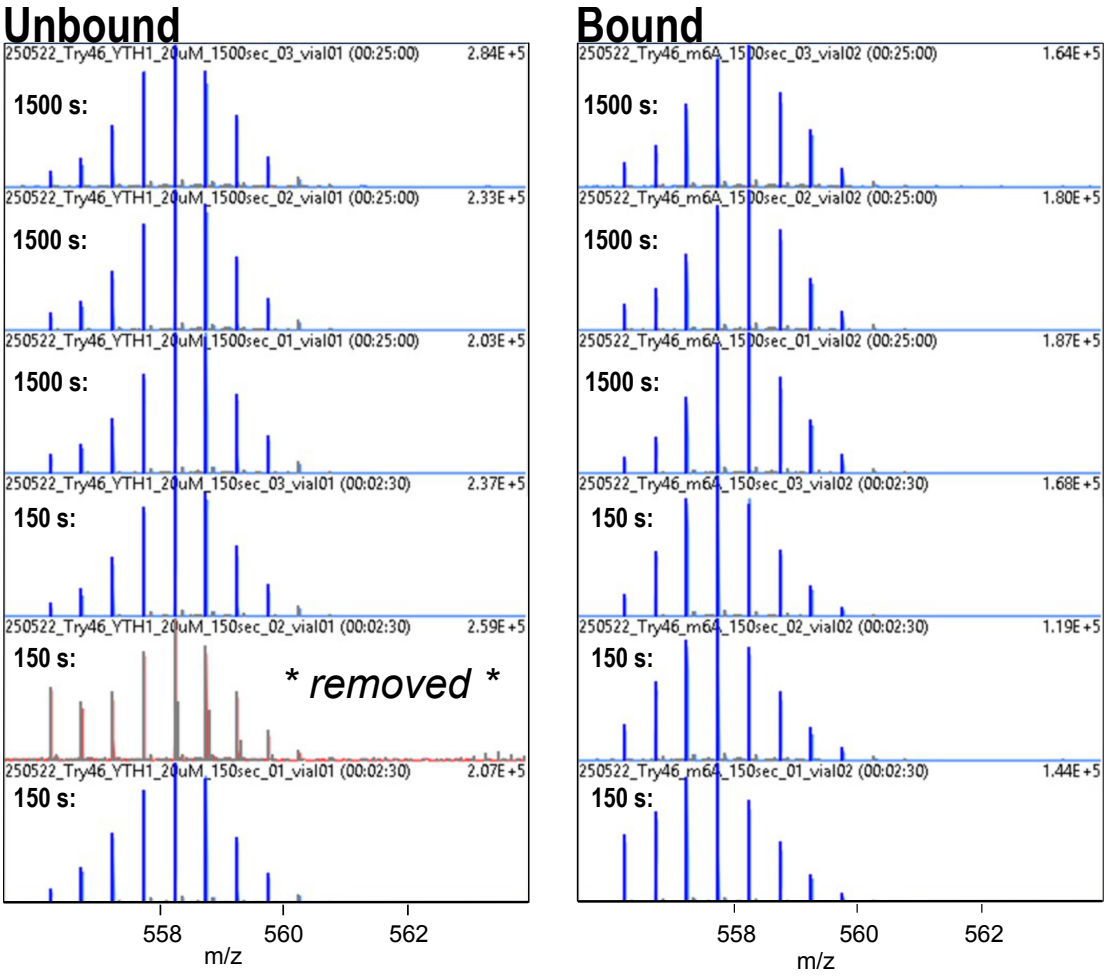


Figure 2.10 | Example of Peptide Omitted from Analysis
In cases of outlier or inconsistent observation, the peptide in question was omitted from analysis. The red color indicates unselected/removed peptides. Blue peaks correspond to peptides included in our analysis in DynamX[40].

2.3.3 YTH-2 showed significant stabilization of the α2 – β3 region while YTH-1 showed no detectable differences in this area

Peptides covering the α2–β3 region showed stabilization from m^6A binding in YTH-2, but no statistically significant changes were observed for YTH-1. Areas which showed different conformational dynamics between the YTH domains, such as the α2–β3 region, are highlighted in yellow in **Figure 2.5**. As an example, in **Figure 2.11**, we have also included the raw isotopic distributions and the kinetic plots for one of the comparable peptides from this region; this peptide

is marked by the red square in **Figure 2.5**. Peptide 425-437 in YTH-1, which corresponds to peptide residue numbering of 446-458 in YTH-2, showed no differences in the isotopic distribution in the unbound versus m^6A-bound state in YTH-1; however, in YTH-2, the m^6A-bound state visited the 'open'[3] conformation much less frequently compared to the unbound YTH-2 across all selected timepoints (**Figure 2.11**). The kinetic plots for this peptide also highlight overlapping lines between the unbound and bound states for YTH-1 (i.e. no statistical significance across any of the timepoints); however, in YTH-2, the lines representing the unbound versus bound state are statistically different at 3-fold standard deviation propagated error (3 σ), with significantly less deuterium uptake in the bound state relative to the unbound state (**Figure 2.11**). Proteins undergo open/close conformations; our data shows that in the m^6A-bound state, the 'occupancy of the open state'[3] is reduced in the α2–β3 region for YTH-2, but no difference is seen in YTH-1.

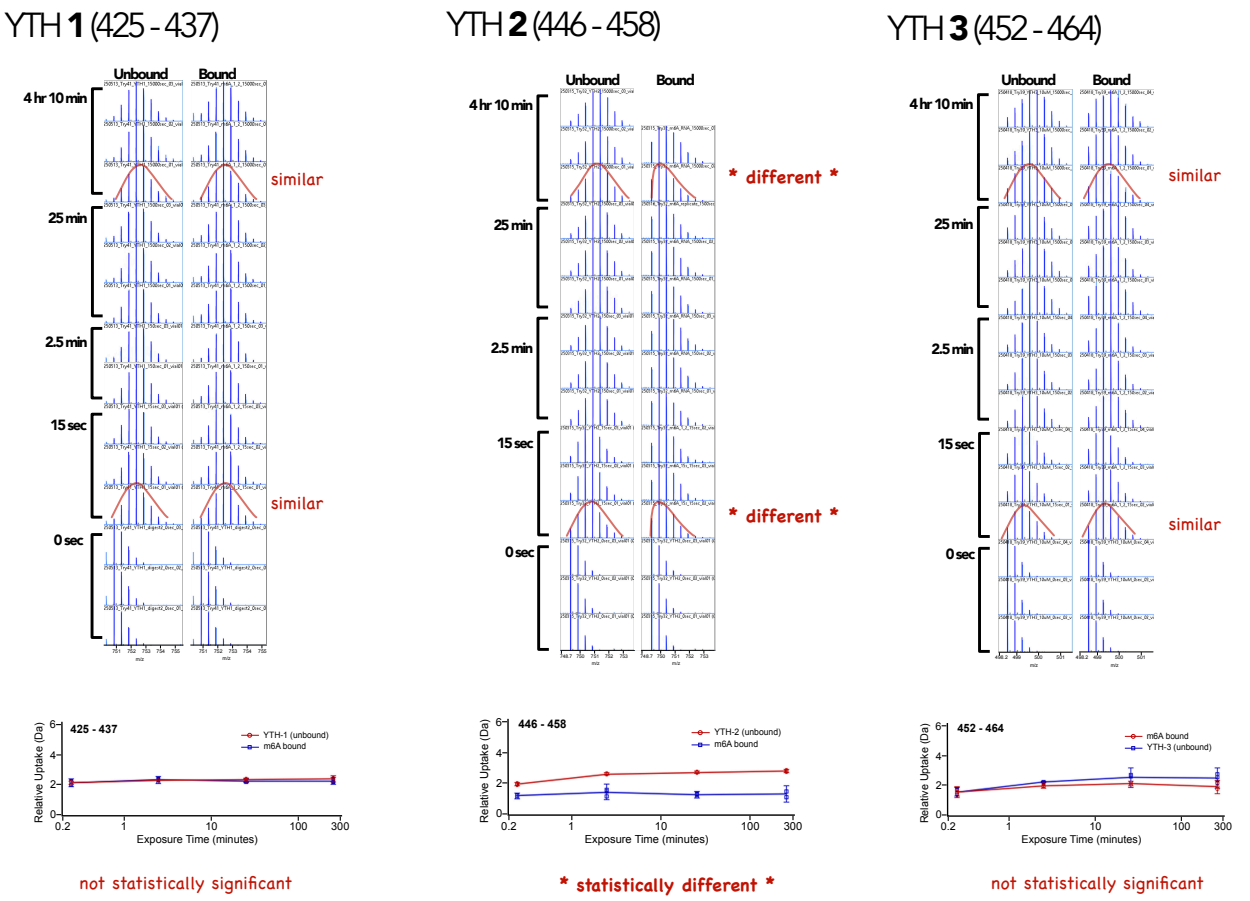


Figure 2.11 | Example of Comparable Peptide from the α2 – β3 Region
Raw isotopic distribution and kinetic plots for peptide 425 – 437 in YTH-1 (which corresponds to residue numbering 446 – 458 in YTH-2 and 452 – 464 in YTH-3) is shown. These are screenshots for Dynamx software with modifications[40].

The detection of different conformational dynamics between YTH-1 and YTH-2 could imply differing functions for the proteins, similar to supports from the He Lab[26,28]. As discussed earlier, the 2014 and 2015 paper from the He lab initially presented differing roles for YTHDF1 and YTHDF2[26,28]. More recently, in their 2023 study[36], which was written in response to the Zacarra et al. 2020 report on the unified model[35], they continue to provide data supporting the differing functions of the YTHDFs[35,36]. They note, among other things, and as also commonly understood in general biochemistry, primary amino acid *sequences* dictate protein *structure*, and

structure impacts function[36]. When Zou et al (2023) compared the *sequence* conservation between the YTHDF proteins; YTHDF2 was most different[36]. The higher order *structure* between the proteins was also different. In addition, the localization patterns for the proteins also matched with their differing functions. For example, YTHDF1 and YTHDF2 co-localized with eIFs (translation initiation factors) versus CNOTs (decay associated), respectively. This aligns with our observation of detecting differing conformational dynamics in the α2–β3 region (i.e. the data for YTH-2 is most different compared to YTH-1 or YTH-3). Since this site does not correspond to the binding pocket (as per prior knowledge from crystal structures)[31,37], allosteric stabilization of this region may be important for the function of YTH-2, but not YTH-1. The allosteric stabilization induced in the α2–β3 region of YTH-2 could imply a different fate for the target m^6A RNA ligand compared to when bound by YTH-1.

This raises an important question, why do some authors (i.e. Zacarra et al.) observe improved RNA stabilization from triple knockdown of the three YTHDFs (and why a redundant role of YTHDFs in mRNA decay was presented)[35]? This has also been addressed by Zou et al. (2023) - they show that a triple knockdown of YTHDFs can indirectly promote higher processing bodies (P-bodies) which then stabilize the mRNA in their paper titled "The mechanism underlying the redundant functions of the YTHDF proteins" - their paper title nicely captures the essence of their story[36]. Therefore, improved stability of the mRNA (based on He Lab studies) does not seem to be a *direct* effect of YTHDF triple knockdown. However, we still do not understand the contradiction between studies from the Chuan He lab and Zacarra et al. in terms of differences in the observed protein-protein interactions and the localization patterns of the three YTHDFs in the cell[26,28,29,35,36]. We realize some contradictions may exist; our HDX-MS perspective on the

function debate aligns more with the Chuan He lab studies, which we believe is the more widely accepted model.

2.3.4 The β4 strand appears dynamic in YTH-1 but shows permanent structural changes in YTH-2

Peptides in the β4 strand appear to present permanent structural change from m^6A binding in YTH-2; however, in YTH-1, the β4 stand appears dynamic in the bound state relative to the unbound state. The β4 strand (as well as other areas of differing conformational dynamics) is highlighted in yellow in **Figure 2.5**. As an example, in **Figure 2.12**, we have also included the raw isotopic distributions and the kinetic plots for one of the comparable peptides from the β4 strand region; this peptide is marked by the blue circle in **Figure 2.5**. Peptide 438-448 in YTH-1, which corresponds to peptide residue numbering of 459-469 in YTH-2 is highlighted in **Figure 2.12**. In both YTH-1 and YTH-2, the isotopic distribution of this comparable peptide is statistically different between the unbound and m^6A bound state during shorter timepoints such as 15 seconds and 2.5 minutes of incubation in deuterium. However, as the proteins were incubated in deuterium for longer times such as 25 minutes and up to ~ 4 hours, the kinetic profiles develop differently. In YTH-1, the isotopic distributions between the states becomes similar with time; however, in YTH-2, the isotopic distribution of the unbound state remains statistically different from the bound state (**Figure 2.12**). The kinetic plots also show converging lines between the states for YTH-1, but the lines representing the unbound versus bound state remain divergent in YTH-2 over our selected time range (**Figure 2.12**). This observation could either be owed to differences in kinetics (i.e. k_{open} versus k_{close}) or due to differences in ligand association/dissociation (i.e. k_{on}/k_{off}). If we

assume the former, it implies that the β4 strand behaves dynamically in YTH-1 (i.e. undergoing fast open/close cycles). This may be justifiable as YTH-1 is predominantly viewed for roles in translation[26], and likely interacts with m^{6}A RNA earlier in the temporal order (compared to YTH-2)[36]. The dynamic nature of the β4 strand, which precedes the binding pocket, could assist in 'scanning' like activities – this is not unusual for the molecular activity of translation-associated proteins (especially because these proteins can also 'pass' the m^{6}A RNA between each other)[26,36]. For YTH-2, the slower rate of 'opening' may correlate with its recognized role in decay and later involvement in the temporal order as reported[36]. However, the latter case might also be the true, and we cannot distinguish the scenarios in our data. Kinetics in HDX can be used as a guide to discriminate different dissociation rates, as per reference [41][41]. Nonetheless, we note that observed differences between the unbound and bound states in the β4 strand converge with time in YTH-1 but remain divergent in YTH-2; and if differing binding kinetics are in play, it correlates well with their differing functions under the prevailing model.

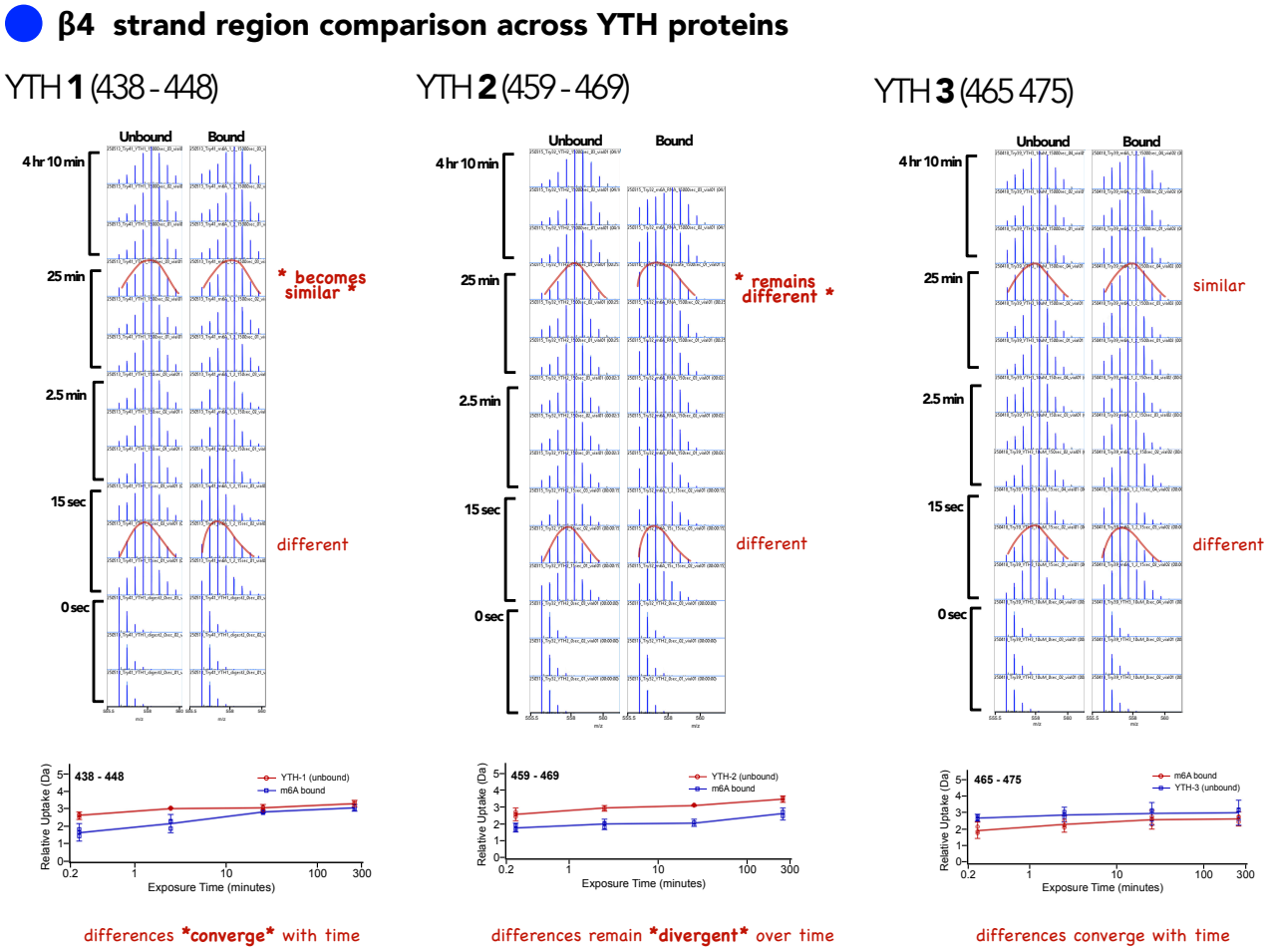


Figure 2.12 | Example of Comparable Peptide from the β4 Strand
Raw isotopic distribution and kinetic plots for peptide 438 – 448 in YTH-1 (which corresponds to residue numbering 459 – 469 in YTH-2 and 465 – 475 in YTH-3) is shown. These are screenshots from Dynamx software with modifications[40].

2.3.5 YTH-1 '*may*' have dynamic regions in the β4 – β5 loop, unlike YTH-2

The β4–β5 loop (also known as the m^{6}A recognition loop) is a recognized important region of the YTH domains and contains the conserved binding pocket. This loop contains two of the three key conserved tryptophan residues thought to play roles in stabilizing direct binding to the m^{6}A RNA ligand, as per crystal structures[31,37]. The presence of a 'tryptophan cage' in the binding

pocket has been emphasized[31]. Residues W465 & W470 in YTH-1 (which corresponds to W486 & W491 in YTH-2 and W492 & W497 in YTH-3) are in the m^{6}A recognition loop and are critical for ligand binding (when these residues were mutated to alanine, m^{6}A RNA binding was disturbed)[31]. The dynamics of this loop have also been emphasized in MD simulations as well[33,34]. It is also thought, due to the importance of this region, that potential inhibitors may interfere with the conformation of this loop to disturb ligand binding[42]. We recognize the importance of the β4 – β5 loop region and were especially intrigued to compare this region across the YTH domains; unfortunately, our peptide coverage also suffered in this regions and our observations are limited to the peptides we did have (none of them were truly 'comparable' across all the YTH domains when all peptides were aligned by sequence in **Figure 2.5**).

The β4 – β5 loop generally showed 'permanent' changes across all YTH domains; however, like the β4 strand discussed earlier, parts of the β4 – β5 loop *may* also be dynamic in YTH-1, but not in YTH-2. Peptides from the m^{6}A recognition loop are highlighted in red in **Figure 2.5**. Generally, we observed permanent stabilization in this loop in the bound state for all YTH domains – 'permanent' changes in structure are those that become more significant with increased time spent in deuterium. For the case of YTH-2, only peptides showing permanent stabilization of the loop were observed (**Figure 2.5** and **Figure 2.7**). However, for YTH-1, some peptides, particularly those containing one or both key tryptophan residues discussed above, appeared to show a change in *dynamic* in observations shown in **Figure 2.6**. First, in our biological run from Aug 20, 2024; we observed that some regions of the m^{6}A recognition loop may be *dynamic*; these peptides are marked by purple/black stars in **Figure 2.6.** However, this was one of our very preliminary data sets and we only had two timepoints; we aimed to attain at least 4 timepoints,

especially when commenting on 'kinetics' (i.e. how peptide profile change over *time*). Then, in our biological replicate from Mar 18, 2025, we were able to reproduce the 'dynamic' behavior of these same peptides over at least four different timepoints (**Figure 2.6**). These 'dynamic' peptides are marked by stars; and kinetic plots are shown for peptides marked by black stars in **Figure 2.13.** The kinetic plots for selected YTH-1 peptides showed that the unbound and m^6A bound state were statistically different during shorter timepoints such as 15 seconds and 2.5 minutes of incubation in deuterium. However, the lines quickly converge from 25 min onwards. Because the observed differences diminish quickly; these peptides are thought to be *dynamic*. This is aligned with data from MD simulations which also suggests that the m^6A recognition loop in YTH-1 is very flexible[33,34]. This kinetic profile would also correlate with the scanning-like translation associated functions of YTH-1 under the prevailing model (as discussed above)[26,36]. However, we also re-iterate again that our observations could alternatively also be a byproduct of difference in ligand association/dissociation patterns instead of dynamics. However, based on the MD simulations, we are likely to believe that these differences could potentially be a byproduct of a change in dynamics[33,34]. Nonetheless, we aimed to reproduce this observation in at least one more biological replicate with at least four timepoints. We collected three additional biological runs in May 2025 (**Figure 2.6**); however, we could not attain coverage in these peptides again, unfortunately. Therefore, we have omitted these peptides from the final *comprehensive* difference plot in **Figure 2.5** due to low confidence. Limited coverage and inability to get truly comparable peptides made it difficult to compare the conformational dynamics of the m^6A recognition loop on a peptide-by-peptide basis across the YTH domains. Nonetheless, we are not aware of any other bioanalytical assays comparing the conformational dynamics between these domains, and even though our data is limited in this important loop region, it is still a step closer to eventually fully unravelling the

mechanism of this loop in the future. Here, we report (for the peptides we did attain) that the m^{6}A recognition loop appears to show permanent structural stabilization across all YTH domains (**Figure 2.5**), and some dynamic behavior was detected for YTH-1 in preliminary observations (**Figure 2.6** and **Figure 2.13**).

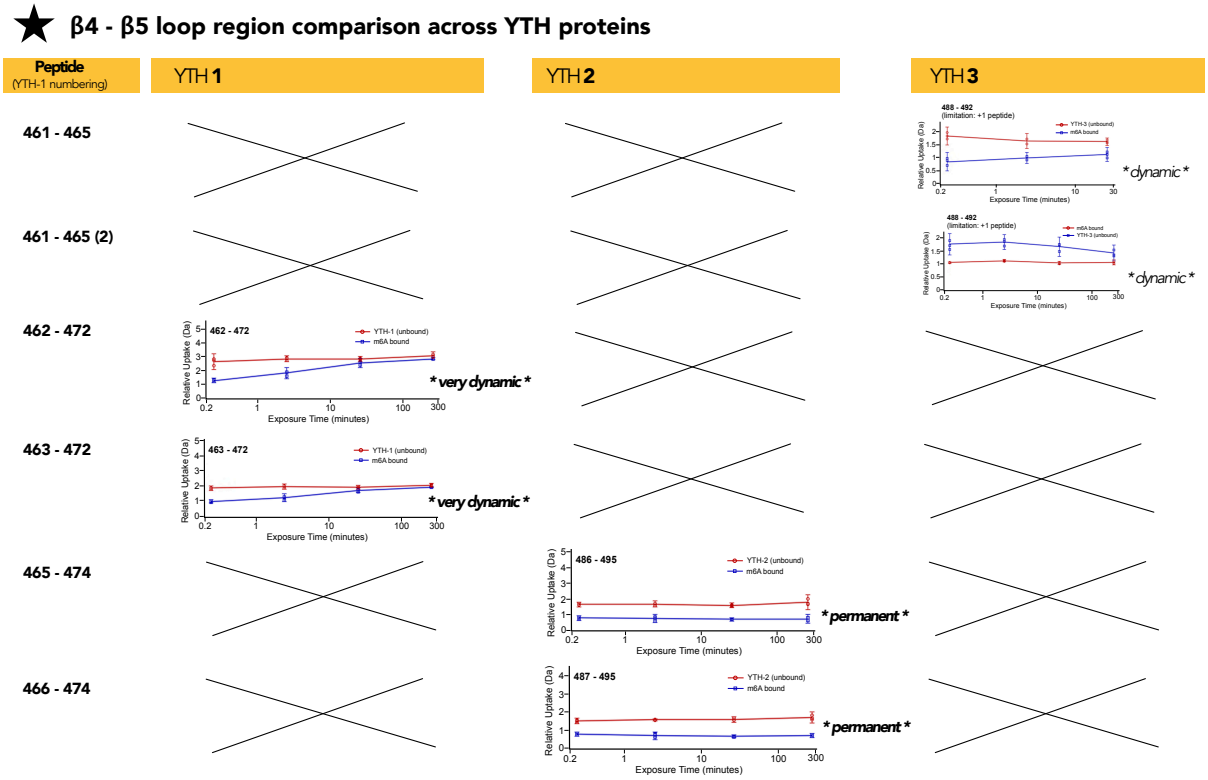


Figure 2.13 | Example of Kinetic Plots of Peptides from the m^{6}A recognition loop
Kinetic plots from DynamX[40] shown for peptides marked by a black star in Figures **2.5-2.7** for YTH 1-3 proteins.

2.3.6 Conformational Dynamics of YTH-3 appear like YTH-1

Conformational dynamics in YTH-3 appear like YTH-1 (this aligns with Zou et al. study as well)[36]. First, like YTH-1, YTH-3 did not show statistically significant differences between its unbound and m^{6}A-bound state in the α2–β3 region (**Figure 2.5**). One of the comparable peptides from α2–β3 region is marked by the red square, and the raw isotopic distribution and kinetic plot

for this peptide has been included in **Figure 2.11**. The raw data also show similarity between the YTH-1 and YTH-3 domains. Second, the β4 strand in YTH-3 also appeared to be dynamic. One of the comparable peptides from the β4 strand region is marked by the blue circle in **Figure 2.5** and the raw isotopic distribution and kinetic plot for this peptide have been included in **Figure 2.12**. In YTH-3, the kinetic plots for the selected peptide show non-overlapping lines between the unbound and bound state at 15 seconds of incubation in deuterium, and the isotopic distribution between the states also appears different at this time point; however, this difference becomes insignificant with increasing time. YTH-3 appears to present similar kinetic profiles to YTH-1 for the selected peptide; however, we note that peptides in YTH-3 generally had higher errors (for example, for the β4 strand, only the earliest timepoint appeared significant in any case and the gray shaded area representing the average error is higher than the bars showing average uptake across all timepoints for YTH-3 in **Figure 2.5**. And we reiterate that any observations made on kinetics and *dynamic* nature of any regions could also be owed to differences in k_{on}/k_{off} instead. Third, for all YTH domains, permanent structural changes were seen in the m^6A recognition loop; however, (similar to YTH-1), some peptides, particularly those containing the key tryptophan residue appeared dynamic in YTH-3. In **Figure 2.8**, peptide marked by the black star, which contained one of the key residues, W492, appeared dynamic. The kinetic plot associated with this peptide is also shown in **Figure 2.13** which shows converging lines between the unbound and bound states for YTH-3 (like YTH-1)**.** Unfortunately, this was a +1 peptide (i.e. poor resolution); we only included this in our preliminary observations due to the importance of peptides with the key Trp residues in the m^6A recognition loop (and such peptides have been omitted from the *comprehensive* difference plot in **Figure 2.5**). Overall, when comparing conformational dynamics in α2 – β3 region, β4 strand, and the β4–β5 loop; YTH-3 (even though it presents higher error)

generally appears like YTH-1. Homology analysis in Zou et al. study also reported similar observations, and we quote, "YTHDF2 is most different, while YTHDF1 and YTHDF3 are more similar to each other"[36]. The Zoe et al. data provide orthogonal confidence in our observation, YTH-1 and YTH-3 appear similar[36].

We do realize that the scope of these YTHDF reader proteins is full of complexities; besides the 'prevailing' and 'unified' model, a 'combined' model is also possible, which is discussed in the attached paper[42]. This is outside the scope of our work. Briefly, when we discuss the function of the three YTHDFs, we are referring to their function in the cytoplasm (which is where they usually reside). However, under different cellular conditions, these proteins may translocate to other areas of the cell and take part in alternative functions and/or behaviors as captured by the combined model (for example, YTHDF2 protects m^6A RNA in the nucleus in response to heat shock)[42]. Basically, in response to different cellular conditions (e.g. heat shock, oxidative stress etc.), these YTHDFs can demonstrate different behaviors – they can form condensates, and under oxidative stress, they can aggregate to stress granules (SGs)[42]. We found it interesting that even in response to various stresses, YTH-1 and YTH-3 respond in similar ways, while YTH-2 responds differently. For example, the condensates of the YTH-2 protein were most different compared to YTH-1 and YTH-3 condensates[42]. In addition, even when SGs are formed, YTH-2 localizes inside the SGs whereas YTH-1 and YTH-3 was found on the periphery of the SGs as per reference [42][42]. Overall, even under the combined model or different cellular scenarios, YTH-1 and YTH-3 appear to behave similarly (compared to YTH-2) which gives us more confidence in our observations (the conformational dynamics of YTH-1 and YTH-3 appear similar).

The conformational dynamics of YTH-3 can also be thought to be 'marginal' between YTH-1 and YTH-2 (this also aligns with Zou et al. study)[36]. For example, in the α2 – β3 region (**Figure 2.5**), the bar representing the relative average uptakes are overall presenting insignificant differences between the unbound and bound states for both YTH-1 and YTH-3 (only YTH-2 showed statistical significance). However, in YTH-3, the height of these bars appears 'marginal' with the longest timepoint presenting statistical significance (while all selected timepoints appear insignificant in YTH-1). In the Zou et al. (2023) study, localization patterns of YTH-3 also appeared to be 'marginal', and we quote their observation: "YTHDF3 in the margin between CNOTs and eIFs" (decay and translation associated proteins, respectively)[36]. Under the prevailing model, YTH-1 is involved in translation, YTH-2 in decay, and YTH-3 in *both*; the detection of 'marginal' conformational dynamics for YTH-3 aligns with the predictions of the prevailing model.

2.4 Conclusions

The function of the three YTHDF proteins has been debated in the scientific community; various studies from the He lab attached supported the prevailing view[26,28,29,36], while authors such as Zaccara et al.[35] presented the unified view. According to Arguello et al. study and various crystallography studies attached, the mechanism of YTH binding to m^{6}A RNA is very conserved and similar[25,31,37]. This was of interest to us – because even if the conformational dynamics in the binding pocket are similar; the allosteric changes could potentially be different between the YTHDFs, suggesting different functions. HDX-MS is especially useful for detecting allostery (**Figure 2.14**). Here, we observed allosteric stabilization in α2–β3 region upon m^{6}A binding in YTH-2, but not in YTH-1 and YTH-3. We observed that YTH-1 and YTH-3 appear similar, but YTH-3 can also be thought to present 'marginal' conformational dynamics between YTH-1 and YTH-2. All of these observations align with various conclusions from the He lab, and the prevailing view: YTH-1 for translation, YTH-2 for decay and YTH-3 for both (i.e. 'marginal') [26,28,29,36]. Less importantly, the β4 strand could also possibly be less dynamic in YTH-2 (compared to YTH-1 and YTH-3). And unfortunately, we could not make conclusive observations on the m^{6}A recognition loop, but occasionally, some peptides in YTH-1 appeared dynamic in this loop region consistent with its expected flexibility in molecular dynamics studies[33,34]. Finally, we re-iterate our limitations: a dimeric species was present on gels for YTH-2 and YTH-3; the size of YTH domains is different due to differing levels of truncations on the C-terminal helice that was purchased, we occasionally have two (instead of three technical replicates) for some timepoints, in-house pepsin column presented challenges in consistently getting similar quality of peptides,

limited coverage (means limited view of conformational dynamics), m^{6}A RNA sequence of choice (derived from Arguello et al.) may not necessarily reflect *in vivo* scenarios[25], and human-error is also possible. Within the confines of our limitations, our data may support the prevailing view of the function debate.

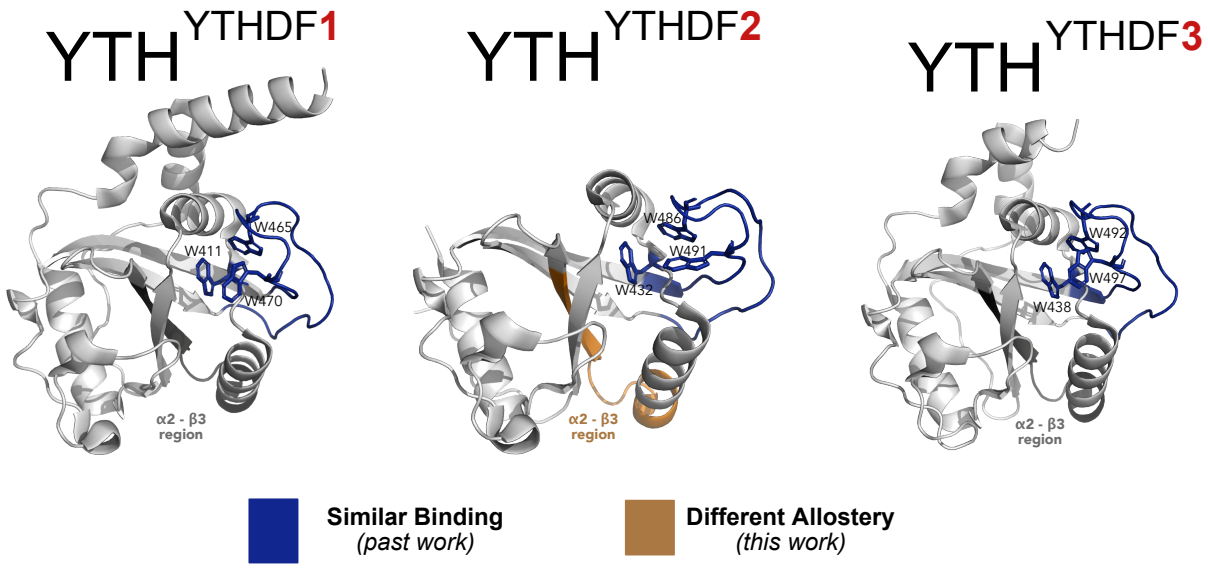


Figure 2.14 | HDX Perspective on the YTHDFs Function Debate
Chapter 2 summary figure – hydrogen-deuterium exchange mass spectrometry perspective on the function debate of the three YTHDF proteins. Key observations are plotted on PDB structures (4RCJ, 4RDO, 6ZOT)[31,33,38].

2.5 Methods

2.4.1 General

We purchased the plasmids for the YTH-1, YTH-2, and YTH-3 domains from addgene with the following plasmid ID numbers, respectively: 64653, 228776, 228777. YTH-1 plasmid was deposited by Cheryl Arrowsmith; and YTH-2 and YTH-3 was deposited by Zhou Lab and discussed in reference [43][43]. Each YTH domain is N-terminally histidine tagged, but the C-terminal helice is differently truncated between the YTH domains. The protein sequence for the

three YTH domains can be found in **Table 2.1**. Each YTH protein was expressed in BL21 cells at 18℃ for 16-20 hours. Ni^{2+} affinity chromatography was performed to purify the histidine (His) tagged YTH proteins over 15 – 500mM imidazole gradient. The YTH proteins were stored in the following buffer: 20mM Tris, 0.4M NaCl, 300 mM – 500 mM imidazole, 5% glycerol, pH 7.5. Final concentration of 1mM DTT was added in attempts to reduce di-sulfide linked dimers. The samples were stored at -80℃ for storage. Proteins were characterized through SDS-PAGE and Native-MS.

2.4.2 m^6A RNA sequence

The m^6A RNA sequence is as follows: GGCCG(m6A)TCTGA (visually depicted in **Figure 2.2**); the sequence identity was retrieved from the Arguello et al. study[25]. This sequence was purchased from horizon (Dharmacon Custom RNA Synthesis). The RNA ligand was suspended in RNase-free water and stored at -80℃. It is recommended to wipe lab stations with RNaseZap Spray and to use barrier-tip pipettes when handling RNA (although we did not always practice this, except when making initial stocks).

2.4.3 Native MS

YTH proteins were buffer exchanged into 100 mM ammonium acetate (A637) using the Millipore Amicon Ultra centrifugal filter, 10 kDa MW cut-off. For YTH-1, ~10 μM protein was ionized with electrospray ionization and with trapping ~ 3500 m/z. For YTH-2 and YTH-3, 10 μM protein concentration was flowed through the Hamilton 500 ml HPLC syringe using the Harvard Pump11 Elite Syringe Pump at 20 μl/min flow rates, and to conduct electrospray ionization, a nano spray source was set up on the Waters G2-S Synapt instrument.

2.4.4 Conventional HDX

For all the sample preparations and each of the YTH proteins (YTH-1, YTH-2, and YTH-3), similar YTH:m^{6}A initial concentrations (i.e. before dilution with deuterium) and ratio of 10 μM : 20 μM was used, with the exception of the following YTH-1 runs: Aug 20, 2024 and May 22, 2025 experiments (difference plots for all individuals runs are shown in **Figure 2.6-2.8**). In the Aug 20, 2024 and May 22, 2025 runs, a YTH-1:m^{6}A ratio of 5 μM : 20 μM and 20 μM : 40 μM was used, respectively. Nonetheless, the percentage bound (in diluted 90% D_2O buffer), as calculated using the equation also noted in reference [41][41], was kept within a similar range, 76 - 86 % bound. This was calculated using the m^{6}A RNA sequence validated in the Arguello et al. study with a K_D of 0.51 μM for YTH-1, and similar K_D is assumed for YTH-2 and YTH-3 (m6A RNA sequence shown in **Figure 2.2**)[25]. For the HDX reaction, the composition of the bound states was as follows (for most experiments): 10 μM YTH, 20 μM m^{6}A RNA, and 5 mM $MgCl_2$ added to YTH storage buffer (as per above). In some case, 1mM DTT was added to reduce disulfide bonds. For the unbound state, the same conditions were used, except, instead of adding the m^{6}A RNA ligand, the same volume of the blank solvent that the RNA ligand is stored in was added (i.e. RNAse free water). Conventional HDX-MS was performed for each sample as per below.

Each prepared sample was labeled in 90% D_2O buffer, and the following timepoints were used: 15 seconds, 2.5 minutes, 25 minutes, and 250 minutes. The HDX reaction was quenched in 100 mM phosphate buffer (pH 2.5) for 2 minutes. For protein digestion, pepsin (porcine gastric mucosa, Sigma-Aldrich) was linked to NHS activated beads (Pierce, Sigma-Aldrich) in-house and was packed in column in-house. Peptide coverage maps are plotted on the crystal structures shown in **Figure 2.5**. Peptides were eluted through the C18 column over 5-35% acetonitrile gradient.

Select Series Cyclic IMS (Waters) instrument was used. Peptides were ionized with electrospray ionization. MS/MS was done with CID to produce (cyclized) b and (linear) y fragments (as common for bottom-up proteomics with CID). The ToF mass analyzer is included with Select Series Cyclic IMS (Waters) instrument for separation of the generated fragments. ProteinLynx Global Server (PLGS) was used for peptide identification, and Dynamx for peptide analysis. Aligned with the recommendations of the attached paper[22], we have prioritized data transparency.

Chapter 3 | Mechanism of Recent YTH Inhibitors and the Dynamics of the Intrinsically Disordered DF Domain

3.1 Abstract

The YTH proteins are notorious for implications in various cancers – YTHDF1 was also framed as a 'tumour promoter' protein with association in over 15 different cancers. Currently, there are limited inhibitors for the YTH proteins, and their mechanism remains elusive. We studied the inhibitory mechanisms of Ebselen, Tegaserod, and Salvianolic Acid C against YTH. Our ion mobility analysis, in addition to millisecond and conventional HDX-MS, indicated increased collisional cross-section and disorder in the YTH domain from Ebselen. Reduced electron density has also been reported in previous studies and YTH-Ebselen cocrystals. For Tegaserod, we conducted millisecond HDX, which is ideal for characterizing weaker interactions. However, we did not detect statistically significant findings for this compound. Salvianolic Acid C was of particular interest to us because the explanation for the selectivity of Salvianolic Acid C for YTH-1 over YTH-2 remains to be articulated. Our HDX-MS attempts revealed unique conformational dynamics between YTH-1 versus YTH-2 from Salvianolic Acid C. It has been challenging to achieve selectivity between the YTH domains due to their structural similarity. The intrinsically disordered N-terminal DF domain has never been crystallized and is especially unique among the three YTHDFs. Elucidating the dynamics of the DF domain can also facilitate the progress of selective inhibitors. As crystallography has failed, we aimed to reveal the intricacies of this domain using HDX-MS. Expression and purification of the full-length YTHDF1 presented various challenges; nonetheless, we successfully purified the full-length YTHDF2 and YTHDF3 proteins. We have included our HDX-MS data for YTHDF2 when bound to m^6A RNA (endogenous ligand) versus the Ebselen inhibitor. The m^6A RNA likely induces the 'Type 2' mode of action in the DF2 domain. Ebselen induced disorder in the YTH domain, but no detectable differences in the DF

domain. All results presented in this chapter are preliminary and would require extensive reproduction as a future step.

3.2 Introduction

The YTH proteins are notorious for their implications in various cancers; we currently have limited inhibitors, and their mechanism of action remains elusive. Sikorski et al. (2023) review the function of YTHDFs and, notably, their involvement in diseases. The role of the three YTHDFs in numerous malignancies has been summarized (refer to Box 2 in reference [42])[42]. For example, YTHDF1 (due to its recognized role in translation) is thought to promote tumorigenicity by accelerating the translation of cancer-associated proteins[27]. In breast cancer, patients expressing high levels of YTHDF1 had significantly ($p = 0.011$) lower overall survival than patients expressing lower levels[44]. Knockdown of YTHDF1 improved survival rates[44]. YTHDF1 was also framed as a 'tumour promoter' protein with dysregulations reported in over 15 different cancers[42,44]. While we have limited inhibitors, the following few have gained interest: Ebselen, Tegaserod, and Salvianolic Acid C (SAC)[42].

We speculated that Ebselen would inhibit by promoting disorder in all the YTH domain-containing proteins, as per previous reports and co-crystals[45]. Micaelli et al. (2022) conducted high-throughput screening (HTS) of ~2000 FDA-approved drugs and presented Ebselen as a covalent and non-specific inhibitor for all YTH domains[45]. Ebselen has previously also been reported for the context of the main protease M^{PRO} enzyme of the SARS-CoV-2 virus and inhibits by forming selenenyl-sulfide bonds with the thiol group of the cysteine residue[46]. For the case of the YTH domain, it directly blocks m^6A RNA–YTH interaction by forming a covalent bond with

C412, which is a residue in spatial proximity to the hydrophobic binding pocket of the YTH protein[45]. The binding of Ebselen close to the binding pocket also distorts the dynamics of the β4–β5 loop[45]. The β4–β5 loop normally undergoes induced-fit to accommodate m^6A in the binding pocket. However, reduced electron density has been reported in this loop in the YTH-Ebselen cocrystals as well[45]. We anticipated that Ebselen would present an accelerated exchange profile in our HDX-MS data.

Like above, Tegaserod is also predicted to inhibit by directly blocking the hydrophobic binding pocket of the YTH protein[47]. The Hong et al. (2023) study conducted virtual screening and molecular docking analysis using the Glide Schrödinger software and identified Tegaserod as an inhibitor of the YTH proteins[47]. We have limited knowledge of the exact mechanism of action of Tegaserod, partly because there are no cocrystals of Tegaserod with YTH. Nonetheless, we predicted accelerated H/D exchange in the β4–β5 loop, because inhibitors that bind spatially close to the binding pocket usually induce distortions in the conformation of this loop.

SAC was of particular interest to us because the explanation for the selectivity of SAC for YTH-1 over YTH-2 remains to be articulated, and bridging this gap can facilitate the design of paralog selective inhibitors[42]. The first report of SAC as an inhibitor of the YTH proteins was presented in the Zou et al. 2023 study – interestingly, also from the Chuan He Lab (who is recognized for his significant contribution in this field, as also discussed in Chapter 2)[48]. Fragile X Mental Retardation Protein (FMRP) suppresses YTHDF1 translation under physiological conditions; the function of FMRP is pivotal in ensuring balanced translation levels in the cell. However, FMRP is absent in Fragile X Syndrome (FXS), leading to 'hyper-translation' by YTHDF1[48]. Zou et al. conducted HTS and reported SAC as an inhibitor of the YTHDF1 protein.

They demonstrated that inhibition by SAC recovered cellular physiology to 'normal translation' from a 'hyper-translation' state. They note that the K_D of SAC for YTHDF1 is in the range of 5.3 - 6.3 μM. Interestingly, they observed significant selectivity of SAC for YTHDF1 over YTHDF2; the IC_{50} of SAC for YTHDF1 was ~1.4 μM versus ~30 μM for YTHDF2. SAC forms hydrogen bonds with the R506 residue of YTHDF1; however, this residue is not present in YTHDF2[42,48]. The mechanism of action of SAC and its reported selectivity for YTH-1 over YTH-2 remains to be elucidated. The Wang et al. (2024) study identified a new pan-inhibitor for the YTH proteins and compared their findings against SAC (likely due to the emerging relevance of this inhibitor); they note that SAC also had selectivity for YTHDF3[43]. We found this quite relevant to us; earlier, we discussed that YTHDF1 and YTHDF3 presented comparable conformational dynamics (Chapter 2), and it appears that they share similar selectivity for the SAC inhibitor as well (this further reinforces our earlier conclusions). The underlying mechanism for the selectivity of SAC for YTH-1 over YTH-2 is unclear; here, we aim to uncover this using HDX-MS.

The intrinsically disordered N-terminal DF domain has never been crystallized and is especially unique among the three YTHDFs; elucidating the dynamics of the DF domain can also facilitate the progress of selective inhibitors. The DF domain of the YTHDF proteins represents an intrinsically disordered region (IDR). IDRs are present in ~1/3 of all proteins and play numerous roles in driving biological activities, such as participating in protein-protein interactions[4]. For the case of the YTHDFs, the DF domains are thought to recruit downstream proteins to dictate the fate of target m^6A RNA. YTHDF1 can promote translation of the m^6A RNA via the recruitment of translation initiation factors such as elF3A and elF3B[26,28]. And YTHDF2 can regulate degradation through association with degradation-associated proteins such as CNO1[26,28]. According to Wang

et al., the DF2 domain also plays roles in the translocation of the m^{6}A RNA into p-bodies[26,28]. It has been challenging to achieve selectivity between the YTH domains due to their structural similarity. As the DF domains contain numerous non-conserved residues between the YTHDFs, we can leverage this towards the design of paralog-selective inhibitors. It is essential to also explore the dynamics of the DF domain from m^{6}A RNA (endogenous ligand) to rationalize how their normal mode of action can be inhibited[42]. In *Mmi* (which is the YTHDF protein equivalent in yeast), the DF domain wraps around the YTH domain to further stabilize the interaction with m^{6}A RNA[49]. Based on this, we hypothesized that the intactness of the hydrogen bond network and decreased solvent accessibility in the m^{6}A RNA-bound state would lead to the observation of reduced deuterium uptake in our HDX data[49]. Due to the importance of the DF domain in driving biological activities and because crystallography has failed, we aim to explore the intricacies of this domain using HDX-MS – useful for extrapolating the structure-activity relationship.

Here, we conduct HDX-MS to uncover the mechanism of recent YTH inhibitors, as well as the dynamics of the intrinsically disordered DF domain. First, we explored the inhibitory mechanism of action of Ebselen, Tegaserod, and SAC against YTH. For Ebselen, we conducted ion mobility analysis as well as millisecond and conventional HDX-MS for orthogonal evidence. HDX was conducted with YTH-1 protein, but ion mobility was performed with YTH-2 instead (as the signal intensity for YTH-1 was not optimal). Because Ebselen is a non-specific inhibitor, all the YTH domains likely share a comparable mechanism of action[45]. For Tegaserod, we only conducted millisecond HDX using the in-house apparatus developed by Dr. Joseph Anacleto, as described in reference [21][21]. Millisecond HDX is ideal for characterizing weaker interactions. For SAC, we conducted conventional HDX for YTH-1 versus YTH-2. All three inhibitors are recent

(i.e. they were described for YTH in 2022 or later)[45,47,48]. We aimed to advance their understanding through HDX-MS and hypothesize that our data may present accelerated H/D exchange in the β4–β5 loop. Second, we explore the dynamics of the DF domain using HDX-MS. We have included our HDX-MS data for YTHDF2 when bound to m^6A RNA (endogenous ligand) versus the Ebselen inhibitor. We expected reduced deuterium uptake in the DF domain from m^6A RNA binding as per previous studies on *Mmi* (discussed above)[49]. We predict our findings may be relevant to the development of paralog-selective inhibitors. We emphasize that all data presented in this chapter are preliminary and would require extensive reproduction as a future step.

3.3 Results and Discussion

3.3.1 The α3 helix shows the most disorder from Ebselen at both millisecond and conventional timepoints

Ebselen induced significant disorder in the YTH-1 protein, including the β4–β5 loop region (as predicted), and the most accelerated H/D exchange was observed in the α3 region at both millisecond and conventional (seconds – minutes) timepoints (**Figure 3.1**). The difference plots are included in **Figure 3.1,** and both the β4–β5 loop region as well as the α3 region are marked on the plot; in addition, various regions of the crystal structure are assigned numbers to show which regions correspond to which bars on the difference plot. We were surprised by the comparability of our detected conformational dynamics at conventional timepoints (collected using the Select Series Cyclic IMS) and millisecond timepoints (collected with the in-house apparatus on G2-S Synapt)[21]. We observed accelerated H/D exchange kinetics in the β4–β5 loop region; this observation matches with previous studies. As discussed earlier, reduced electron density was

observed in the β4–β5 loop in the Ebselen-YTH cocrystal[45]. We also observed that the most accelerated H/D exchange occurred on the α3 region (while the β4–β5 loop only showed subtle increases in uptake in comparison). The α3 region approximately spans from residue 513 to 527 and consistently showed pronounced destabilization from Ebselen binding at both millisecond and conventional timepoints. Ebselen appears to promote global disorder in the YTH-1 protein, and it is not unusual to observe a less localized profile in HDX[1]. Overall, both our millisecond and conventional timepoints show comparable data, accelerated HDX in the β4–β5 loop region matched with cocrystal predictions, and global disorder is seen with pronounced destabilization in the α3 region.

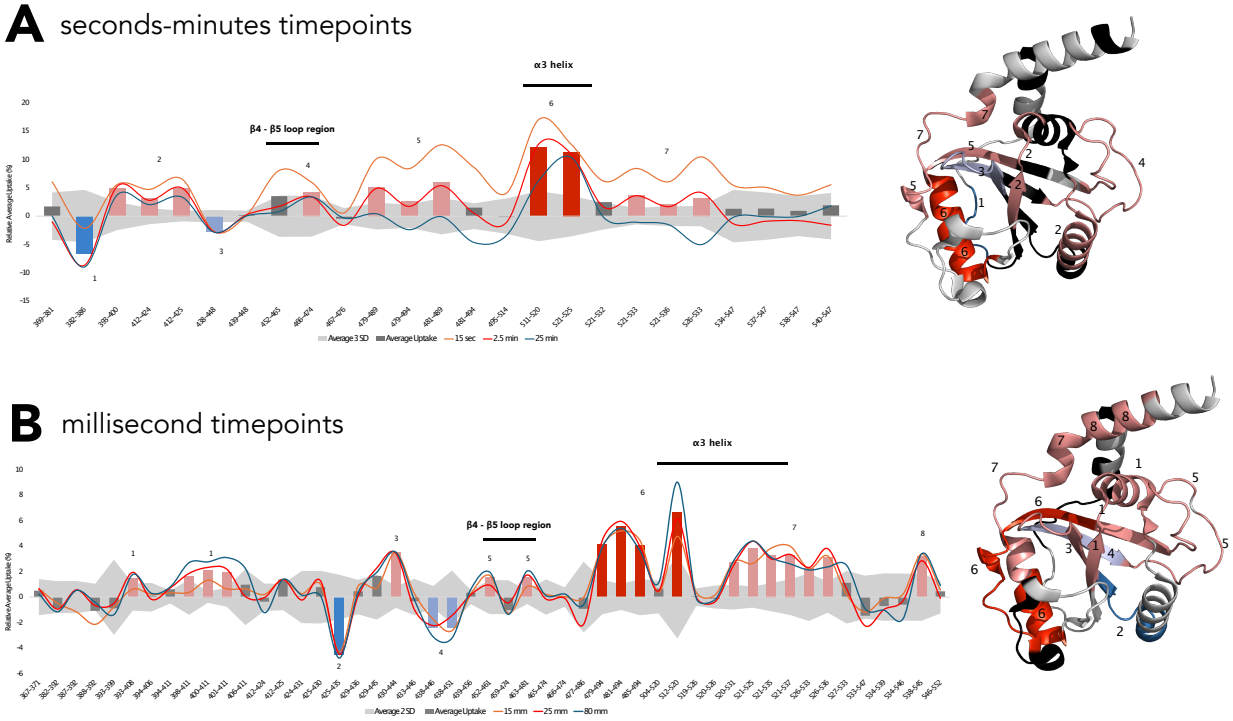


Figure 3.1 | Ebselen Induced Disorder in YTH-1 at Conventional (A) and Millisecond (B) timepoints.
The bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the Ebselen-bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints as indicated under each plot. The gray shaded area indicates the 3 σ error for conventional timepoints. Numbers are included on the PDB structure (4RCJ) to show the corresponding location on the difference plot[31]. Red colors indicate accelerated exchange, and

blue shades indicate decreased uptake in the bound state relative to the unbound state. Note that regions 2, 3 and 4 on the plot shared overlapping peptides displaying differences in different magnitudes and directions, making crystal painting a bit subjective (see the difference plot when interpreting protein regions 2-4).

3.3.2 Ion mobility suggests an increased collision cross-section of the YTH domain from Ebselen binding

We detected binding between YTH-2 and Ebselen in our Native MS attempts. First, we intended to acquire Ebselen-bound state natives for YTH-1; however, the signal intensity of YTH-1 was not optimal for bound state natives (or for ion mobility later). Hence, we studied Ebselen with YTH-2 instead; fortunately, YTH-2 showed significantly resolved and intense peaks at approximately ~ e^5 counts. In **Figure 3.2,** YTH-2 unbound (pink circles) and Ebselen-bound (pink and red circles) natives are aligned (note that in the unbound state, the same volume of dimethyl sulfoxide (DMSO) solvent was used as a negative control). As Ebselen is a covalent inhibitor, 30-minute incubation was allowed to ensure complete binding. In the bound state, almost all the peaks associated with the monomeric YTH-2 had diminished, and peaks corresponding to the bound complex became prominent (**Figure 3.2).** In our bound state, we noticed that multiple Ebselen compounds may be binding to a single YTH-2 protein. We detected peaks corresponding to YTH-2 bound to one, two, and three Ebselen compounds (only the most intense peaks where three Ebselen compounds are bound are labelled in **Figure 3.2**). One limitation of CRM and ESI-MS is the artifact of observing non-specific binding (i.e. ligand and protein may be non-specifically bound in the same droplet as it shrinks). Therefore, interactions that occur in the gas phase may not always accurately represent solution-phase chemistries. Alternatively, this can potentially be a byproduct of high Ebselen concentration used; the concentrations and ratios usually need to be optimized to maximize the likelihood of a 1:1 stoichiometry. In our case, we used a 1:10 ratio of YTH-2 and Ebselen; the ratio and/or incubation times could have been optimized further.

Nonetheless, due to the small size of the Ebselen molecule (relative to the protein), the binding of multiple Ebselen is *not* expected to compromise IMS, so we proceeded with this next.

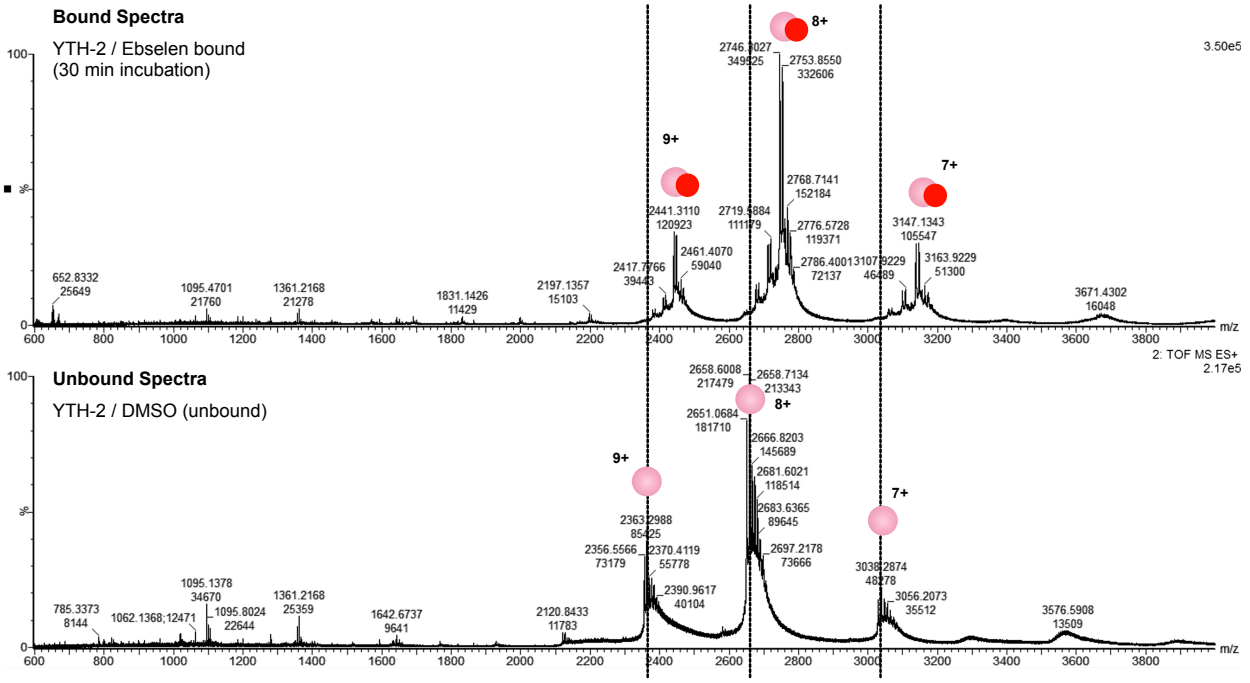


Figure 3.2 | Ebselen-bound YTH-2 Natives
Bound spectrum for YTH-2 with Ebselen is shown on top, and the unbound spectrum is aligned based on m/z below. 1:10 YTH-to-Ebselen was used for the bound state; for the unbound state, the same volume of DMSO solvent was used with YTH-2. Each sample was incubated for 30 minutes prior to data acquisition. Pink circles indicate unbound YTH-2 monomer. Pink and red circles indicate the YTH2 protein bound to Ebselen at a 1:3 stoichiometry; peaks corresponding to 1:2 and 1:1 binding are also visible (not labelled).

Ion mobility data suggests increased collisional cross section (usually indicative of more disorder) in the Ebselen-bound state, which provides orthogonal confidence in our HDX data above. Ebselen does not impact the drift times (due to its small size); however, if the binding of Ebselen induces disorder in the YTH protein, it will shift drift times. In IMS, proteins that adopt a disordered conformation present a higher collisional cross-section, which leads to delayed drift time measurements. Proteins with higher CCS collide with background gases more, which slows

their movement. Compared to the unbound state, we observed that the drift time for YTH-2 is slower in the Ebselen-bound state. In **Figure 3.3**, we have presented our drift times comparisons for the +7 and +9 charge. For the +7-charge state, the drift time of YTH-2 (with the same volume of DMSO as a negative control) is 8.18 ms, whereas in the bound state, the drift time of YTH-2 bound to Ebselen is 8.49 ms – the protein travels significantly slower here (as per increased disordered and CCS). Similar observations were noted for the +9-charge state (i.e. drift time is delayed in the presence of Ebselen). Previous crystals, our millisecond and conventional HDX, and ion mobility observations suggest that Ebselen inhibits by increasing disorder in the YTH protein.

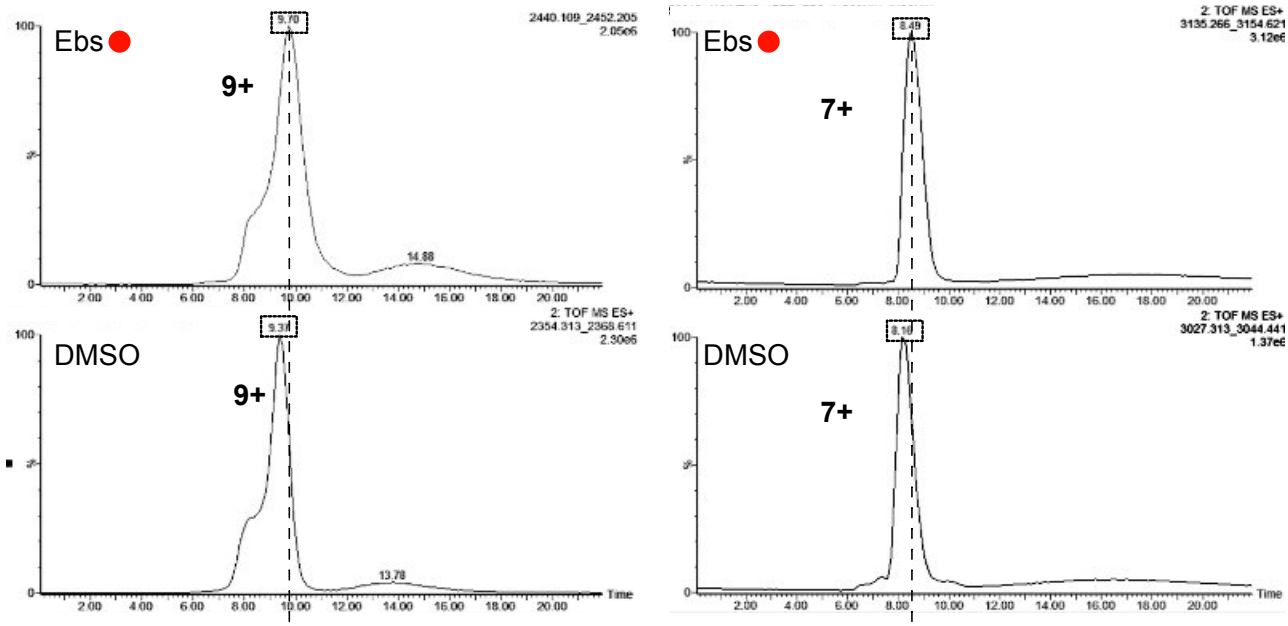


Figure 3.3 | Ebselen-bound YTH-2 Ion Mobility Spectrometry
Bound spectrum for YTH-2 with Ebselen is shown on top (with red circles indicating Ebselen-bound state), and the unbound spectrum is aligned below. 1:10 YTH-to-Ebselen was used for the bound state; for the unbound state, the same volume of DMSO solvent was used with YTH-2. Each sample was incubated for 30 minutes prior to data acquisition. IMS data is included for two charge states: +9 (left), and +7 (right). Dashed lines are used to highlight shifts in drift times between unbound and bound states.

3.3.3 No statistical significance for Tegaserod with millisecond HDX

We did not detect statistically significant observations for the YTH-1 protein with Tegaserod. The difference plots are included in **Figure 3.4,** and all the uptake bars were within error at 2-fold standard deviation propagated error (2 σ) for millisecond timepoints. We generally use 3 σ for conventional timepoints and 2 σ for millisecond measurements. We only detected reduced deuterium uptake in one peptide, 425 – 435. However, we did not have any redundancy in this region, and it is likely that it may just be an outlier. More importantly, as this was just preliminary data, we only used a 1:2 ratio. In the future, using higher ligand concentrations such as a ratio of 1:10 may favor ligand-protein complexation more. Here, we did not detect statistical significance H/D exchange profiles for Tegaserod in our preliminary data – this does not mean that binding does not occur, it may just be outside our detection limit (or needs to be optimized experimentally).

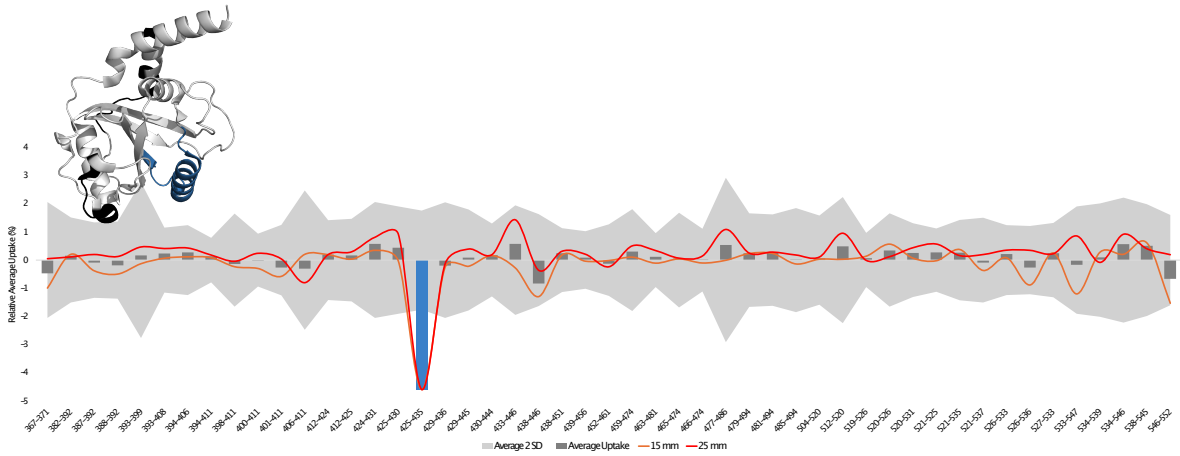


Figure 3.4 | Tegaserod-bound YTH-1 Conformational Dynamics
The bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the Tegaserod-bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints as indicated under each plot.

The gray shaded area indicates the 2 σ error for millisecond timepoints. Sky blue color indicates statistically significant decreases plotted on the PDB structure (4RCJ)[31].

3.3.4 Salvianolic Acid C, selective inhibitor, 'may' induce unique detectable differences between YTH-1 versus YTH-2

SAC induced unique conformational dynamics in YTH-1 versus YTH-2. The difference plots are included and mapped on the crystal structures for YTH-1 and YTH-2 in **Figure 3.5**. In YTH-1, peptides covering the α2 and α4 regions showed significant stabilization from SAC, but no statistically significant changes were observed in these regions for YTH-2. Areas which showed reduced deuterium uptakes are painted in light blue on the crystal structures (**Figure 3.5**). SAC was predicted to form hydrogen bonds with R506 residue of YTH-1, as per previous studies[42,48]. Although we did not detect statistical significance in peptides containing this residue over our selected timepoints, we did observe significant decreases in peptides ranging from 521-536, which are spatially close to R506. For YTH-2, an accelerated H/D exchange profile was observed in the β4–β5 loop region (**Figure 3.5**). Regions which showed accelerated deuterium uptakes are painted in red on the crystal structures (**Figure 3.5**). The β4–β5 loop normally undergoes induced-fit to accommodate m^6A in the binding pocket; it is not unusual that an inhibitor can act by inducing distortions in the conformation of this loop. The magnitude of our average uptake bars was slightly above the error at 3 σ (i.e. 99.7% probability that the signal is real in properly processed data; in reality, this may be lower). YTH-2 data showed the most statistical significance at 2.5 minutes; the timepoint selection can be optimized in the future to enhance the magnitude of the signals. In our preliminary HDX data, we observed that SAC acts via allosteric stabilization of the α2 and α4 regions in YTH-1, and distortions in the β4–β5 loop of YTH-2; reproducing this data can eventually help address current gaps in the action and selectivity of SAC.

Due to the differences in SAC's affinity for YTHDF1 versus YTHDF2, it was challenging to identify truly comparable conditions for HDX. The IC_{50} of SAC for YTHDF1 was ~1.4 μM versus ~30 μM for YTHDF2. While functional responses from cellular assays and IC_{50} values do not always correlate with K_D or molecular-level chemistries, we do assume that the interaction of SAC with YTHDF2 is weaker based on approximation from the IC_{50} values (and as per our preliminary HDX attempts). Weaker interactions often necessitate the use of higher ligand concentrations (and more DMSO content). Small molecules were stored in DMSO; and higher DMSO content can impact the protein's structure (i.e. excess DMSO can unfold the protein). 2% final DMSO concentrations were used for YTH-2; whereas 1% final DMSO content was present in the YTH-1 samples. These variations are likely negligible because we are *not* comparing the YTH-1 and YTH-2 domains *between* each other. Rather, we study the dynamics for each domain between the unbound state versus the bound state. Besides this, the selection of timepoints was also different between the YTH-1 and YTH-2 domains. Shorter timepoints were deemed more suitable for elucidating weaker interactions, such as SAC with YTH-2, whereas longer timepoints were employed for YTH-1. In the future, it is important to optimize experimental conditions to ensure that conformational dynamics are as comparable as feasible.

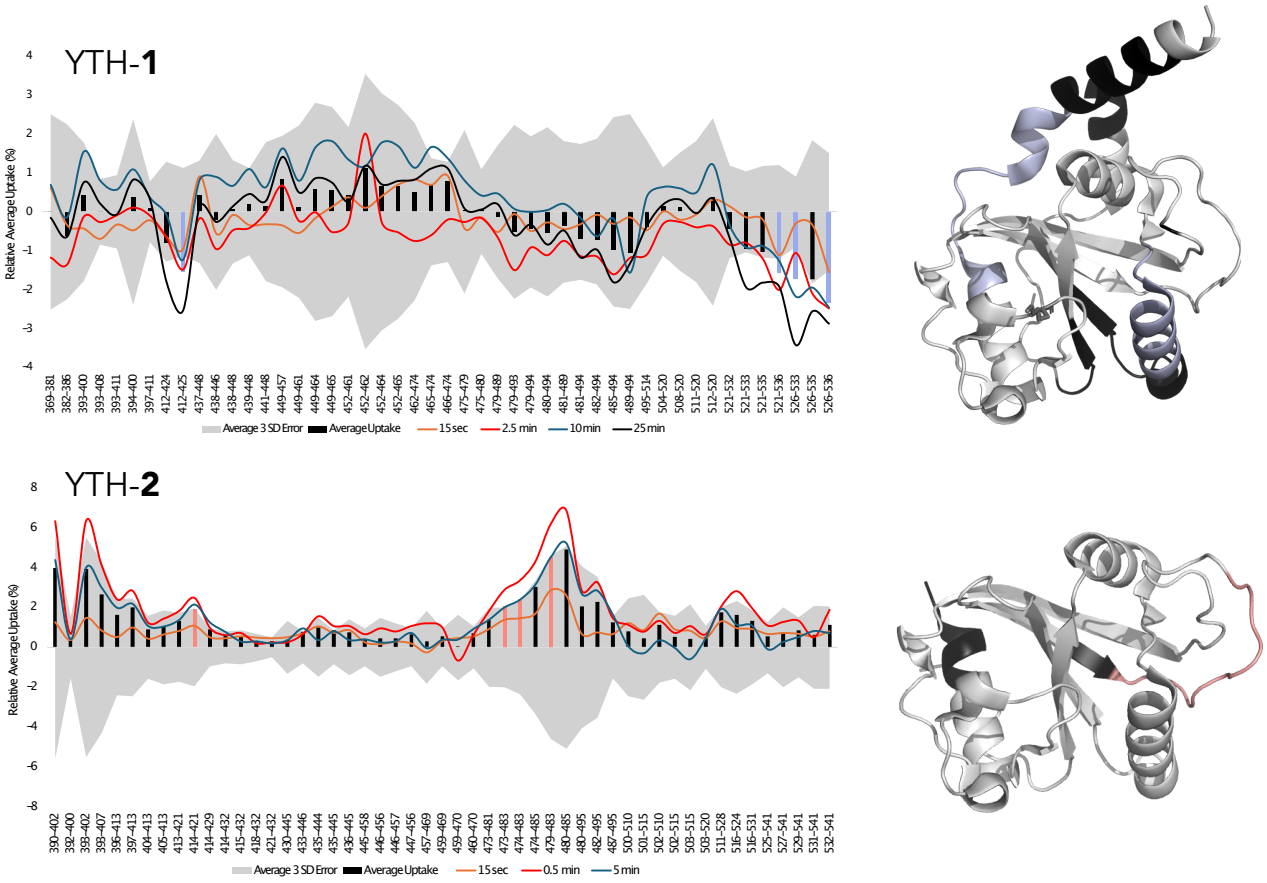


Figure 3.5 | SAC-induced Conformational Dynamics in YTH-1 (A) versus YTH-2 (B)
The bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the SAC-bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints as indicated under each plot. The gray shaded area indicates the 3 σ error for conventional timepoints. Red color indicates accelerated exchange, and blue indicates decreased uptake in the bound state relative to the unbound state on the PDB 4RCJ and 4RDO[31,38].

3.3.5 YTHDF1-3 Full Length Protein Expression and Purification

YTHDF1-3 were characterized on SDS-PAGE (**Figure 3.6**). The protein sequence of the YTHDF1-3 proteins can be found in **Table 3.1**. Glutathione-S-Transferase (GST) tagged YTHDF1 matched with the expected monomeric mass of 87.5 kDa on the gel; however, a major degradation

profile was visible. We predicted that the GST tag was dissociating from the YTHDF1 protein, and various truncations were seen in the lower MW **range (Figure 3.6).** Nonetheless, we successfully purified the full-length YTHDF2 and YTHDF3 proteins. N-terminally His-tagged YTHDF2 and YTHDF3 matched with their expected monomeric mas of 64.5kDa and 66.3kDa, respectively, and exhibited ~70% purity on gels (**Figure 3.6**). We have included our HDX-MS data for YTHDF2 when bound to m^{6}A RNA (endogenous ligand) versus the Ebselen inhibitor.

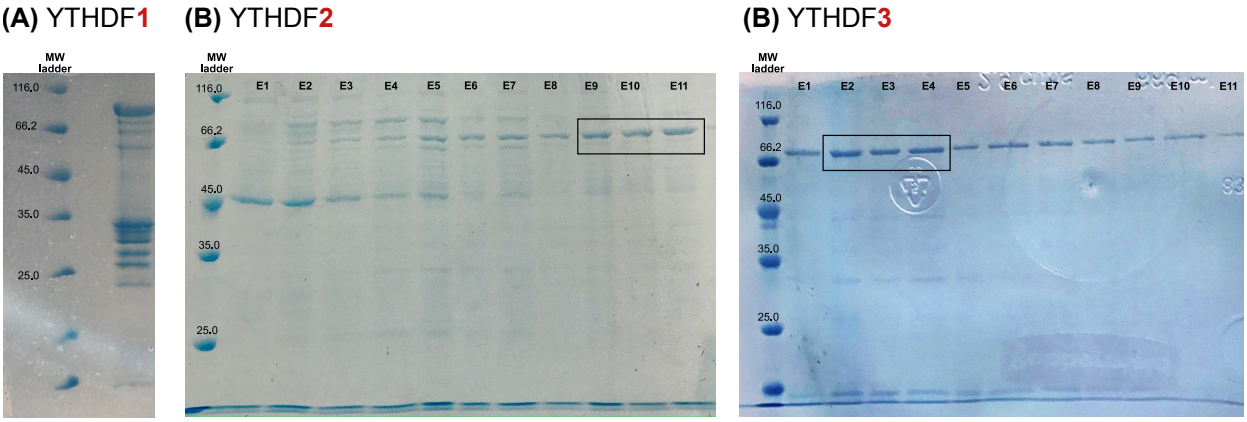


Figure 3.6 | YTHDF1-3 SDS-PAGE
GST-YTHDF1 (A), His-YTHDF2 (B), and His-YTHDF3 (C) are shown on SDS-PAGE. All proteins presented bands matching their expected monomeric mass: 87.5 kDa, 64.5kDa, 66.3kDa. GST tagged YTHDF1 presented a major degradation profile in the lower MW range, the GST tag likely fell off, and various truncations were seen. YTHDF2 and YTHDF3 appeared around ~70% pure on SDS-PAGE.

Table 3.1 | YTHDF 1-3 Protein Sequences.
GST-YTHDF1 (top), His-YTHDF2 (middle), and His-YTHDF3 (bottom) sequences are shown. Protein tags (GST/His) and linkers are in blank, the DF domains are shown in blue, and YTH domain in green.

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPK ERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAF PKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPRGSPEFMSATSVDTQRTKGQDNKVQNGSLHQKDTVHD NDFEPYLTGQSNQSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPWSTAGDPPIPYLTTYGQLSNGDHHFMHDAVFGQPGGLGNNIYQHR FNFFPENPAFSAWGTSGSQGQQTQSSAYGSSYTYPPSSLGGTVVDGQPGFHSDTLSKAPGMNSLEQGMVGLKIGDVSSSAVKTVGSV VSSVALTGVLSGNGGTNVNMPVSKPTSWAAIASKPAKPQPKMKTKSGPVMGGGLPPPPIKHNMDIGTWDNKGPVPKAPVPQQAPSP QAAPQPQQVAQPLPAQPPALAQPQYQSPQQPPQTRWVAPRNRNAAFGQSGGAGSDSNSPGNVQPNSAPSVESHPVLEKLKAAHSYN PKEFEWNLKSGRVFIIKSYSEDDIHRSIKYSIWCSTEHGNKRLDSAFRCMSSKGPVYLLFSVNGSGHFCGVAEMKSPVDYGTSAGVW SQDKWKGKFDVQWIFVKDVPNNQLRHIRLENNDNKPVTNSRDTQEVPLEKAKQVLKIISSYKHTTSIFDDFAHYEKRQEEEEVVRK ERQSRNKQ
MGSSHHHHHHSSGLVPRGSHMSASSLLEQRPKGQGNKVQNGSVHQKDGLNDDDFEPYLSPQARPNNAYTAMSDSYLPSYYSPSIGF SYSLGEAAWSTGGDTAMPYLTSYGQLSNGEPHFLPDAMFGQPGALGSTPFLGQHGFNFFPSGIDFSAWGNNSSQGQSTQSSGYSSNY AYAPSSLGGAMIDGQSAFANETLNKAPGMNTIDQGMAALKLGSTEVASNVPKVVGSAVGSGSITSNIVASNSLPPATIAPPKPASWADI ASKPAKQQPKLKTKNGIAGSSLPPPPIKHNMDIGTWDNKGPVAKAPSQALVQNIGQPTQGSPQPVGQQANNSPPVAQASVGQQTQPL PPPPPQPAQLSVQQQAAQPTRWVAPRNRGSGFGHNGVDGNGVGQSQAGSGSTPSEPHPVLEKLRSINNYNPKDFDWNLKHGRVFIIK

SYSEDDIHRSIKYNIWCSTEHGNKRLDAAYRSMNGKGPVYLLFSVNGSGHFCGVAEMKSAVDYNTCAGVWSQDKWKGRFDVRWIF
VKDVPNSQLRHIRLENNENKPVTNSRDTQEVPLEKAKQVLKIIASYKHTTSIFDDFSHYEKRQEEEESVKKERQGRGK

MGSSHHHHHHSSGLVPRGSHMASMSATSVDQRPKGQGNKVSVQNGSIHQKDAVNDDDFEPYLSSQTNQSNSYPPMSDPYMPSYYA
PSIGFPYSLGEAAWSTAGDQPMPYLTTYGQMSNGEHHYIPDGVFSQPGALGNTPPFLGQHGFNFFPGNADFSTWGTSGSQGQSTQSS
AYSSSYGYPPSSLGRAITDGQAGFGNDTLSKVPGISSIEQGMTGLKIGGDLTAAVTKTVGTALSSSGMTSIATNSVPPVSSAAPKPTSW
AAIARKPAKPQPKLKPKGNVGIGGSAVPPPPIKHNMNIGTWDEKGSVVKAPPTQPVLPPQTIIQQPQPLIQPPPLVQSQLPQQQPQPPQP
QQQQGPQPQAQPHQVQPQQQQLQNRWVAPRNRGAGFNQNNGAGSENFGLGVVPVSASPSSVEVHPVLEKLKAINNYNPKDFDWN
LKNGRVFIIKSYSEDDIHRSIKYSIWCSTEHGNKRLDAAYRSLNGKGPLYLLFSVNGSGHFCGVAEMKSVVDYNAYAGVWSQDKWK
GKFEVKWIFVKDVPNNQLRHIRLENNDNKPVTNSRDTQEVPLEKAKQVLKIIATFKHTTSIFDDFAHYEKRQEEEEAMRRERNRNKQ

3.3.6 DF2 domain exhibits ‘Type 2’ mode of action from m^{6}A RNA binding

The binding of m^{6}A RNA potentially induces large structural rearrangements in the DF2 domain. As discussed earlier, *Mmi* is a yeast protein that contains the YTH domain and contains a low complexity domain (which is equivalent to the DF2 domain)[49]. Stowell et al. observed increased stabilization of this domain from m^{6}A RNA binding[49]. Based on this, we initially expected reduced deuterium uptake for the DF2 domain in the bound state. Our difference plots are included and mapped on the crystal structures for the full-length YTHDF2 protein in **Figure 3.7**. We represented our conventional HDX data at 2 σ (we usually employ 3 σ for conventional timepoints, with the exception here). Within these parameters, statistical significance was detected at the 15-second timepoints, and only this timepoint is mapped on the crystal structure (**Figure 3.7**). Peptides spanning the entire DF2 presented an accelerated H/D exchange profile, which did not match our expectations. We rationalized that the DF2 domain likely demonstrates ‘Type 2’ mode of action, which is described by Konermann as a scenario where accelerated H/D exchange is due to large structural rearrangement events that may have taken place[3]. Therefore, the observation of accelerated may not necessarily mean that the DF2 has become even more disordered. Here, we employed a 1:4 protein-to-ligand ratio and conventional timepoints to

elucidate the DF2 domain; in the future, higher ratios or the selection of millisecond timepoints can enhance the magnitude of detectable signals.

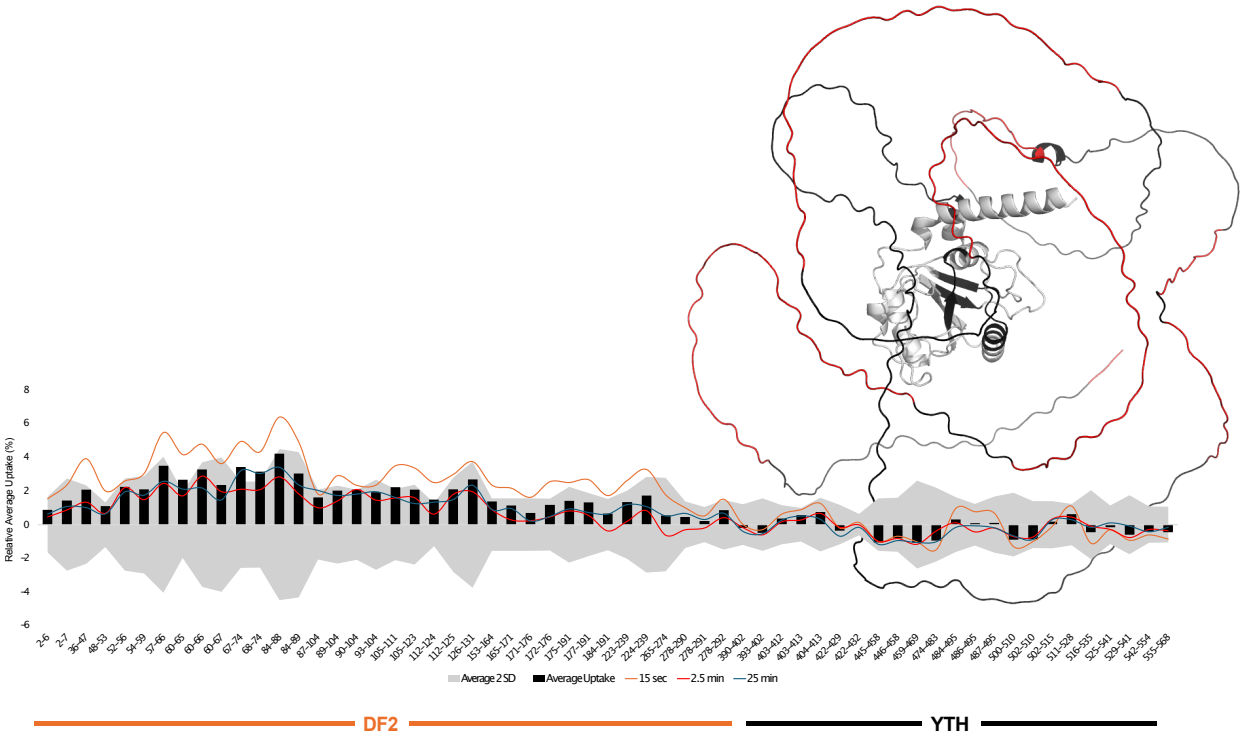


Figure 3.7 | Disordered DF domain from m^{6}A RNA binding
The bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the m^{6}A RNA-bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints as indicated under each plot. The gray shaded area indicates the 2 σ error. Only the 15-second time point is mapped on the YTHDF2 AlphaFold structure[50]. Red color indicates regions of statistically significant accelerated H/D exchange.

3.3.7 DF2 domain is not impacted by Ebselen inhibitor

Ebselen induced significant disorder in the YTH domain; however, no statistical significance was seen in the DF domain (**Figure 3.8**). We have studied Ebselen with the truncated YTH domain (**Figure 3.2**), as well as the full-length YTHDF protein (**Figure 3.8**). In either case, we observed comparable conformational dynamics in the YTH domain. For example, the most accelerated H/D exchange was consistently observed in the α3 region of YTH. In some cases,

truncations of protein domains can potentially impact the conformational dynamics; however, the reproducibility of our data between the YTH and full-length YTHDF protein offers confidence that truncations to this protein have not impacted their conformational dynamics. For the DF domain, no statistically significant observations were detected, and it appears that Ebselen-induced disorder was spatially restricted to only the YTH domain within our selected timepoints (**Figure 3.8)**

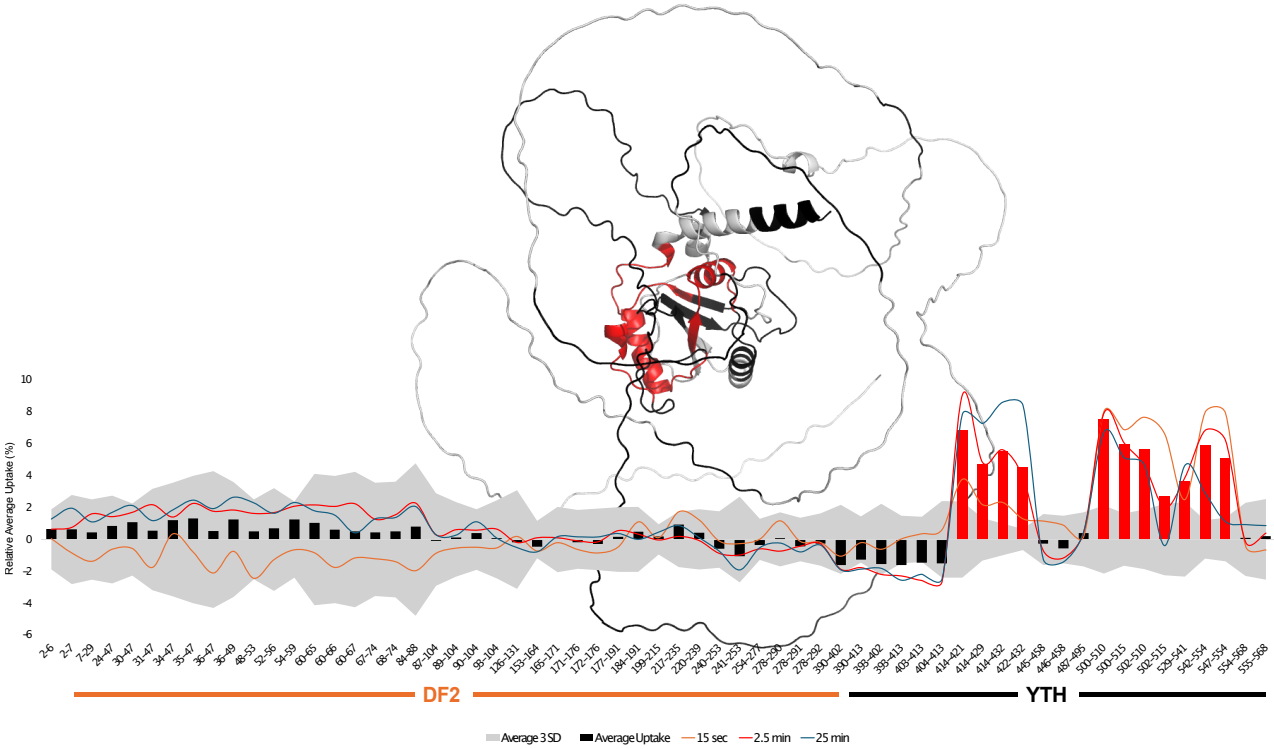


Figure 3.8 | Ebselen binding YTHDF2 for disordered YTH domain
The bars indicate relative average deuterium uptake represented in percentage and calculated by subtracting the relative deuterium uptake in the Ebselen-bound state from the ligand-free state, and then the differences of all time points were averaged. Individual timepoint traces are also shown for selected timepoints as indicated under each plot. The gray shaded area indicates the 3 σ error for conventional timepoints. Red color indicates regions of statistically significant accelerated H/D exchange on YTHDF2 AlphaFold Structure[50].

3.4 Conclusions

We studied the inhibitory mechanisms of Ebselen, Tegaserod, and Salvianolic Acid C against YTH, and aimed to reveal the intricacies of the DF domain using HDX-MS. We probed the recent YTH inhibitors: Ebselen (Micaelli et al. 2022)[45], Tegaserod (Hong et al. 2023)[47], and SAC (Zou et al. 2023)[48]. Ebselen increased collisional cross-section and promoted disorder in the YTH domain, with most pronounced differences in the α3 helix. Our observations aligned with reports from cocrystals and previous studies[45]. For Tegaserod, we conducted millisecond HDX, which is ideal for characterizing weaker interactions. However, we did not detect statistically significant findings for this compound. In our preliminary HDX data, we observed that SAC acts via allosteric stabilization of the α2 and α4 regions in YTH-1, and distortions in the β4–β5 loop of YTH-2; reproducing this data can eventually help address current gaps in the action and selectivity of SAC. Elucidating the dynamics of the DF domain are also thought to facilitate the progress of selective inhibitors. The binding of m^6A RNA potentially induces large structural rearrangements in the DF2 domain; however, the DF2 domain did not present statistically significant observations from Ebselen binding. All results presented in this chapter are preliminary and would require extensive reproduction as well as various experimental optimizations as a future step, such as maximizing the bound state complex by fine-tuning the ligand-protein ratios and strategic selection of timepoints to enhance the magnitude of detectable differences

3.5 Methods

3.5.1 General

We purchased the plasmids for the truncated YTH domains, YTH-1, YTH-2, and YTH-3 from Addgene. The protein sequences for the three YTH domains can be found in **Table 1.1** (see Chapter 2). The expression and purification of YTH-1, YTH-2, and YTH-3 are included (see Chapter 2). We purchased the plasmids for the full-length GST-YTHDF1, His-YTHDF2, and His-YTHDF3 domains from Addgene. The expression and purification of the full-length YTHDF1 presented various challenges (methods not included). The expression and purification of His-YTHDF2 and His-YTHDF3 were similar. Each YTHDF protein was expressed in BL21 cells at 18°C for 16-20 hours. Ni^{2+} affinity chromatography was performed to purify the His-YTHDF proteins over a 15– 500 mM imidazole gradient. All proteins were stored in the following buffer: 20 mM Tris, 0.4 M NaCl, 300 mM–500 mM imidazole, 5% glycerol, pH 7.5, and the samples were stored at -80°C. The samples were stored at -80°C. Proteins were characterized through SDS-PAGE.

3.5.2 m^6A RNA & Inhibitors

The m^6A RNA sequence is as follows: GGCCG(m6A)TCTGA; the sequence identity was retrieved from the Arguello et al. study[25]. This sequence was purchased from horizon (Dharmacon Custom RNA Synthesis). The RNA ligand was suspended in RNase-free water and stored at -80°C. The inhibitors, Ebselen, Tegaserod, and SAC were purchased from Cedarlane. Each compound was purchased as a 10 mM stock prepared in 1 ml of dimethyl sulfoxide (DMSO). Upon receiving, aliquots were prepared (to avoid too many freeze/thaw cycles) and stored at -80°C.

3.5.3 Native MS and IMS

Native MS and IMS data were acquired for YTH2 with Ebselen in this Chapter. The YTH-2 domain was buffer-exchanged into 100 mM ammonium acetate (A637) using the Millipore Amicon Ultra centrifugal filter with a 10 kDa molecular weight (MW) cut-off. A 1:10 YTH2-Ebselen ratio was used. In the bound state, 10 μM YTH-2 was incubated with 100 μM Ebselen for 30 minutes before data acquisition. Experimental conditions were kept consistent for the unbound state; except that, instead of Ebselen, the same volume of DMSO solvent was used a negative control. The sample was flowed through the Hamilton 500 ml HPLC syringe using the Harvard Pump11 Elite Syringe Pump at 20 μl/min flow rates, and to conduct electrospray ionization, a nano spray source was set up on the Waters G2-S Synapt instrument.

3.5.4 Conventional HDX Protocol

Conventional HDX data was collected for: (1) YTH1 with Ebselen, (2) YTH1 versus YTH2 with SAC, (3) YTHDF2 with m^6A RNA, and (4) YTHDF2 with Ebselen.

First, for Ebselen (covalent inhibitor), the samples were incubated for ~6 hours to ensure maximal binding. A 1:4 ratio of YTH1:Ebselen was used in the bound state (and same amount of 2% DMSO solvent was added as a negative control in the unbound state). 90% deuterium labelling collected for the following timepoints: 15 seconds, 2.5 minutes, 25 minutes.

Second, for YTH1 with SAC (K_D in the range of 5.3 - 6.3 μM), a 1:10 ratio was used in the bound state. And in the unbound state, the same volume of 1% final DMSO was used as a negative control. 90% deuterium labelling was collected for the following timepoints: 15 seconds, 2.5

minutes, 10 minutes, 25 minutes. For YTH2 with SAC (IC_{50} ~30 μM), a 1:10 ratio was used in the bound state. In the unbound state, the same volume of 2% final DMSO as the negative control. 50% deuterium labelling was used instead of 90%; due to the weaker interaction, less D_2O dilution was deemed suitable. The following timepoints were collected: 15 seconds, 0.5 minutes, 5 minutes.

Finally, for YTHDF2 with m^6A RNA (3) and Ebselen (4), a 1:4 ratio was used in the bound state. In the unbound state, the same volume of RNAse-free water and DMSO was used, respectively. Labelling was conducted with 90% deuterium. Dynamics from m^6A RNA and Ebselen were elucidated over the same timepoints: 15 seconds, 2.5 minutes, 25 minutes.

Each of the above HDX reactions was quenched in 100 mM phosphate buffer (pH 2.5) for 2 minutes. For protein digestion, pepsin (porcine gastric mucosa, Sigma-Aldrich) was used, and peptides were eluted through the C18 column over 5-35% acetonitrile gradient. Select Series Cyclic IMS (Waters) instrument was used. Peptides were ionized with electrospray ionization. ProteinLynx Global Server (PLGS) was used for peptide identification, and Dynamx for peptide analysis.

3.5.5 Millisecond HDX Protocol

Millisecond HDX was conducted for: (1) YTH1 with Ebselen (2) YTH1 with Tegaserod. For each of these experiments, a 1:2 ratio protein-to-ligand ratio was used in the bound state; and in the unbound state, the volume of DMSO was used as a negative control (2% final DMSO content). The apparatus follows the methods as described by Dr. Joseph Anacleto (see Chapter 1)[21]. For millisecond time points, online deuterium labelling is performed by adjusting the length of the protein line (inner capillary) relative to the deuterium line (outer capillary). The outer

capillary was pulled to 15 mm, 25 mm, and up to 80 mm. For quench, 3% formic acid (FA) was used. For protein digestion, pepsin (porcine gastric mucosa, Sigma-Aldrich) was linked to NHS-activated beads (Pierce, Sigma-Aldrich) in-house and was packed in columns in-house. ProteinLynx Global Server (PLGS) was used for peptide identification, and Dynamx for peptide analysis.

References

(1) Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem. Soc. Rev.* **2011**, *40* (3), 1224–1234. https://doi.org/10.1039/C0CS00113A.

(2) Konermann, L.; Scrosati, P. M. Hydrogen/Deuterium Exchange Mass Spectrometry: Fundamentals, Limitations, and Opportunities. *Mol. Cell. Proteomics* **2024**, *23* (11). https://doi.org/10.1016/j.mcpro.2024.100853.

(3) Konermann, L.; Rodriguez, A. D.; Sowole, M. A. Type 1 and Type 2 Scenarios in Hydrogen Exchange Mass Spectrometry Studies on Protein–Ligand Complexes. *Analyst* **2014**, *139* (23), 6078–6087. https://doi.org/10.1039/C4AN01307G.

(4) Stofella, M.; Seetaloo, N.; St John, A. N.; Paci, E.; Phillips, J. J.; Sobott, F. Recalibrating Protection Factors Using Millisecond Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2025**, *97* (5), 2648–2657. https://doi.org/10.1021/acs.analchem.4c03631.

(5) McLafferty, F. W. A Century of Progress in Molecular Mass Spectrometry. *Annu. Rev. Anal. Chem.* **2011**, *4* (1), 1–22. https://doi.org/10.1146/annurev-anchem-061010-114018.

(6) Wilson, D. Principles and Applications of Mass Spectrometry, 2024.

(7) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9. https://doi.org/10.1021/ac302789c.

(8) Yamashita, M.; Fenn, J. B. Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.* **1984**, *88* (20), 4451–4459. https://doi.org/10.1021/j150664a002.

(9) Dodds, J. N.; Baker, E. S. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2185–2195. https://doi.org/10.1007/s13361-019-02288-2.

(10) Zhang, Z.; Smith, D. L. Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci.* **1993**, *2* (4), 522–531.

(11) Weis, D. D. Hydrogen Exchange Mass Spectrometry of Proteins.

(12) Nguyen, D.; Mayne, L.; Phillips, M. C.; Walter Englander, S. Reference Parameters for Protein Hydrogen Exchange Rates. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (9), 1936–1939. https://doi.org/10.1007/s13361-018-2021-z.

(13) Bai, B. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins Struct. Funct. Bioinforma.* **1993**, *17.1*, 75–86.

(14) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998**, *120* (13), 3265–3266. https://doi.org/10.1021/ja973478k.

(15) Anacleto, J. Personal Communication - Fragmentation in ECD and Data Analysis Using Zeno TOF 7600 System, 2024.

(16) Potriquet, J.; Pribil, P.; Winter, D. Improving Complex Phosphopeptide Characterization with Hybrid EAD/CID MS/MS Fragmentation. *SCIEX*.

(17) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Anal. Chem.* **2000**, *72* (3), 563–573. https://doi.org/10.1021/ac990811p.

(18) Roder, H.; Elove, G. A.; Englander, S. W. Structural Characterization of Folding Intermediates in Cytochrome c by Ff-Exchange Labelling and Proton NMR.

(19) 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* **2003**, *42* (19), 5896–5905. https://doi.org/10.1021/bi034285e.

(20) 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. https://doi.org/10.1021/ac300365u.

(21) Anacleto, J.; Lento, C.; Sarpe, V.; Maqsood, A.; Mehrazma, B.; Schriemer, D.; Wilson, D. J. Apparatus for Automated Continuous Hydrogen Deuterium Exchange Mass Spectrometry Measurements from Milliseconds to Hours. *Anal. Chem.* **2023**, *95* (9), 4421–4428. https://doi.org/10.1021/acs.analchem.2c05003.

(22) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16* (7), 595–602. https://doi.org/10.1038/s41592-019-0459-y.

(23) Luo, J.; Liu, H.; Luan, S.; He, C.; Li, Z. Aberrant Regulation of mRNA m6A Modification in Cancer Development. *Int. J. Mol. Sci.* **2018**, *19* (9), 2515. https://doi.org/10.3390/ijms19092515.

(24) Adhikari, S.; Xiao, W.; Zhao, Y.-L.; Yang, Y.-G. M^6 A: Signaling for mRNA Splicing. *RNA Biol.* **2016**, *13* (9), 756–759. https://doi.org/10.1080/15476286.2016.1201628.

(25) Arguello, A. E.; Leach, R. W.; Kleiner, R. E. In Vitro Selection with a Site-Specifically Modified RNA Library Reveals the Binding Preferences of N6-Methyladenosine Reader Proteins. *Biochemistry* **2019**, *58* (31), 3386–3395. https://doi.org/10.1021/acs.biochem.9b00485.

(26) Wang, X.; Zhao, B. S.; Roundtree, I. A.; Lu, Z.; Han, D.; Ma, H.; Weng, X.; Chen, K.; Shi, H.; He, C. *N*6-Methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **2015**, *161* (6), 1388–1399. https://doi.org/10.1016/j.cell.2015.05.014.

(27) Chen, Z.; Zhong, X.; Xia, M.; Zhong, J. The Roles and Mechanisms of the m6A Reader Protein YTHDF1 in Tumor Biology and Human Diseases. *Mol. Ther. - Nucleic Acids* **2021**, *26*, 1270–1279. https://doi.org/10.1016/j.omtn.2021.10.023.

(28) Wang, X.; Lu, Z.; Gomez, A.; Hon, G. C.; Yue, Y.; Han, D.; Fu, Y.; Parisien, M.; Dai, Q.; Jia, G.; Ren, B.; Pan, T.; He, C. N6-Methyladenosine-Dependent Regulation of Messenger RNA Stability. *Nature* **2014**, *505* (7481), 117–120. https://doi.org/10.1038/nature12730.

(29) Shi, H.; Wang, X.; Lu, Z.; Zhao, B. S.; Ma, H.; Hsu, P. J.; Liu, C.; He, C. YTHDF3 Facilitates Translation and Decay of N6-Methyladenosine-Modified RNA. *Cell Res.* **2017**, *27* (3), 315–328. https://doi.org/10.1038/cr.2017.15.

(30) Wiedmer, L.; Eberle, S. A.; Bedi, R. K.; Śledź, P.; Caflisch, A. A Reader-Based Assay for m6A Writers and Erasers. *Anal. Chem.* **2019**, *91* (4), 3078–3084. https://doi.org/10.1021/acs.analchem.8b05500.

(31) Xu, C.; Liu, K.; Ahmed, H.; Loppnau, P.; Schapira, M.; Min, J. Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *J. Biol. Chem.* **2015**, *290* (41), 24902–24913. https://doi.org/10.1074/jbc.M115.680389.

(32) Patil, D. P.; Pickering, B. F.; Jaffrey, S. R. Reading m6A in the Transcriptome: m6A-Binding Proteins. *Trends Cell Biol.* **2018**, *28* (2), 113–127. https://doi.org/10.1016/j.tcb.2017.10.001.

(33) Li, Y.; Bedi, R. K.; Moroz-Omori, E. V.; Caflisch, A. Structural and Dynamic Insights into Redundant Function of YTHDF Proteins. *J. Chem. Inf. Model.* **2020**, *60* (12), 5932–5935. https://doi.org/10.1021/acs.jcim.0c01029.

(34) Cazzanelli, G.; Dalle Vedove, A.; Spagnolli, G.; Terruzzi, L.; Colasurdo, E.; Boldrini, A.; Patsilinakos, A.; Sturlese, M.; Grottesi, A.; Biasini, E.; Provenzani, A.; Quattrone, A.; Lolli, G. Pliability in the m^6 A-Binding Region Extends Druggability of YTH Domains. *J. Chem. Inf. Model.* **2024**, *64* (5), 1682–1690. https://doi.org/10.1021/acs.jcim.4c00051.

(35) Zaccara, S.; Jaffrey, S. R. A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. *Cell* **2020**, *181* (7), 1582-1595.e18. https://doi.org/10.1016/j.cell.2020.05.012.

(36) Zou, Z.; Sepich-Poore, C.; Zhou, X.; Wei, J.; He, C. The Mechanism Underlying Redundant Functions of the YTHDF Proteins. *Genome Biol.* **2023**, *24* (1), 17. https://doi.org/10.1186/s13059-023-02862-8.

(37) Zhu, T.; Roundtree, I. A.; Wang, P.; Wang, X.; Wang, L.; Sun, C.; Tian, Y.; Li, J.; He, C.; Xu, Y. Crystal Structure of the YTH Domain of YTHDF2 Reveals Mechanism for Recognition of N6-Methyladenosine. *Cell Res.* **2014**, *24* (12), 1493–1496. https://doi.org/10.1038/cr.2014.152.

(38) Li, F.; Zhao, D.; Wu, J.; Shi, Y. Structure of the YTH Domain of Human YTHDF2 in Complex with an m6A Mononucleotide Reveals an Aromatic Cage for m6A Recognition. *Cell Res.* **2014**, *24* (12), 1490–1492. https://doi.org/10.1038/cr.2014.153.

(39) Narang, D.; Lento, C.; J. Wilson, D. HDX-MS: An Analytical Tool to Capture Protein Motion in Action. *Biomedicines* **2020**, *8* (7), 224. https://doi.org/10.3390/biomedicines8070224.

(40) DynamX HDX Data Analysis Software 3.0.

(41) Wolf, E.; Herasymenko, O.; Kutera, M.; Lento, C.; Arrowsmith, C.; Ackloo, S.; Wilson, D. Quantitative Hydrogen–Deuterium Exchange Mass Spectrometry for Simultaneous Structural Characterization and Affinity Indexing of Single Target Drug Candidate Libraries. *Anal. Chem.* **2024**, *96* (32), 13015–13024. https://doi.org/10.1021/acs.analchem.4c01001.

(42) Sikorski, V.; Selberg, S.; Lalowski, M.; Karelson, M.; Kankuri, E. The Structure and Function of YTHDF Epitranscriptomic m6A Readers. *Trends Pharmacol. Sci.* **2023**, *44* (6), 335–353. https://doi.org/10.1016/j.tips.2023.03.004.

(43) Wang, C.-H.; Zhou, H. Discovery of a New Inhibitor for YTH Domain-Containing m6A RNA Readers. *RSC Chem. Biol. 5* (9), 914–923. https://doi.org/10.1039/d4cb00105b.

(44) Sun, Y.; Dong, D.; Xia, Y.; Hao, L.; Wang, W.; Zhao, C. YTHDF1 Promotes Breast Cancer Cell Growth, DNA Damage Repair and Chemoresistance. *Cell Death Dis.* **2022**, *13* (3), 1–11. https://doi.org/10.1038/s41419-022-04672-5.

(45) Micaelli, M.; Dalle Vedove, A.; Cerofolini, L.; Vigna, J.; Sighel, D.; Zaccara, S.; Bonomo, I.; Poulentzas, G.; Rosatti, E. F.; Cazzanelli, G.; Alunno, L.; Belli, R.; Peroni, D.; Dassi, E.;

Murakami, S.; Jaffrey, S. R.; Fragai, M.; Mancini, I.; Lolli, G.; Quattrone, A.; Provenzani, A. Small-Molecule Ebselen Binds to YTHDF Proteins Interfering with the Recognition of N6-Methyladenosine-Modified RNAs. *ACS Pharmacol. Transl. Sci.* **2022**, *5* (10), 872–891. https://doi.org/10.1021/acsptsci.2c00008.

(46) Haritha, C. V.; Sharun, K.; Jose, B. Ebselen, a New Candidate Therapeutic against SARS-CoV-2. *Int. J. Surg.* **2020**, *84*, 53–56. https://doi.org/10.1016/j.ijsu.2020.10.018.

(47) Hong, Y.-G.; Yang, Z.; Chen, Y.; Liu, T.; Zheng, Y.; Zhou, C.; Wu, G.-C.; Chen, Y.; Xia, J.; Wen, R.; Liu, W.; Zhao, Y.; Chen, J.; Gao, X.; Chen, Z. The RNA m6A Reader YTHDF1 Is Required for Acute Myeloid Leukemia Progression. *Cancer Res.* **2023**, *83* (6), 845–860. https://doi.org/10.1158/0008-5472.CAN-21-4249.

(48) Zou, Z.; Wei, J.; Chen, Y.; Kang, Y.; Shi, H.; Yang, F.; Shi, Z.; Chen, S.; Zhou, Y.; Sepich-Poore, C.; Zhuang, X.; Zhou, X.; Jiang, H.; Wen, Z.; Jin, P.; Luo, C.; He, C. FMRP Phosphorylation Modulates Neuronal Translation through YTHDF1. *Mol. Cell* **2023**, *83* (23), 4304-4317.e8. https://doi.org/10.1016/j.molcel.2023.10.028.

(49) Stowell, J. A. W.; Wagstaff, J. L.; Hill, C. H.; Yu, M.; McLaughlin, S. H.; Freund, S. M. V.; Passmore, L. A. A Low-Complexity Region in the YTH Domain Protein Mmi1 Enhances RNA Binding. *J. Biol. Chem.* **2018**, *293* (24), 9210–9222. https://doi.org/10.1074/jbc.RA118.002291.

(50) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.