THE UTILITY OF HYDROGEN DEUTERIUM EXCHANGE MASS SPECTROMETRY FOR CHARACTERIZATION OF PROTEINS AND THEIR INTERACTIONS.

ESTHER ZVIA WOLF

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

GRADUATE PROGRAM IN CHEMISTRY

YORK UNIVERSITY

TORONTO, ONTARIO

May 2024

ABSTRACT

The Mass Spectrometer (MS), a versatile bioanalytical instrument with origins dating back to discoveries of isotopes and electrons, has seen increasing application in structural biology. Particularly, the coupling of Hydrogen Deuterium Exchange (HDX) to MS has become an attractive user-friendly tool for characterizing the structure and interactions of proteins, particularly those that are dilute and/or dynamic. Although HDX-MS is not explicitly atomic resolution like gold standard methods such as X-Ray Crystallography, it is readily applied in comparative studies to extract the details of binding sites of reasonable affinity ligands. Through the advent of robotic tool change coupled autosamplers, the speed of information turnaround time has only grown more rapid. Additionally, ease of access to high confidence *in silico* structure prediction has made HDX-MS easier to visualize for more difficult protein systems, making the interpretation of data far more friendly to the extended community. HDX-MS has a future in fast paced environments, such as the early stages of drug discovery, where higher throughput is a necessity. Although the awareness, support, and application of HDX-MS grows, the instrumentation can be cost prohibitive. It is therefore very exciting to experience how many interesting, relevant scientific questions can be resolved through collaboration. This work follows many different scenarios where HDX-MS and MS can be employed, including protein-peptide interactions, affinity ranking and prediction of a focused molecule library, small molecule drug binding characterization, the detection of bifunctional molecule-induced ternary complex formation, and epitope mapping, made possible by the collaboration of the Dickinson Lab (University of Chicago), Icosagen (Estonia), the Ingerman-James Lab (University of North Carolina at Chapel Hill), and the Structural Genomics Consortium.

ACKNOWLEDGEMENTS

First, I would like to express my gratitude toward Derek Wilson, who, dating back to 2018, gave me the life changing opportunity to learn about mass spectrometry and its many applications. Also, thank you to Gerald Audette and Ryan Hili for being insightful, compassionate course instructors, who, as my committee members, have consistently encouraged thought-provoking discussion during research evaluations. Likewise, a big thank you to Lars Konermann and Mark Bayfield for serving as my external and at-arms-length examiners.

Thank you to my mentors over the years: Cristina Lento, Nicholas Bragagnolo, Irina Oganesyan, Xiaojing Huang, Lisa Szymkowicz, Kerene Brown, John van Nostrand, Banafsheh Mehrazma, and Dominic Narang. Thank you to all my lab mates, Vimanda, Ayesha, Venita, Joe, and Alex, for their constructive scientific discussions. You are all so bright and hard working, and I am looking forward to seeing what your next steps will be.

Thank you to professors Kyle Belozerov, Derek Jackson, Pierre Potvin, Arturo Orellana, Nicole Nivillac, Cora Young, and Demian Ifa for being inspiring, supportive teachers. Thank you to Sofia Lendor, Dave Wolfenden, and the folks at Waters for having our backs.

I want to emphasize my deep appreciation for Suzanne Ackloo for uniting our lab with Cheryl Arrowsmith, Rachel Harding, and the incredible folks at the SGC. Thank you all for giving me a chance and completely changing my academic trajectory.

I do not think words can express my gratitude toward Cristina Lento. A pillar of strength who improves the life of everyone she meets. Her unwavering support, brilliance, generosity, and curiosity make up the foundation of the lab. Her friendship is the greatest gift of this journey.

Lazar, my rock, thank you for being an unbelievably remarkable teammate. Thank you for giving me hope. Thank you for creating our little bubble.

Finally, thank you to my family for their support despite the Ottawa visits spent doing data analysis at the kitchen table, the missed messages, all the funny sounding words, the endless tasks to be stressed over, and until recently, no ideas about what the future would bring.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BCL-2 – B Cell Lymphoma 2 Protein

BH – Bcl-2 homology domain

BLI – Biolayer Interferometry

CCS – Collision Cross Section

CIU – Collision Induced Unfolding

COVID-19 – Coronavirus disease 2019

CRDMO – Contract Research, Development, Manufacturing Organization

CRM – Charged Residue Model

Cryo-EM – Cryogenic Electron Microscopy

CTD – Carboxy-Terminal Domain

DUB - Deubiquitylases

DUBTAC – Deubiquitylase-Targeting Chimeras

ESI – ElectroSpray Ionization

FBXO22 – F-box protein 22

FWHM – Full-Width at Half Maximum

H3K36Me2 – Histone 3 Lysine-36 Dimethylation

hACE2 – Human Angiotensin-Converting Enzyme 2

HDX – Hydrogen Deuterium Exchange

IMS – Ion Mobility Separation

IPTG – Isopropyl B-D-1-Thiogalactopyranoside

K_D – Dissociation Coefficient

LB – Luria Bertani

m/z – mass-to-charge

mAb – Monoclonal Antibody

MCL-1 – Induced Myeloid Leukemia Cell Differentiation Protein

MLL1 – Histone-lysine N-methyltransferase 2A

MS – Mass Spectrometry

NMR – Nuclear magnetic resonance spectroscopy

NSD2 – Nuclear SET domain-containing protein 2

NTD – Amino-Terminal Domain

PDB – Protein data bank

PLGS - ProteinLynx Global Server

PMMA – Polymethylmethacrylate PROTAC – Proteolysis Targeting Chimeras

PWWP1 – Pro-Trp-Trp-Pro domain 1

RbBP5 – Retinoblastoma-binding Protein 5

RBD – Receptor Binding Domain

RMSD – Root Mean Square Deviation

S/N – Signal-to-Noise

SAR – Structure-Activity Relationship

SARS-CoV-2 – Severe Acute Respiratory Syndrome Related Coronavirus 2

SDS-PAGE – Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

SEC – Size Exclusion Chromatography

SPR – Surface Plasmon Resonance

TCEP – Tris(2-Carboxyethyl)Phosphine

TRESI – Time-Resolved ElectroSpray Ionization

UBD – Ubiquitin Binding Domain

UBR – Ubiquitin Recognition Boxes

UPLC – Ultra Performance Liquid Chromatography

USP – Ubiquitin-Specific Peptidase

VoC – Variant of Concern

WBM – WDR5 Binding Motif

WDR – Trp-Asp 40 (WD40) Repeat

WIN – WDR5-Interacting Motif

XL – Cross-linking

ZnF – Zinc Finger

LIST OF PUBLICATIONS

1. **Wolf E**, Herasymenko O, Kutera M, Arrowsmith CH, Ackloo S, Wilson D. Quantitative HDX-MS for Structural Characterization and Affinity Indexing of WDR5-Targeting Ligand Library. *Submitted for Publication.*
2. Nie DY, Tabor JR, Li J, Kutera M, St-Germain J, Hanley RP, **Wolf E**, Paulakonis E, Kenney TMG, Duan S, Shrestha S, Owens DDG, Pon A, Szewczyk M, Lamberto AJ, Menes M, Li F, Barsyte-Lovejoy D, Brown NG, Barsotti AM, Stamford AW, Collins JL, Wilson DJ, Raught B, Licht JD, James LI, Arrowsmith CH. Recruitment of FBXO22 for Targeted Degradation of NSD2. *In Press.*
3. Mann MK, **Wolf E**, Silva M, Kwak HA, Wilson B, Bolotokova A, Wilson DJ, Harding RJ and Schapira M (2024) "Small Molecule Screen Identifies Non-catalytic USP3 Chemical Handle", *ACS Omega , 9*(1), 917–924.
4. Hermet P, ... **Wolf E**, ... Wilson DJ, ... Ustav Jr. M (2023) "Broadly neutralizing humanized SARS-CoV-2 antibody binds to a conserved epitope on Spike and provides antiviral protection through inhalation-based delivery in non-human primates", *PLoS pathogens, 19*(8), e1011532.
5. **Wolf E**, Lento C, Pu J, Dickinson BC and Wilson DJ (2023) Innate Conformational Dynamics Drive Binding Specificity in Anti-Apoptotic Proteins Mcl-1 and Bcl-2", *Biochemistry, , 62*(11), 1619–1630.
6. Chow V, **Wolf E**, Lento C and Wilson DJ (2023) "Developments in rapid hydrogen–deuterium exchange methods", *Essays in biochemistry, 67*(2), 165–174.

CHAPTER 1

Introduction

1.1 Studying Proteins as Therapeutic Targets

Many proteins have unique primary and higher order structures which link to their function and enable their modulation. Accordingly, studying the structure of proteins and their interactions is at the core of understanding disease mechanisms and developing medical treatment.

Genetically linked diseases often manifest as proteins with aberrant expression, activity, and/or sequences. For example, the overexpression of antiapoptotic proteins Bcl-2 and Mcl-1 or epigenetic methyl reader WDR5 is implicated in various cancers.[1,2] Alternatively, our endogenous, normal functioning proteins may be hijacked to spread infectious diseases. A relevant example would be the human ACE2 receptor facilitating endocytosis of the SARS-CoV-2 viral particle.[3] To correct for aberrant proteins, therapeutic agents modify protein structure, interactions, and abundance. Traditional approaches include inhibition using small molecule drugs or antibodies while new, emerging strategies such as Proteolysis Targeting Chimeras (PROTACs)[4] and Deubiquitinase Targeting Chimeras (DUBTACs)[5] aim to alter protein abundance.

These developments are enabled by structural biologist's who study macromolecular structure, dynamics, and function using various biophysical instruments. While X-ray Crystallography and Nuclear Magnetic Resonance Spectroscopy (NMR) are structural elucidation gold standards that offer atomic resolution, scenarios exist where they fall short: proteins or complexes which are massive, dynamic, dilute, and/or weakly populated, etc.[6] Thus, incorporation of Mass Spectrometry (MS) into the analytical toolbox can help overcome structural biology obstacles through orthogonal means. Moreover, as populations grow and insights are made into new and old diseases, higher throughput is increasingly necessary to meet research and development demands. For example, MS enables rapid, sensitive, and accurate sample characterization, all qualities of great utility in early/preclinical stages of drug discovery, where the aim is to screen for novel protein targeting agents (i.e., hit finding and identification) and characterize and validate their impact on protein structure and/or the proteome.[7]

1.2 Mass Spectrometry

The analytical technique of Mass Spectrometry (MS) measures gas-phase ion mass-to-charge (*m/z*). MS originated at the hands of physicists such as J. J. Thomson (1906 Nobel Prize

in Physics for electron discovery), E. Everett, and Francis Aston (1922 Nobel Prize in Chemistry for discovery of isotopes) in the early 1900s as they aimed to answer fundamental questions about atomic structure while bridging theoretical and experimental physics.[8] By the late-1950s, mechanisms of fragmentation during MS were understood and employed to elucidate organic molecule structure; fragment ions in mass spectra related to precursor structure and rearrangement.[9] Today, MS is a widely used analytical tool across various scientific disciplines to reveal the identity and composition of pure or highly complex samples. MS is essential in the studies of proteomics, metabolomics, and other -omics fields.[10] Pixel-by-pixel ionization has enabled the imaging of tissues and cells.[11] Additionally, MS has enabled the gas-phase study of masses ranging from elements to viruses.[12]

To begin with some basics, a mass spectrometer is usually composed of three general parts: (i) an ion source, (ii) mass analyzer(s), and (iii) a detector. To travel through the MS, analytes must be ionized (i) into the gas-phase from gas, solution, or solid state. A variety of ionization methods exist and have distinct applications, including gentle, macromolecular structure preserving ElectroSpray Ionization (ESI, see section 1.2.1). Mass analyzers (ii) come in different forms: Quadrupole (Q), Time-of-Flight (TOF), Fourier-Transform Ion Cyclotron Resonance (FT-ICR), etc., and modern-day MS instruments often have hybrid functionality, such as a Q-TOF (see section 1.2.2.). In the most general terms, analyzers are compartments (cells) which apply principles of physics to gaseous ions passing through them, enabling separation and mass-to-charge identification. Each cell operates below atmospheric pressure (i.e., vacuum) requiring the instrument to be fitted with pumps that withdraw ambient air. (iii) At the instrument terminus, a detector converts the ion presence into an electrical signal monitored by a computer. For example, microchannel plates (MCPs) will release secondary electrons upon collision with an impinging analyte ion, causing an amplifying cascade that registers as a signal.[13,14]

The basic forms of data output from mass spectrometer are a mass spectrum or chromatogram. A mass spectrum is a plot of signal intensity as a function of m/z, where the net charge, z, is of either positive or negative polarity and ion mass, m, is the sum of the analyte's neutral mass (molecular weight), M, and z masses (i.e., $M+zH = m$ in positive ion mode, H = 1.00784 amu). For example, an ion with m of 200 amu and z of 2+ will generate a MS peak at 100 m/z, while its M is 198 Da. Most MS instrumentation measures ion ensembles/packets, rather than one at a time, so a mass spectrum will reflect analyte composition, including matrix components or contaminants, and isotopic abundances. On the other hand, a chromatogram plots the cumulative signal intensity per spectrum over time (Total Ion Chromatogram, TIC), or the base

peak intensity over time (Base Peak Intensity, BPI). The base peak is the highest intensity peak in a spectrum with multiple ion species.

Resolving power is the capability of the MS instrument to distinguish between similar masses such as isotopes (e.g., C-12 and C-13, H-1 and H-2, etc.). Peak resolution (R) can be determined for a given *m/z* using the peak's Full-Width at Half-Maximum intensity (R_{FWHM}) as shown in **Fig. 1.1a**, although resolution can be more stringently determined using 10 or 5% maximum (as opposed to 50%).[15,16] As R increases for a given peak, the isotopic distribution of the peak becomes more clear (**Fig. 1.1bc**).[16] An isotopically distributed signal has a (i) monoisotopic *m/z*, composed of only the highest abundance isotope of each element, a (ii) average *m/z*, dependent on population weighted isotopic abundance, and in the event of a poorly resolved distribution, (iii) a peak top, determined by the mode average measurement.[16]

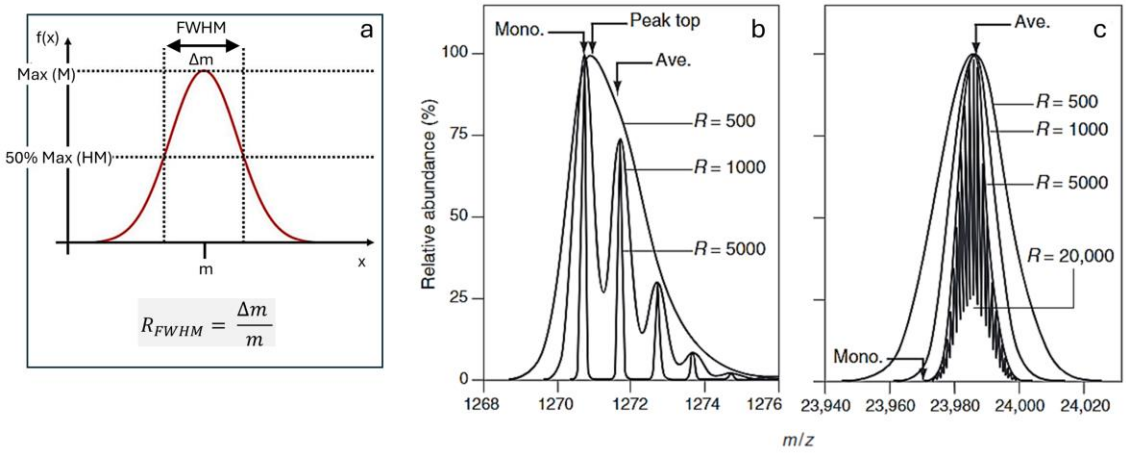


Figure 1.1 | **Mass Spectral Signal Peak Resolution and Isotopic Distribution.** (a) Full-Width at Half-Maximum Resolution. Adapted from Tanaka et al. (2023) with new labels under CC BY license 4.0.[15] (b) $C_{60}H_{86}N_{16}O_{15}$ and (c) $C_{1080}H_{1697}N_{268}O_{325}P_5S_6$ adapted from Carr & Annan (1996) with permission, © 2010 John Wiley & Sons, Inc.[14]

Adjacent isotopic peaks have a difference of ~1 amu because each successive peak is generated by a ionic species with 1 additional heavy isotope ($m_{neutron}$ = 1.0087 amu).[17] In the case of NH_4^+, its monoisotopic mass would be approximately 18 amu, but the next largest isotopic peak mass would be 19 amu if there is a single N-15 or H-2, and so on. If we consider the monoisotopic peak m_1/z and its adjacent isotopic peak m_2/z, the spacing between isotopic peaks $\Delta m/z$ can be used to determine mass:

$$m_2 - m_1 = 1 \quad ; \quad \Delta\frac{m}{z} = \frac{m_2 - m_1}{z} = \frac{1}{z} \quad ; \quad [M + zH]^{z+} = \frac{m}{z} \cdot z \qquad (1.1)$$

Taking the reciprocal of the peak spacing and multiplying by the m/z reveals the *M+zH* and subsequent subtraction of *zH* yields the *M*. When a species is low molecular weight, the mass of the monoisotopic peak is the real mode average of the ensemble (**Fig. 1.1b**) whereas, for larger species, the peak top is the mode and mean average (**Fig. 1.1c**) because the likelihood of containing one or more heavy isotopes increases with size. Moreover, due to the inverse relationship between isotopic peak spacing and net charge, and the tendency for net charge to increase with analyte size, the observation of isotopic distributions in larger analytes requires an even higher resolution (**Fig. 1.1bc**).

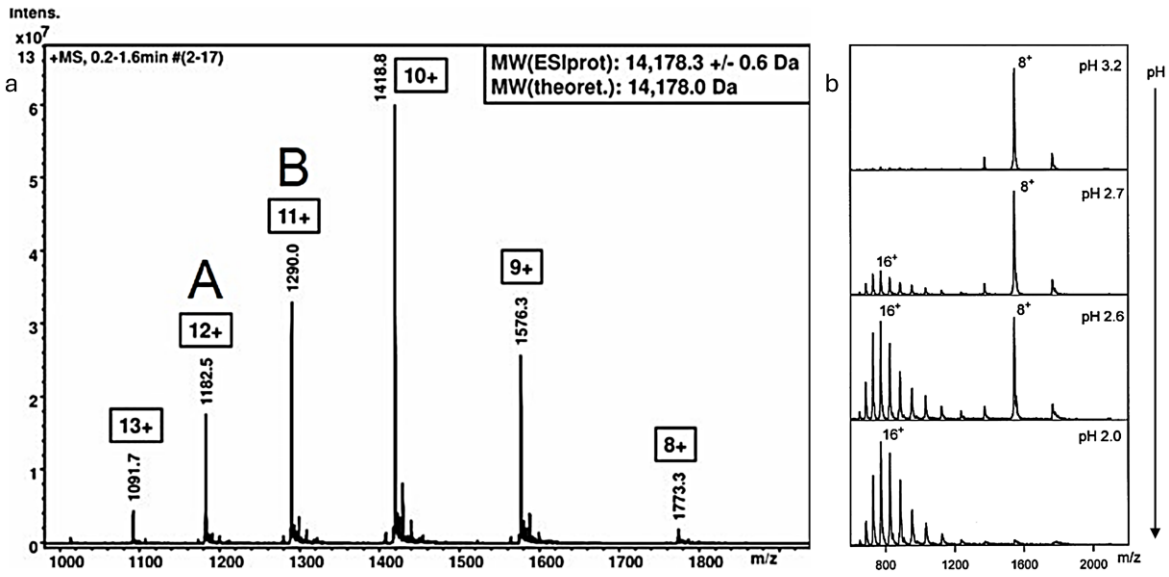


Figure 1.2 | **Charge State Envelopes**. (a) ESI-MS of α-lactalbumin adapted from Winkler (1996) with permission.[18] (b) Unfolding of Cytochrome *c* with decreasing pH adapted with permission from Konermann & Douglas (1998).[19]

Also, as mass and conformational freedom increase, charged states begin to appear (**Fig. 2**) due to the charge deposition process during ElectroSpray Ionization (ESI), discussed further in Section 1.2.1. The spacing between adjacent charge state peaks in an envelope can be derived using a using a simplified case of peaks A and B (**Fig. 1.2a**) where z_A and z_B of $(m/z)_A$ and $(m/z)_B$ are sequential integer values:

$$z_A - 1 = z_B \quad ; \quad \Delta\left(\frac{m}{z}\right)_{AB} = \left(\frac{m}{z}\right)_B - \left(\frac{m}{z}\right)_A = \frac{M+z_B H}{z_B} - \frac{M+z_A H}{z_A} = \frac{M}{(z_A)^2 - z_A} \quad (1.2)$$

This inverse relationship explains why the spacing between peaks increases from left to right as the charge state goes down. Also, this relationship can be used to predict and identify charge state peaks in spectra such as those with many obscuring signals (e.g., protein purification).

Alternatively, a tool called ESIprot, widely used in the field, was developed by Winkler (2010) which enabled the determination of *M* and prediction of charge states from two or more *m/z*'s corresponding to adjacent charge state peaks.[18] ESIprot automated *M* determination from $(m/z)_A$, $(m/z)_B$, … etc. by sampling all integer values of *z* until an average *M* is found with the lowest standard deviation.[18] This is particularly useful when the resolution of a charge state peak is insufficient to extract *z* from the isotopic distribution.

A charge state envelope, or the extent of charge deposition, can describe the extent of folding, and multiple envelopes can indicate conformational sub-populations or oligomers. Accordingly, a very broad charge state envelope would indicate a vast conformational space that is sampled. Proteins exposed to denaturing conditions such as acidic or organic solvents will unfold, pick up more charge, and their entire envelope will shift to the left of the *m/z* axis (**Fig. 1.2b**).[19] Since MS is an ensemble measurement, samples undergoing folding or oligomerization equilibria will produce multiple charge state envelopes.

1.2.1 ElectroSpray Ionization of Proteins

Early stages of MS relied on ionization methods which would cause fragmentation of macromolecules. The 2002 Nobel Prize in Chemistry was awarded to John Fenn and Koichi Tanaka for their respective developments of soft, atmospheric pressure ionization methods, ElectroSpray Ionization (ESI) and Matrix Assisted Laser Desorption Ionization (MALDI). While MALDI involved the irradiation of analyte from a solid matrix, the ESI process, whose origins date back to 1966 when Malcolm Dole was inspired by a spray-painting apparatus used in car manufacturing[20,21], ionizes from the solution-phase and is illustrated in **Fig. 1.3**.[22]

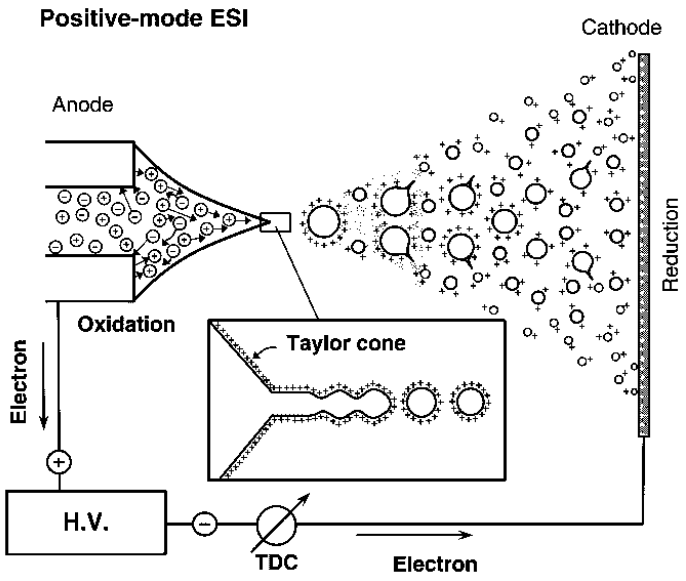


Figure 1.3 | **Positive Mode ElectroSpray Ionization Process.** Reproduced with permission from Hiroaka (2013).[22]

At the start of the ESI process, an electrolytic solution containing the analyte of interest is passed through an ESI emitter at ~1-100s µL/min. The ESI emitter is a sub-millimeter diameter, electrically conducting capillary under a high electric field, particularly at the capillary tip.[22] In the case of positive mode ESI, the voltage applied to the capillary (anode) is positive (2-3 kV), and the MS orifice (cathode) toward which the spray is directed, often assisted by a nebulizing gas (e.g., N_2), has a negative voltage. Thus, as bulk solution flows, anions are pulled toward the capillary wall, where oxidation occurs, while cations are repelled away and out of the emitter, leading to high like-charge density and formation of a Taylor cone.[23] The Taylor cone is punctuated at its apex by a sub-micrometer jet which then breaks up into small, charged droplets called electrospray, which form due to coulombic repulsion overcoming surface tension.[21,22,24] Surface tension at the water-air interface is due to water molecules forming stronger intermolecular bonds with each other than the air (i.e., cohesion > adhesion).[25] As a parent droplet moves toward the cathode, it undergoes desolvation, decreasing droplet volume and increasing charge density. The polarized droplet again forms a cone or protrusion as surface tension is exceeded by repulsion, and progeny droplets emerge.

The net critical charge, Q, of a droplet with radius, r, where repulsion exceeds surface tension has been modeled to be theoretically equal, or exceptionally close, to Q_R in the "Rayleigh Equation" (**Eq. 1.3**).[26]

$$Q_R = 8\pi\sqrt{\varepsilon_0 \gamma {r_R}^3} \quad (1.3)$$

Q_R is defined as a function of the droplets radius, r_R, surface tension, γ, and permittivity of the vacuum, ε_0.[27] Dating back to 1882, Lord Rayleigh explained, *"When Q is great* [relative to surface tension]*, the spherical form is unstable* […]. *Under these circumstances the liquid is thrown out in fine jets…"*.[27] Droplets undergoing this form of jet fission during the ESI process are said to be at or beyond the "Rayleigh Limit".[28]

NanoESI harnesses this relationship by using capillaries with tip diameters of several micrometers to generate even smaller charged droplets (i.e., nano-electrospray) than ESI, which require fewer iterations of jet fission to generate naked gas-phase analytes. It is accordingly considered to be better at preserving weaker interactions and multi-subunit complexes, and more sensitive due to reduced flowrates (nL/min) and low sample volumes.[29] A caveat of nanoESI is that such narrow tips are prone to clogging, and while constant purchase from commercial sources may be expensive, in-house capillary pulling takes a lot of practice to master the skill.[30]

To aid in both analyte desolvation and reduction of the salt adduction that would otherwise convolute the mass, the solutions used for ESI rely on vaporizable/volatile solvents (e.g., methanol, acetonitrile, and water) and salts (e.g., ammonium acetate). Typically, a substance will be considered volatile if it weakly interacts with itself or its solvent. Volatility is also inversely proportional to mass; lower relative mass and low electron density result in weaker London dispersion forces/polarizability and can be quantified using boiling point (b.p.) or vapour pressure.[31,32] In the case of adduction, higher purity solvents and effective buffer exchange during sample preparation generate sharper, narrow m/z peaks due to, for example, the elimination of sodium and chloride ions from prior purification stages.

While the mechanistic steps of droplet evolution during electrospray have been detailed, the final stages of naked gas-phase ion formation are less obvious. To start, where does the charge come from? Analytes can have in-solution 'permanent' charge such as Na^+, while the charge of organic molecules is provided by a proton, H^+, as a function of pKa. The charge of larger molecules, such as proteins, is less intuitive in mass spectra, as they display distributions of charges and/or exceed their net charge predicted by primary sequence. For example, in **Fig. 1.2a**, bovine α-lactalbumin has a charge state distribution of 8+ to 13+, but its theoretical net charge in solution is 7- (i.e., 20 Asp/Glu vs. 13 Arg/Lys).[33] During ESI, excess positive charge in the droplet can come from protons formed from water oxidation at the metal capillary wall (e.g., $2\ H_2O \rightarrow 4\ H^+ + 4\ e^- + O_2$), but then, ultimately, how is additional charge deposited on analytes as they

escape the droplet? Also, notably, despite excess protons present in ESI droplets, folded protein structures are maintained when stock solution conditions reflect biologically relevant pH. This likely indicates that the ESI process is too rapid for acid denaturation of proteins (i.e., droplet lifetime is ~ns-μs)[28], though MS-compatible neutral pH buffers have been described for native structure preservation.[34]

Excess positive charges in an isolated conductor such as droplet will repel and reside on the surface.[28,35] This surface localization is possible for small, low charge ions or even non-polar, hydrophobic species; however, large ions with distributed charge would be insufficiently solvated at the surface. In that case, the ability for a species with abundant like-charges to remain sequestered within a droplet despite forces of repulsion could be rationalized by "projected charges": chains of water-dipoles radiating from the ion out to the droplet surface.[28] With that said, there are three commonly encountered analytes in bioanalytical samples: (i) elemental or small molecular ion, (ii) large globular protein, and (iii) disordered protein. Konermann and colleagues have endeavored to illustrate these scenarios using molecular dynamics (MD) simulations.[28,36–38]

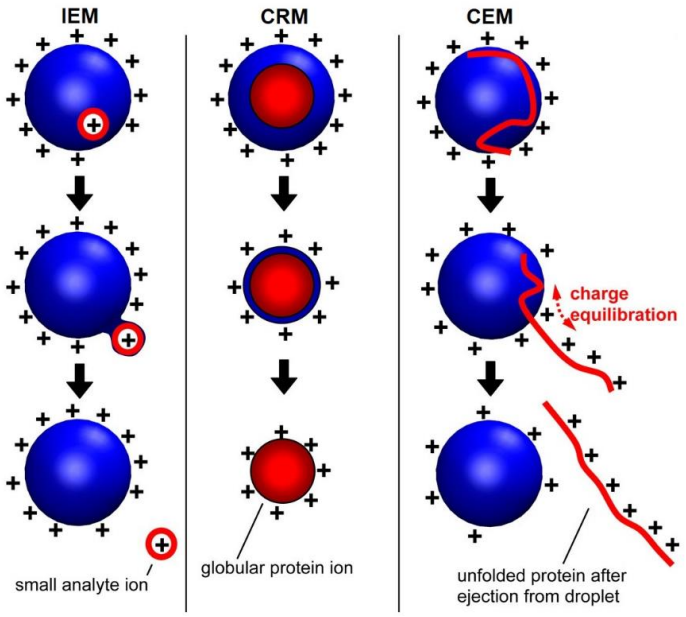


Figure 1.4 | **Gas-Phase Analyte Formation.** Ion Evaporation Model (IEM), Chain Ejection Model (CEM), and Charged Residue Model (CRM). Adapted with permission from Konermann et al. (2013), © 2013 American Chemical Society.[28]

Three models of gas-phase analyte formation are depicted in **Fig. 1.4**: Ion Evaporation Model (IEM), Chain Ejection Model (CEM), and Charged Residue Model (CRM).[28] While IEM best

applies to small chemical moieties (organic molecules, elemental ions, compounds), CRM and CEM are best applied to globular and unstructured/unfolded proteins, respectively.

In IEM, the small ion experiences opposing forces of like-charge repulsion and solvation until the electric field is strong enough to eject the minimally solvated ion, which is morphologically analogous to jet fission observed during the broader ESI process.[36] Thus, typically, the ion charge reflects the solution-phase charge of the analyte.

In CRM, a well solvated globular analyte is at the core of a droplet undergoing successive evaporation events, while small matrix ions undergo IEM, until the remaining surface charge is deposited on the analyte.[28] In this case, the electric field is not strong enough to induce ejection of the relatively larger globular analyte itself. Since droplet decomposition is slower than ion ejection for gaseous ion formation, duration of CRM is longer than IEM (μs vs. ns).[28] Konermann et al. (2013) noted that, consistent with CRM where gaseous ions form from minimally solvated droplets ($r_{protein} \approx r_R$, **Eq. 1.3**), spectra of globular proteins have charge states close to Q_R, and charge state distributions are a function of the droplet size they started from.[28] Also, for proteins sprayed from complex matrices (intentional or not), the CRM model explains why protein ions, despite matching the theoretically expected charge, will often manifest in broad m/z peaks due to varying or higher than expected mass: (i) water clusters may remain on the gas-phase analyte due to slow droplet breakdown, (ii) the analyte picks up the excess charges near the droplet surface, which will reflect the matrix composition (i.e., sodium or other impurities). This can be extended to include non-specific protein-ligand interactions observed in mass spectra induced by excess ligand-to-protein stoichiometries.

Finally, CEM deals with unstructured/unfolded proteins which accordingly have patches of hydrophobic residues that would otherwise be sequestered in a globular protein. This hydrophobic character causes the analyte to reside near the droplet surface where solvation is less efficient.[28] During droplet fission, the polypeptide chain terminus protrudes out of the droplet due to like-charge repulsion. This positions both hydrophobic and hydrophilic patches along the droplet surface, resulting in the chain continually ejecting away from the water and picking up surface charges. Like IEM, CEM is more efficient at producing gas phase ions because the slow process of droplet evaporation is superseded by analyte ejection. This is consistent with experimental observations that unfolded proteins have higher intensity signals in mass spectra than their folded counterparts, indicating more analyte is successfully ionized.[37]

Once the analyte of interest has been ionized and entered the gas-phase, it is in a state that can be analyzed by MS. Adjusting the conditions at the MS front-end can improve ion transmission

by efficient nebulization and desolvation, but may be confounding.[39] Unidirectional, concentric gas flow (e.g., N_2) will help nebulize the Taylor cone while increased counter flow gas could deflect smaller mass ions; however, increased pressure can also increase collisions amongst analyte molecules. The ion source itself can heat up to improve droplet drying but may impact analyte integrity. Positioning the ESI emitter perpendicular to the orifice can reduce transmission of neutral species. The emitter-MS orifice distance can inversely impact the efficiency of desolvation compared to ion loss due to ESI plume; longer distance results in more iterations of fission at the expense of the spray expanding beyond the orifice diameter.[39]

1.2.2. The Synapt G2-Si by Waters: A Hybrid Mass Spectrometer

There are many mass analyzers currently available commercially, but the downstream focus will be driven by those of the mass spectrometer used at the earliest stages of my studies, the Waters Synapt G2-S. The Synapt G2-S is a hybrid instrument made up of two distinct mass analyzers, a quadrupole, and a TOF (**Fig. 5**). It has additional ion activation and size filtration compacities by virtue of pressurized collision and ion mobility separation (IMS) cells.

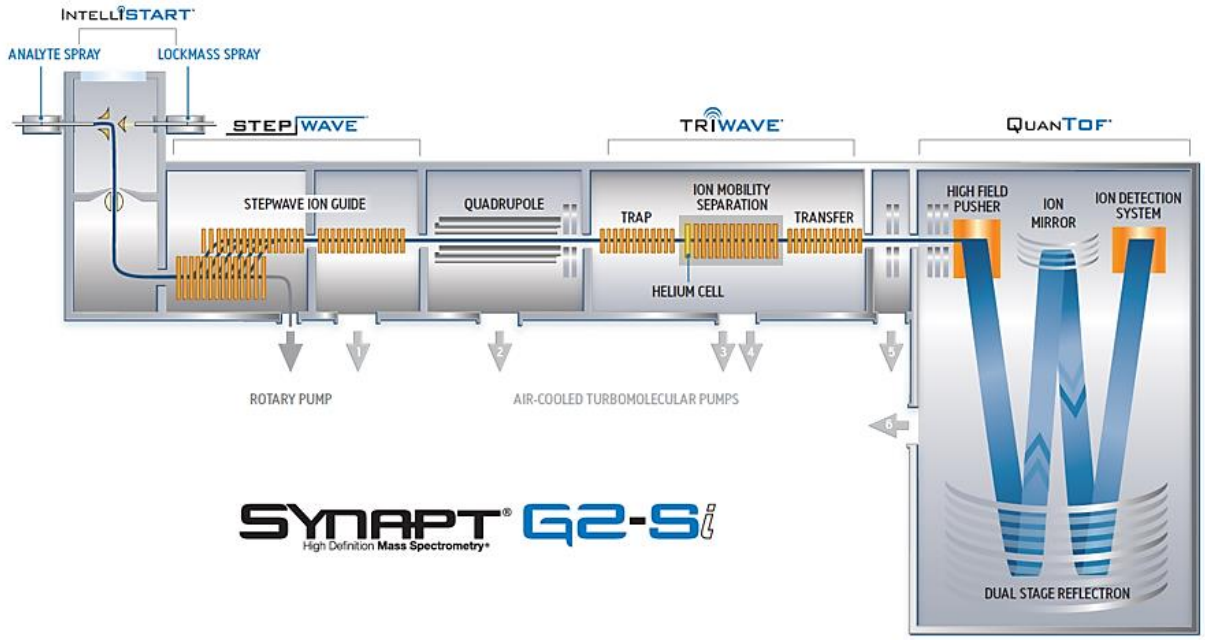


Figure 5 | **The Waters Synapt G2-Si.**[40] The hybrid mass spectrometer has both a quadrupole and TOF mass analyzer, separated by three pressurized cells for fragmentation and size separation.

Throughout the instrument, stacks of ring-shaped electrodes called Ion Guides confine gas phase ions in space to counter their repulsion forces and form the pathway through which ions

fly. Near the MS entrance, the StepWave, an offset ion guide, filters out neutral gas-phase species given that they do not feel the change in path/electric field. The narrowing of the second stage of the StepWave focuses the ions further. The ions then enter the Quadrupole which acts as a mass filter (see **section 1.2.3**), followed by 3 collision cells: Trap, Ion Mobility Separation, and Transfer. Collision cells contain gas and have tunable voltages, enabling analyte activation and fragmentation, as well as gas-phase size separation (see **section 1.2.4** and **1.2.5**). Finally, the ions pass through a TOF tube (see **section 1.2.3**) and reach the detector, which in this instance is a microchannel plate (MCP) that generates signal by converting ion impact into an electron cascade.[41]

1.2.3. Mass Analyzers: Time-of-Flight and Quadrupole Filtration

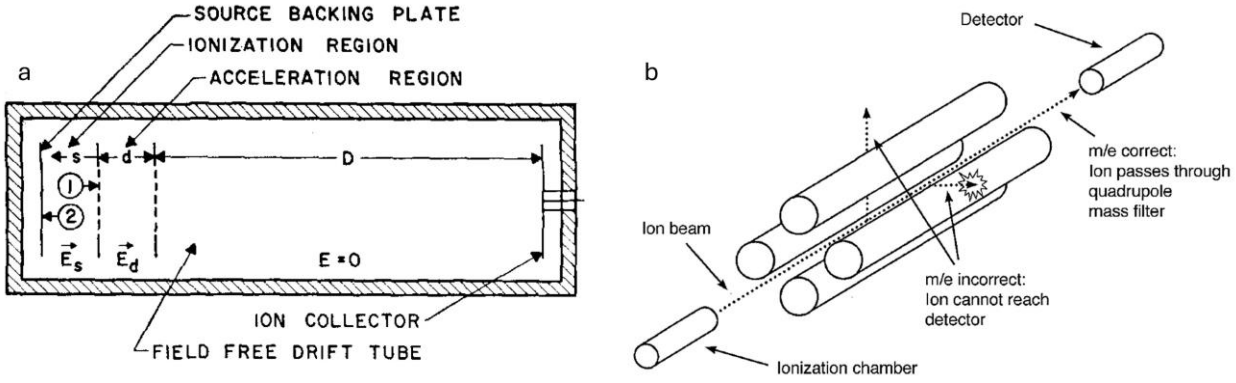


Figure 1.6 | **Mass Analyzers.** (a) Linear Time-of-Flight reproduced with permission from Wiley & McLaren (2004).[42] (b) Quadrupole reproduced with permission from Honour (2003). [43]

Time-of-Flight involves (i) the brief acceleration of ions and (ii) their subsequent movement through an empty, field-free space (**Fig. 1.6a**). Consider an ion with total charge total charge *q* where *z* is number of charges and *e* is elementary charge quantity, 1.60×10^{-19} C (**Eq. 1.4**). This is possibly the simplest form of MS to describe because in (i), a stationary ion will have a potential energy *PE* equal to its total charge *q* along an electric potential difference ΔV, and then upon moving, since energy is not created or destroyed, *PE* is converted to kinetic energy *KE* of the ion which is a function of its mass m and velocity *v*.[41] (**Eq. 1.5**)

$$q = ze \quad (1.4)$$

$$ze\Delta V = \frac{mv^2}{2} \quad (1.5)$$

The chamber is under vacuum pressure such that there are no species present to cause collision, leaving the mean free flight path of the ion to be equal to the length of the TOF tube, *D*. Since *v*

is distance *D* over time *t*, and *D* and ΔV are known constants, *m/z* is a function of *t* during TOF analysis (**Eq. 1.6**).

$$\frac{m}{z} = \frac{e2\Delta V}{D^2} t^2 \qquad (1.6)$$

In a simple example, consider two ions with the same charge (*z* =1) but $m_1 > m_2$. Given that $m \propto t$, the heavier ion, m_1, would take a longer time *t* to traverse *D*.

Modern TOF analyzers are typically Reflectrons which are TOFs fitted with ion mirrors that elongate the flight path (typically in "V" or "W" patterns). This increases the observable mass range because high *m/z* ions, with similar, slow velocities and long drift times, will diverge in their relative positions over longer distance, enabling separation prior to detector impact. The main purpose of ion mirrors is to correct for ion spread in the packet due to relative positioning of the ions along electric field in the accelerating region (i).[44] Ions penetrate the ion mirrors at depths proportional to their kinetic energy and are reflected back out such that ions with the same *m/z* are narrowly distributed upon reaching the detector, generating sharper, resolved peaks in mass spectra. This greatly improves peak resolution between similar *m/z* ions and even an analyte's isotopic distribution.

Quadrupole mass analyzers, often abbreviated "Q", are composed of four parallel rods that generate a dynamic electric field (**Fig 1.6b**).[43] Lengthwise (z-direction), the four poles form a corridor through which ions fly, and ion trajectory (in the xz- and yz-planes) is dependent on mass-to-charge.[43,45] Ions with stable trajectories in both the xz and yz-planes survive the corridor and reach the detector, whereas unstable ions veer off course and neutralize by hitting a pole. In modern practice, as a stand-alone analyzer, quadrupoles do not have sufficient resolution for complex samples (~2000 R_{FWHM}) but can be used for mass measurements of moderately pure analytes, particularly at the low *m/z* range.[46,47] Instead, a quadrupole could be placed upstream of other mass analyzers such as a TOF to act as gas-phase filters of bulk sample by isolating ions of narrow *m/z* ranges (like in the Synapt G2).

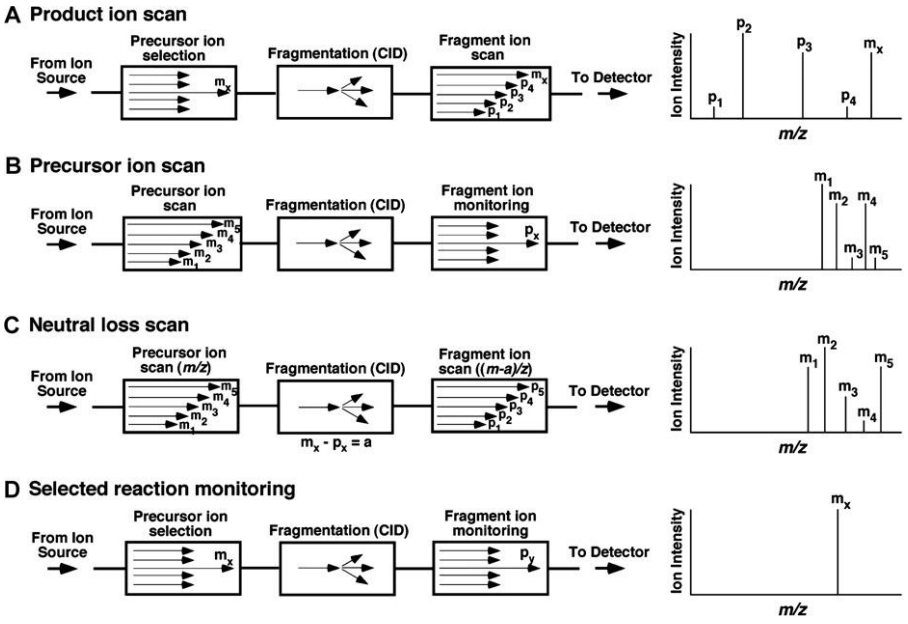


Figure 1.7 | **Triple Quadrupole Tandem MS Experiment Variations.** Reproduced with permission from Han et al. (2013).[48] (a) product ion and (b) precursor ion scans have one of three quadrupoles set to allow all ions through (no filtering) whereas another selects for a specific *m/z*. (c) neutral loss scan monitors the mass difference between precursor and product ions equal to a neutral loss. (d) selects for a specific m/z in both the first and last quadrupole.

Quadrupoles can also be placed in tandem with each other; the Triple Quad (QqQ) can be set to various modes of selection (**Fig. 1.7**).[48] In **Fig. 1.7A**, product ion scanning mode has the first quadrupole select a precursor ion (narrow m/z), the second quadrupole fragments it (pressurized collision cell), and the last one has an open scan for all product ions (all *m/z*) to pass to the detector. This setup is typically used for peptide identification where the peptide *m/z* is selected, and then the product fragments are used to validate the sequence like puzzle pieces (see Section 1.2.5). In another workflow called select/multiple reaction monitoring (SRM or MRM), both the precursor and product ions are selected for (**Fig. 1.7D**).[48] This strategy is often used in reverse-phase liquid chromatography (RP-LC) method validation. Quantitation in tandem MS can be achieved in many ways, including spiking in a known amount of isotope labeled analyte or comparing intensity to a concentration/calibration curve.[49]

Each mass analyzer type has a particular linear mass range for which mass accuracy is considered sufficient. Mass accuracy is commonly reported in Da units or quantified by error in ppm (**Eq. 1.7**).

$$error(ppm) = \frac{m_{theoretical} - m_{experimental}}{m_{theoretical}} \cdot 10^6\ ppm \quad (1.7)$$

The peak resolution, mass range, and mass accuracy are lowest in quadrupoles, while TOF has the highest theoretical mass range, and ICRs have the highest resolution and mass accuracy (**Table 1.1**).[46,47]

Table 1.1. **Parameters of Common Mass Analyzers***

Mass Analyzer	Mass accuracy	m/z range	Resolution
Quadrupole	0.1 (~100 ppm)	Up to 4,000 Da	Up to 2,000
Time-of-Flight	0.001 (~5 ppm)	Up to 10,000,000 Da	Up to 50,000
Ion Cyclotron Resonance	0.0001 (~1 ppm)	Up to 20,000 Da	Up to 5,000,000

* Rough estimates adapted from Clark et al. (2013) and Cunsolo et al. (2014).[46,47]

1.2.4 Ion Mobility Separation

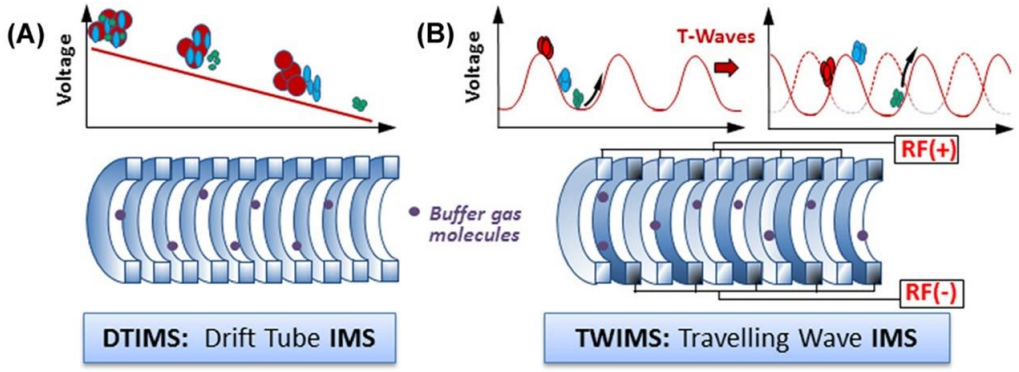


Figure 1.8 | **Ion Mobility Separation Instruments.** (A) Drift Tube and (B) Traveling Wave Ion Mobility Separation reproduced with permission from Delvaux et al. (2021).[50]

Ion Mobility Separation (IMS) utilizes the interrupted flight path of gaseous ions as they pass through a neutral buffer gas along a weak electric field (**Fig. 1.8**).[50] It is commonly compared to polyacrylamide gel electrophoresis (PAGE) in that small *sized* species travel faster than those that are larger. The key principles are that gentle gas collisions increase the mean free flight path of analyte ions without inducing fragmentation, and gas-analyte collision frequency is proportional to the size of the analyte. Low mass, chemically inert gases such as Helium and Nitrogen (4 and 28 amu) do not react to form byproducts or cause analyte fragmentation during collisions. Compared to Helium, Nitrogen is far less expensive, larger, sensitive to temperature (i.e., diatomic, vibrational modes), polarizable, and can undergo transient dipole-dipole interactions. This has been observed to increase peak capacity, the amount of separated species per analytical run, and peak-to-peak resolution.[51]

In IMS, ions are separated by their mobility *K* which is the proportionality constant between analyte ion velocity *v* and the electric field *E* (**Eq. 1.8**).[52]

$$K = \frac{v}{E} \qquad (1.8)$$

However, this mobility relationship is complicated by gas collisions (Mason-Schamp **Eq. 1.9**).[53] In this mathematical model, *K* is influenced by the analyte charge *q*, reduced mass of the ion-gas pair μ (**Eq. 1.10**), the Boltzmann constant k_B, the temperature *T*, the particle density N_{dens}, and the momentum transfer collisional cross section (CCS) of the gas-analyte pair Ω_{CCS}.[52]

$$K = \frac{3}{16}\sqrt{\frac{2\pi}{\mu k_B T}}\left(\frac{q}{N_{dens}\Omega_{CCS}}\right) \qquad (1.9)$$

$$\mu = \frac{m_{ion}\, x\, m_{gas}}{m_{ion} + m_{gas}} \qquad (1.10)$$

Importantly, in this context, the 'drop shadow' cross section σ which can be applied to hard, spherical objects (**Eq. 1.11**) is an oversimplification compared to the momentum-transfer collisional cross-section, which factors in the ensemble of gas-ion collisions, the electric field *E* (potential difference ΔU over distance *L*), and the time dimension, drift time t_{drift}, typically on the order of milliseconds (**Eq. 1.12**).[52]

$$\sigma = \pi(r_{ion} + r_{gas})^2 \qquad (1.11)$$

$$\Omega_{CCS} = \frac{3}{16}\sqrt{\frac{2\pi}{\mu k_B T}}\left(\frac{ezt_{drift}\Delta U}{N_{dens}L2}\right) \qquad (1.12)$$

Even then, caution is encouraged when deriving collisional cross section from IMS because IMS explicitly measures ion mobility as a function of time traversing the cell, not ion surface size.[52,54] Particularly, beyond ion-gas mass (**Eq. 1.10**), CCS is impacted by the buffer gas identity (e.g., as buffer gas polarizability increases, CCS increases) and the degrees of freedom of the analyte (e.g., can enable ion CCS compaction), which are not parametrized in **Eq. 1.12**.[51] As such, IMS is best employed in qualitative or comparative applications, rather than quantitative.[52,54]

Some IMS devices can be distinguished by the way in which voltage is applied (**Fig. 1.8**). The simplest form, Drift Tube IMS (DTIMS), uses a uniform applied potential difference across the length of the flight tube (~1000s V), whereas Travelling Wave IMS moves ions using traveling waves (T-Waves) formed by pulsing voltages (10-40 V) along groups of ring electrodes.[55] Ions in TWIMS are described relative to their position about the T-Wave: high mobility species (few collisions, small) will 'surf the wave' and low mobility species (many collisions, large) 'roll over' the

wave. The time it takes for an ion to traverse a TWIMS is referred to as arrival time, rather than drift time, because mobility is dependent on the T-Wave amplitude and velocity.[50] Given the linear nature of DTIMS, collisional cross-section can be derived mathematically from drift time (**Eq. 1.12**), while TWIMS requires size standards (like a ladder of proteins in SDS-PAGE), derived from DTIMS, to calibrate and determine the collisional cross-section of unknowns.[52]

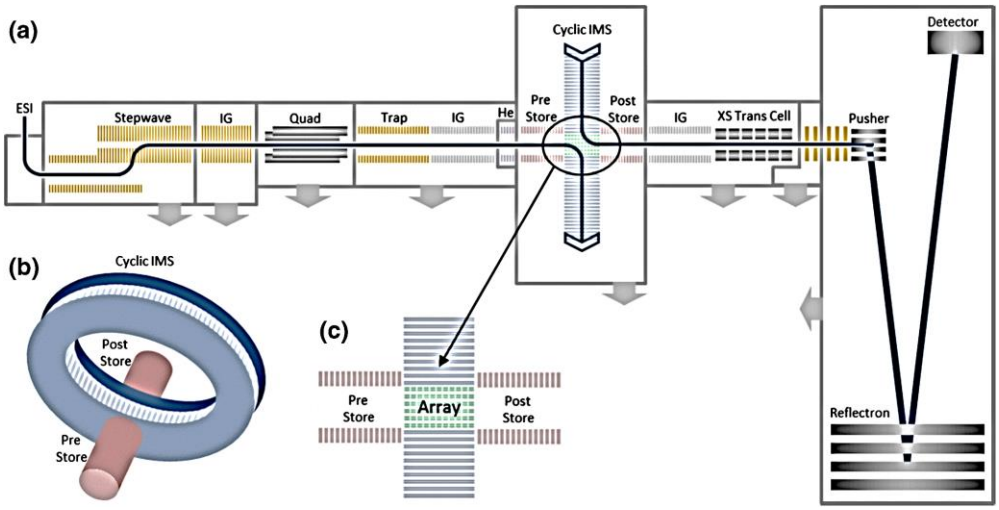


Figure 1.9 | **Waters Select Series Cyclic-IMS MS.** (a) Q-cIMS-TOF Hybrid MS with pre- and post-cIMS collision cells (Trap and XS Transfer), similar to the Synapt G2, (b) the Cyclic IMS "racetrack", (c) the electrode array which directs ions in, out, or along the cIMS. Reproduced with permission from Ujma et al. (2019). Copyright © 2019, American Chemical Society.[56]

Improved peak capacity of TWIMS has been achieved recently through the development of the Waters Select Series Cyclic IMS (cIMS) MS (**Fig. 1.9**).[56] A single pass on the "racetrack" is 4x the length of the Synapt G2 TWIMS (~25 cm). An electrode array at the inlet/outlet of the racetrack uses T-waves to direct ions and enables the analytes to move along the cIMS for multi-pass separation, a theoretically infinite path length. Coupled with LC at the MS front end, single pass mode is sufficient to obtain very high peptide sequence coverage of massive proteins or protein mixtures, beyond the Synapt G2 capacity, although multiple passes reduce coverage.[57] Prolonged residence in the TWIMS can cause ion loss, heating/activation, and/or fragmentation, and therefore may reduce sensitivity; however, this phenomenon can be harnessed for studies of gas-phase protein unfolding.[58] Ions in multi-pass cIMS may also experience wrap around effects; ions with high mobility overtake those with slower mobility.[57]

Major applications of IMS include complex mixture separation (e.g., peptides of a digested protein, many of the -omics', etc.) and Collision-Induced Unfolding (CIU, see section 1.2.5). CIU

harnesses the ability of IMS to separate overlapping m/z signals of structural isomers, particularly for larger *m/z* ions, including proteins sampling distinct conformations/unfolded states.

1.2.5 Ion Activation and Dissociation

Gas-phase analyte fragmentation has a key application in structural elucidation. A major strategy in bottom-up proteomics, for example, is proteolyzing mixtures of proteins (from e.g., co-immunoprecipitations or whole cell lysates) and then using MS to identify which proteins were present. The product ions (fragments) of proteolyzed peptides (precursor ions) can be pieced together to elucidate peptide structure and then primary sequence of an intact protein.

Ion fragmentation begins with activation, or increasing the internal energy of an ion, and can be achieved by collisions with gas molecules, electrons, photons, etc. To facilitate dissociation of ions by gas collision, kinetic energy is converted into internal energy (i.e., inelastic collisions).[59,60] If a collision with gas *M* increases the internal energy (e.g., vibrational, etc.) of a species AB^+, this is considered activation $[AB^+]^*$ (**Eq. 1.13**).

$$Step\ 1: \quad AB^+ + M \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} [AB^+]^* + M \qquad (1.13a)$$

$$Step\ 2: \quad [AB^+]^* \xrightarrow{k_{critical}} [AB^+]^{\#} \qquad (1.13b)$$

$$Step\ 3: \quad [AB^+]^{\#} \xrightarrow{k_2} A^+ + B \qquad (1.13c)$$

When the species has reached a critical internal energy level $[AB^+]^{\#}$, also called the transition state, the species undergoes a unimolecular reaction causing its decomposition into fragments A^+ and *B* (fragmentation pattern can vary).[59] Redistribution of internal energy often occurs faster than the unimolecular reaction ($k_1 >> k_{critical}+k_2$) and $k_{critical}$ depends on the number of degrees of freedom of a given molecule. As such, it takes more collisional activation, through means such as higher kinetic energies of analytes or heavier mass gases, to reach the critical energy of the transition state for larger molecules.

The reversibility of step one of this reaction can be considered at two extremes: high and low pressure of *M*.[60] At the high pressure limit, activated ions can undergo deactivation faster (k_{-1}) than irreversible dissociation (k_2 + $k_{criticial}$) due to the high probability of collisions. At the low-pressure limit, the forward bimolecular collision is the rate limiting step, followed by a faster

unimolecular dissociation (i.e., reverse bimolecular collisions are infrequent). Consider the TriWave region of the Waters Synapt G2: the high pressure of the IMS cell (3 mbar) reduces fragmentation compared to the Trap and Transfer (0.05 mbar).[61]

If activation is energetic enough that it causes immediate bond cleavage without energy redistribution, this is considered to be non-ergodic.[62] However, most collisional activated dissociation methods are considered ergodic, where redistribution of internal energy at all possible internal states (including structural isomerization) occur prior to cleavage.

Collision-Induced Dissociation (CID) is achieved by using a cell lined with focusing electrodes that accelerate ions through a pressurized space (e.g., Ar, He, N_2, etc.). In the context of peptides, unimolecular reaction often happens about the peptide bond after internal energy has sufficiently increased; however, the cleavage mechanism is not definitively understood. Additionally, the overall fragmentation pattern of peptides is not always easy to predict; in practice, not all peptide bonds between various residues undergo cleavage with equal probability.

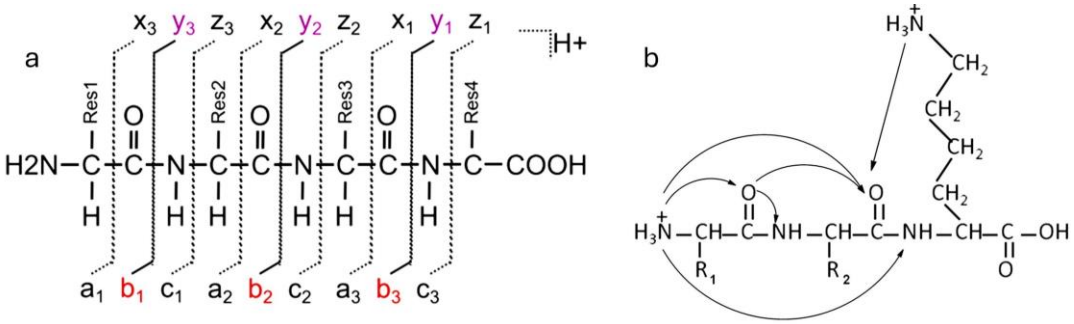


Figure 1.10 | **Peptide Backbone Fragmentation.** (a) Peptide Fragmentation and Roepstorff-Fohlmann-Bremann Nomenclature reproduced with permission from Hjernø & Højrup (2015).[63] and (b) Mobile Proton Model reproduced with permission from Boyd & Somogyi (2010), Copyright © 2010, Elsevier B.V.[64]

Fortunately, there are rules that can be employed to deduce peptide primary sequence, mainly by using mass differences between- and masses of- fragment ion peaks since they should add up to the precursor ion mass. The Roepstorff-Fohlmann-Bremann nomenclature for backbone fragments is illustrated in **Fig. 1.10a**.[63] Cleavage at the peptide bond generates b- and y-ions which contain the peptide N- and C-terminus, respectively. The fragment ions are numbered in subscript chronologically by residue starting from their respective terminus (e.g., b3 and y1 will be at the third and last residue of the peptide). After fragmentation, only positively charged, stable species will continue to fly through the MS to reach the detector which in turn sometimes results in gaps from fragments that went undetected.

Also, secondary forms of fragmentation can cause neutral losses of ammonia (-17 Da) and water (-18 Da), while the loss of CO (-28 Da) from b-ions form a-ions (**Fig. 1.10**). When secondary fragmentation causes loss of both the C and N terminal moiety, this is called an internal fragment, and when this spans only one residue sidechain (R), it is considered a diagnostic immonium ion ($[H_2N=CHR]^+$) that can be used to determine residue identity and corroborate other spectral fragments.[63]

The b- and y-ions are thought to form as a consequence of backbone protonation (e.g., at carbonyl oxygen or amide nitrogen) which weakens the peptide bond. Since ions are desolvated/dry in the gas phase, standard water-mediated acid-base rules do not apply, giving rise to "Mobile Proton Model".[65] Here, proton migration along a peptide in the gas phase depends on the internal energy and basicity of the residue that the proton is originating from. Wysocki et al. (2001) suggested that a proton can move to backbone heteroatoms from a protonated basic sidechain due to sufficiently increased internal energy from collisional activation (**Fig. 1.10b**).[65] This model is supported by the proportionality between peptide basicity and the collisional voltage required for fragmentation. For their work in developing the Mobile Proton Model framework, Wysocki and Gaskell received the 2010 American Society for Mass Spectrometry's Distinguished Contribution Award.[64]

One possible mechanism of Mobile Proton Model involves the protonation of the amide carbonyl oxygen followed by charge-directed nucleophilic cleavage by the adjacent N-terminal carbonyl oxygen on the (now more) electrophilic carbonyl carbon (**Fig. 1.10b**). This pathway is consistent with the *infrequency* of b_1-ions (because there is no C=O N-terminal to the first residue in a peptide) and the infrequency of cleavage C-terminal to proline (whose carbonyl is sterically hindered from nucleophilic attack by upstream carbonyl oxygen).[65] Uniquely, arginine is found to be so basic that it 'sequesters' protons from being mobile and, therefore, when precursor ion charge is not in excess of arginine, less product ions form. Also, acidic residues (Asp, Glu) in arginine-containing peptides often undergo selective C-terminal cleavage and are thought to be the mobile proton donator.

In the Waters Synapt G2, peptide CID is achieved using relatively low voltages (up to ~35 V) and Argon gas. The positioning of the 3 collision cells (Trap/TWIMS/Transfer) downstream of the quadrupole enables high resolution tandem MS (also known as MS/MS, MS^2, etc.) (see **Fig. 1.5**). The quadrupole can select for a narrow m/z range, allowing only select precursor ion(s) to pass, and the trap/transfer cells can form fragments (e.g., product ions) before/after the IMS cell. The transfer cell position enables a dimension of precursor ion separation in the IMS prior to

fragmentation which reduces the complexity of downstream spectra. Efficient peptide identification is particularly achieved by rapidly acquiring spectra at low and high collision energy, resulting in both rich intact peptide and peptide fragment data per analytical run (MS^e mode by Waters). Further dimensions of separation at the front end (e.g., LC) can additionally reduce spectral overlap and complexity, improving sequence coverage and sensitivity to lower abundance ions.

On the other hand, larger protein species and complexes have more degrees of freedom, allowing for redistribution of internal energy and consequential unfolding events rather than fragmentations. In the Waters Synapt G2, Collision Induced Unfolding (CIU) is conducted by ramping collisional voltage (sometimes beyond 100 V) in the Trap cell, followed by TWIMS. Here, unfolded ions of the same *m/z* separate based on collisional cross section, increasing in arrival time with increased unfolding. For a given m/z signal, the drift time peak can be plotted as a function of collisional activation voltage, which can be used as a means to evaluate gas-phase ion stability via unfolding/transition events.

1.3 Hydrogen-Deuterium Exchange of Proteins

Protein Hydrogen-Deuterium eXchange (HDX) is a covalent labeling reaction in which solvent deuterium (e.g., D_2O) replaces protein heteroatom-bound hydrogens. Given the 1 amu difference for every H-to-D swap, the HDX reaction is readily measured by monitoring mass change using mass spectrometry (HDX-MS). The rate of exchange is inherently correlated with protein structure (at all levels, discussed below) and the reagents used in experimentation are becoming increasingly more accessible, propelling HDX-MS forward as a more commonly used structural characterization method. Particularly, HDX-MS is an analyte-dilute, solution phase method which, at the expense of atomic resolution, is less susceptible to some of the drawbacks of gold standard methods like X-ray Crystallography and NMR.

1.3.1 HDX Fundamentals: from residues to proteins.

The HDX reaction can be base-, acid-, and water-catalyzed, and the chance of successful proton transfer increases when the proton affinity, pK_a, of the deuteron donor is lower than the deuteron acceptor ($pK_{a,donor} < pK_{a,\ acceptor}$).[66] The larger this difference in pK_a, the higher the probability that a collision with a given catalyst (D^+, OD^-, D_2O) will transfer a proton/deuteron, and the faster the HDX rate. Since pKa describes the protonation state of a given functional group, pH is a major factor in HDX reaction pathways. The general catalytic pathways of protein HDX are described in **Fig. 1.11**.[66]

During base catalysis in excess D_2O (Fig. 1.11a), deuteroxide (OD^-) abstracts a proton from the amide backbone forming an amidate ion which can then abstract a deuteron from D_2O. On the other hand, acid catalysis may occur either by (i) N-deuteration of the amide nitrogen followed by deprotonation of the nitrogen by water (Fig. 1.11b), or (ii) O-protonation of the amide carbonyl oxygen followed by deprotonation of the imide nitrogen (resonance) by water, and subsequent deuteron abstraction from D_3O^+ (Fig. 1.11c). Of the two, NMR analysis revealed (i) occurs for electron donating sidechains, while (ii) is more commonly observed in amides near electron withdrawing sidechains; however, (ii) requires that the amide C=O and NH swap their hydrogen bond acceptor/donator behaviour in a structured system, implicating unfolding in that region.[67]

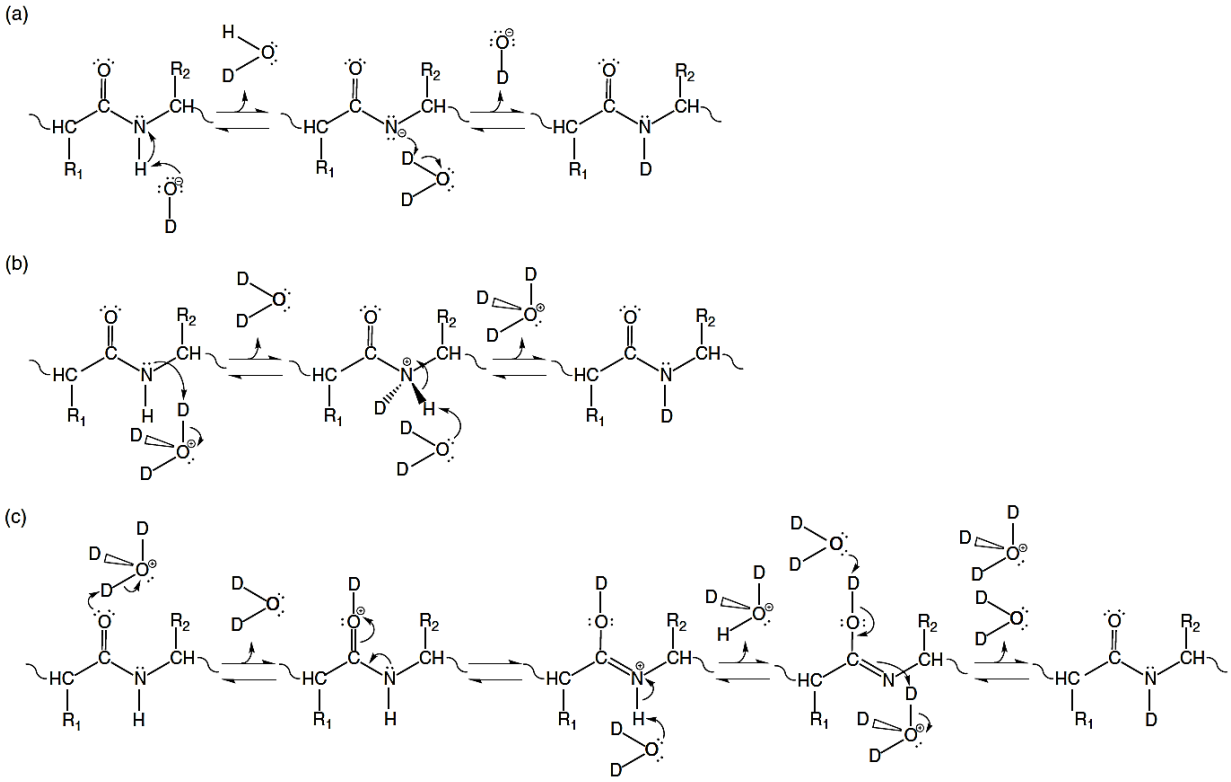


Figure 1.11 | **Intrinsic HDX catalytic pathways**. (a) base catalysis, (b) N-protonation mediated acid catalysis, and (c) O-protonation mediated acid catalysis. Reprinted with permission from Weis (Ed.). (2016) "Hydrogen Exchange: A Sensitive Analytical Window into Protein Conformation and Dynamics." [66]

Since all catalytic pathways can occur, just at different incidences depending on the free catalyst available, the sum of the intrinsic catalytic pathways makes up the overall HDX rate constant k_{ex} (**Eq. 1.14**).[66,68]

$$k_{ex} = k_{int,acid}[D^+] + k_{int,base}[OD^-] + k_{int,water} \quad (1.14)$$

In foundational work by the Englander lab, Bai et al. (1993) experimentally determined the intrinsic rate constants for acid-, $k_{int,acid}$, base-, $k_{int,base}$, and water-, $k_{int,water}$, catalysis of various model peptides including poly-DL-alanine (PDLA) with 1D ^{1}H-NMR; diluting peptides into D_2O and monitoring NH signal decay as a function of time.[68] PDLA was reported to be a randomly synthesized polymer of racemic D and L alanine, with an average degree of polymerization (dp) of ~28 monomeric units. At 25 °C, PDLA intrinsic HDX rate constants were: $k_{int,acid}$ = 41.7 M^{-1} min^{-1}, $k_{int,base}$ = 1.12 x 10^{10} M^{-1} min^{-1}, and $k_{int,water}$ = 0.0316 min^{-1}.[68] The exchange rate constant k_{ex} of PDLA at different pH_{read} is shown in **Fig. 1.12a**.[69] It was calculated using the D_2O molar dissociation rate constant $K_{D2O}(20°C) = 10^{-15.05}$, as well as corrected deuteron and deuteroxide concentrations (i.e., $[D^+]$ and $[OD^-]$ from pD_{corr} and $pOD_{corr,}$ where $pD_{corr} = pH_{read} + 0.4$).[70,71]

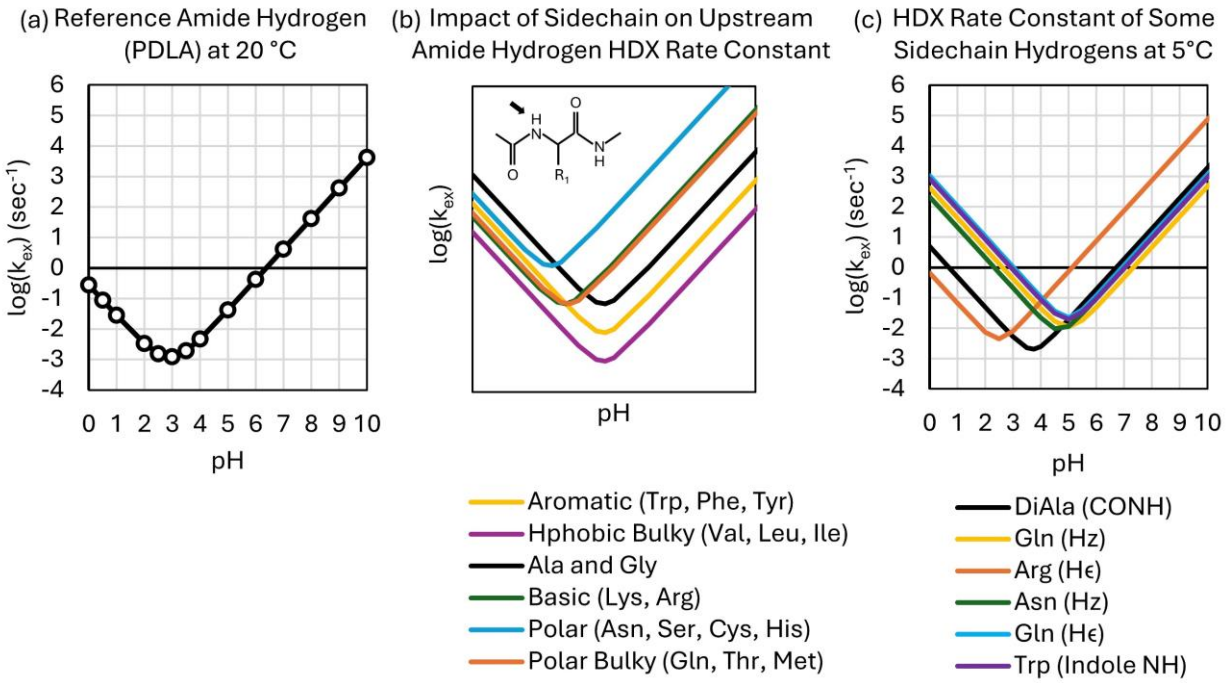


Figure 1.12 | **HDX rate constant, k_{ex}, dependence on pH.** (a) peptide reference poly-DL-alanine (PDLA), (b) amide hydrogens upstream of various amino acids, and (c) sidechain hydrogens. See (b) inset for dipeptide structure. Adapted from Smith, Deng, and Zhang (1997) and Hamuro (2021) based on experimental results of Bai et al. (1993) and Molday et al. (1972).[68,69,72,73] Note that (b) depicts general trends by translation of the V-shaped curve f(x) of Ala: Aromatic is f(x)-a, Hphobic bulky is f(x)-b (where b>a), Basic and Polar Bulky is f(x-c), and Polar is f(x-d)+e (where c<d).

Since PDLA is randomly coiled[74] and lacks exchangeable sidechain hydrogens, three factors about backbone amide exchange were highlighted: (i) the majority of amide HDX is base catalyzed ($k_{int,base}$ >>> $k_{int,acid}$ > $k_{int,water}$), especially at physiologically relevant pH (~ pH 6-8), (ii) there is a pH (~ 2.5-3), denoted pH_{min}, at which the minimum exchange rate constant occurs, and

(iii) acid-catalyzed exchange dominates below pH_{min} due to the low concentration of OD^-, although it is 1000x slower (e.g., at 20°C pH 2 and 7, k_{ex} of PDLA is 0.00316 sec^{-1} and 4.17 sec^{-1}, respectively).

Bai et al. (1993) also examined the backbone exchange of di-peptide-bond models (i.e., $CH_3(CONH_A)CHR(CONH_B)CH_3$) of various residues across the pH scale (**Fig. 1.12b**).[68] Generally, H_B was observed to have a lower rate constant than H_A. Focusing on H_A, relative to an alanine (DiAla) standard, sidechain specific trends in V-shaped curve translation were clear. When R corresponded to the hydrophobic, bulky moieties of Val, Ile, and Leu, the V-shaped curve exhibited a downward translation, indicating reduced k_{ex}, while aromatic residues (Tyr, Trp, Phe) had a more moderate downward shift. These results suggested that sidechains incur both steric and inductive effects on backbone HDX. Additionally, polar residues (Ser, Thr, Cys, Met, Asn, Gln) generally showed a leftward shift in pH_{min}, consistent with an amide proton that is more readily abstracted in base-catalyzed HDX due to electron density withdrawal, whereas acidic residues (Aps, Glu) imparted a horizontal stretch on the V-shaped curve (not shown), resulting from slowed exchange in both acidic and basic conditions.

In an extension of their experiment, Bai et al. (1993) also monitored heteroatom-bound sidechain hydrogen exchange.[68] They observed a right, upward translation in V-shaped curves for several residues relative to amide hydrogen: Gln-NH_ϵ, Gln-NH_z, Asn-NH_z, Trp-N_1H (**Fig. 1.12c**, apart from Arg-NH_ϵ, which had a left, upward translated curve). More recently, Hamuro (2021) found this right, upward translation to be consistent with theoretical and experimental curves of most other sidechain hydrogens (e.g., SH, OH, Phenyl-OH, Arg-NH_η).[72] These findings suggest that in most cases, sidechain hydrogens undergo acid catalyzed HDX more readily than the amide hydrogen, while rates at physiological pH are quite similar to amide hydrogen.

HDX rate is also influenced by temperature. Take for example, a pH probe reading that increases as temperature is reduced. Although it may be tempting to assume the acidity of the solution is changing, it is the *ionization* of <u>both</u> acid and base that are impacted ($K_W = [OH^-][H^+]$). Since ionization is reduced at lower temperature (e.g., At 1 atm, pK_W(0 °C) = 14.95 vs. pK_W(25 °C) = 13.99),[75] the concentration of available catalyst is reduced as well, slowing the HDX rate. This manifests as a downward shift of the entire V-shaped curve.

Overall, the exchange rate constants observed above for short peptide models were indicative of fast exchange at heteroatom-bound hydrogens. Their rapid $k_{int,base}$ (25°C) span 10^9-10^{12} M_{OH}^{-1} sec^{-1} and consequently have k_{ex}(25°C) at pH 7 spanning ~10^2-10^5 sec^{-1}.[72]

Long before these quantitative studies were completed, in the 1950s, Kaj Ulrik Linderstrøm-Lang, the father of protein HDX, began to observe distinctions in HDX rate between proteins and short peptide models; proteins fully exchanged slowly, over mins-hours.[77,78] Inspired by ground breaking work by Pauling, Corey, and Branson (1951) on the structure of the alpha-helix,[79] he reasoned that higher order structure, incurred by hydrogen bonding at the amide backbone, reduced solvent accessibility, slowing amide HDX to an observable rate.[76] He further described that proteins undergo structural isomerization via breathing motions, or local and global unfolding events, which manifest from formation and breaking of non-covalent hydrogen bonds in secondary structure. When sampling an unfolded conformation, amide hydrogens briefly have access to solvent and undergo exchange.

One can appreciate the convenience of studying protein structure with HDX given the uniform distribution of backbone amides. HDX can shed light on: the extent of structure based on the percent of backbone hydrogen exchange, the stability of structures over time (i.e., breathing motions, dynamics), and perturbations to structure in one state relative to another.[76]

To extract meaningful structural information, it is necessary to delineate between *intrinsic* and *observed* HDX in the context of higher order polypeptides and proteins compared to model dipeptides, etc. As seen in **Fig. 1.12**, intrinsic rate, k_{ex}, can be determined from the primary sequence alone because it is dependent on how amide hydrogen acidity/abstractability is influenced by adjacent, covalently linked functional groups, mainly through induction and steric hindrance at a given pH and temperature.

From comprehensive primary structure studies in the Englander lab (Bai et al. 1993 & Molday et al. 1972), quantitative factors were derived, relative to alanine, describing the impact that sidechains left, *L,* and right, *R,* of a peptide amide bond have on the acid *A* and base *B* 'arms' of that amide's V-shaped curve (**Eq. 1.15)**.[68,73] See appendix A for L and R factors in **Table A1**.

$$L \text{ or } R_{A\ (or\ B)} = logk_{ex,A\ (or\ B)\ of\ sidechain} - logk_{ex,A\ (or\ B)\ of\ Alanine} \quad (1.15)$$

These pH and temperature-specific *L* and *R* factors could then be used to determine $k_{ex,NH}$, the site-specific amide exchange constant (**Eq. 1.16**).[68,73]

$$k_{ex,NH} = k_{A,alanine}(L_A \times R_A)[D^+] + k_{B,alanine}(L_B \times R_B)[OD^-] + k_{water,alanine}(L_B \times R_B) \quad (1.16)$$

Finally, the extent of deuterium uptake D_t for a free peptide amide at a given timepoint *t* could be determined using exponential decay (**Eq. 1.17**).[80] Here, D_t=1 when the full population of this amide has exchanged and D_t=0 when *t*=0.

$$D_t = 1 - e^{-k_{ex,NH}t} \quad (1.17)$$

Consequently, a peptide that is entirely unstructured will undergo intrinsic exchange that can be predicted from the sum of each amide's D_t across the peptide.

The difference between intrinsic and observed rate stems from higher order structure, as noted by Linderstrøm-Lang, attributed to hydrogen bonding and solvent inaccessibility. Particularly, backbone amide hydrogens that act as hydrogen bond donors restrict access of the catalyst (D^+, OD^-, D_2O), slowing the rate of exchange below intrinsic, such that increasingly structured regions exchange less frequently with deuterium. The extent of structure relative to a denatured peptide or protein can be quantified using protection factors, *PF* (**Eq. 1.18**). The larger the PF for a given system, the more structured it is.

$$PF = \frac{k_{ex}}{k_{obs}} \geq 1 \quad (1.18)$$

The modulation of intrinsic exchange (k_{ex}) by unfolding events (k_{open}, k_{closed}) at backbone amide hydrogens (NH) was described by the following general reaction mechanism (**Eq. 1.19**).[76,78]

$$NH_{closed} \underset{k_{closed}}{\overset{k_{open}}{\rightleftarrows}} NH_{open} \xrightarrow{k_{ex}} ND_{open} \underset{k_{closed}}{\overset{k_{open}}{\leftrightarrows}} ND_{closed} \quad (1.19)$$

Since hydrogen bond breaking is a unimolecular dissociation, k_{open} has units s^{-1} and k_{closed} has units $M^{-1}s^{-1}$. From here, the observed HDX rate constant (k_{obs}), was derived (**Eq. 1.20**).[76,81,82]

$$k_{obs} = \frac{k_{open} \cdot k_{ex}}{k_{open} + k_{closed} + k_{ex}} \quad (1.20)$$

In a completely disordered protein system, all regions are effectively in their open state *(k_{open} >>> k_{ex}, k_{closed} = 0*), simplifying **Eq. 1.20** to equal the intrinsic rate constant: $k_{obs} = k_{ex}$. In a structured protein system, regions tend to stay closed and folded (k_{closed} >> k_{open}). This simplifies to **Eq. 1.21** because k_{open} minimally contributes to the cumulative term in the denominator and can be removed.[66]

$$k_{obs} = \frac{k_{open} \cdot k_{ex}}{k_{closed} + k_{ex}} \quad (1.21)$$

Two distinct circumstances impacting k_{obs} can be explored.[66,81,82] In the event that exchange is faster than closing (k_{ex} >> k_{closed}), k_{obs} can be simplified to **Eq. 1.22**, where the observed exchange is proportional to the opening event (Hvidt and Nielson (1966), unimolecular exchange kinetics: "EX1").[82]

$$k_{obs} = k_{open} \quad [s^{-1}] \quad (1.22)$$

In practice, this labeling is characteristic of unstructured, open loops that infrequently or briefly sample a structured state and manifest as a is bimodal distribution: minimal deuteration for the closed subpopulation and maximal deuteration for the open subpopulation (**Fig. 1.13a**).

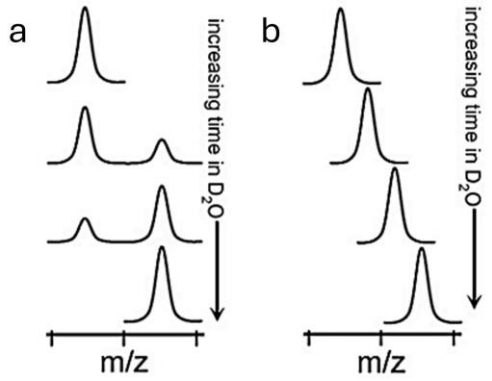


Figure 1.13 | **Population Distributions.** (a) EX1 kinetics represent distinct conformational populations and manifest bimodally, (b) EX2 kinetics occur for an averaged conformation. Reproduced with permission from Weis et al. (2006), Copyright © 2006, American Society for Mass Spectrometry.[83]

On the other hand, when k_{closed} is faster than the exchange rate constant k_{ex} (i.e., $k_{ex} < k_{closed}$), the observed rate is a function of the ratio of opened and closed states (Hvidt and Nielson (1966), bimolecular "EX2" kinetics).[82] Here, an opened state is sampled so briefly that it closes back without becoming completely deuterated. The k_{obs} can be simplified to **Eq. 1.23**:

$$k_{obs} = k_{ex} \cdot \frac{k_{open}}{k_{closed}} \quad [M^{-1}s^{-1}] \quad (1.23)$$

The deuterium incorporation in these protein regions will gradually increase as HDX labeling time is increased, up to a maximum of the exchangeable backbone amides (**Fig. 13b**). The EX2 regime is by far the more common scenario observed in peptides of structured proteins and it is easier to analyze. However, both EX1 and EX2 kinetics provide informative ways to study structural dynamics.

1.3.2 Bottom-Up Workflow for HDX of Proteins

HDX-MS can be analyzed on the global protein-level or local peptide-level, although local HDX is more common given its inherent increased spatial resolution and feasibility for various protein systems. Local HDX-MS can be bottom-up, such that the protein is proteolyzed prior to MS, or top-down, such that the intact protein is fragmented by the instrument. Currently, the most

common and best understood local-HDX workflow is the bottom-up approach, for various reasons including instrument development/limitations. For example, top-down suffers from poor CID fragmentation coverage and spectral overlap for medium-to-large proteins; however, fast fragmentation methods such as electron-capture dissociation (ECD) may reduce deuterium scrambling (e.g., Mobile Proton Model for CID) and achieve residue-level resolution provided it is upstream of a mass analyzer that can separate all the fragments sufficiently.[84]

For relevance, the following discussion will assume a bottom-up framework using proteolysis, RP-LC, and ESI-MS (**Fig. 1.14a**). Front-end implementation of autosamplers or robotic syringe handling systems coupled to temperature-controlled compartments (**Fig. 1.14b**) is ideal for reproducibility and automation of this method. These systems further benefit from high pressure LC pump and valve-based fluid delivery, modulating the flow of solvents to target locations (discussed shortly).

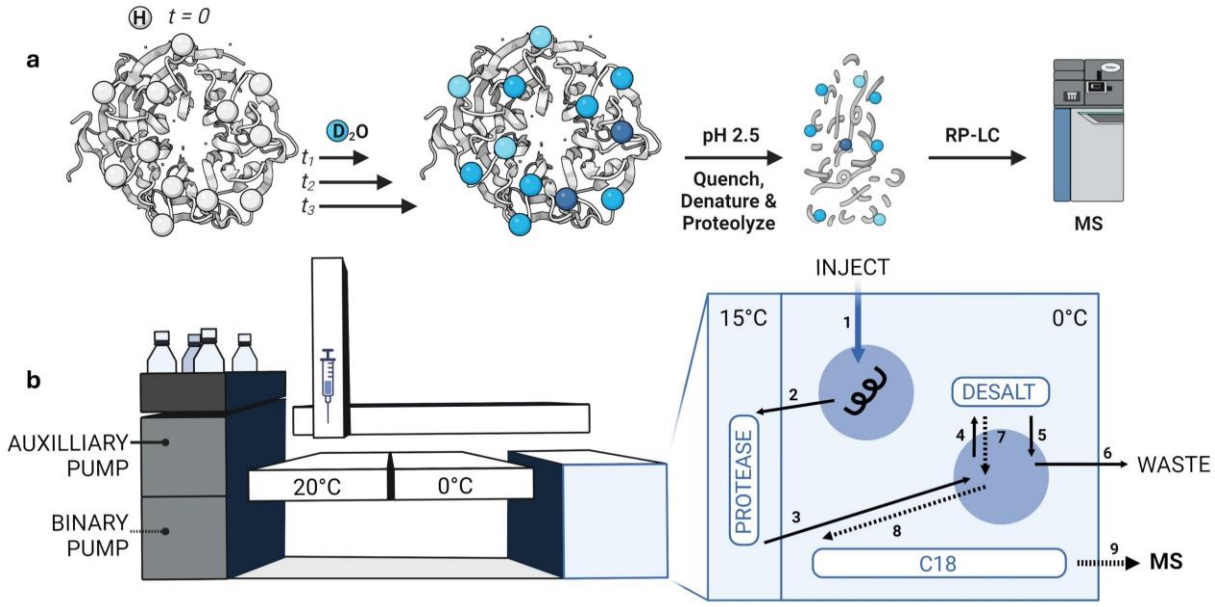


Figure 1.14 | **Bottom-up HDX-MS.** (a) Workflow: A protein undergoes HDX labeling, followed by quenching, unfolding, and proteolysis. Peptides are separated using reverse-phase liquid chromatography (RP-LC), followed by ESI-MS. (b) Temperature-controlled compartments and LC pump fluid delivery coupled to robotic liquid handling system. Samples are stored at 0°C to maintain stability, followed by HDX at 20 °C, and quench back in 0 °C. Quenched sample is injected (1) for proteolysis (2) at 15 °C. Peptides undergo desalting (3-6) and then reverse-phase separation (7-9) prior to ESI-MS. Dashed lines (- - -) indicate aqueous solvent pH 2.5 delivered by the Auxiliary Pump. Solid lines (—) indicate gradient composed of organic and aqueous solvents at pH 2.5 delivered by Binary Pump. Generated using BioRender.com and 3D structure of WDR12 (PDB 2GNQ).

To initiate the HDX reaction, steady-state protein sample is mixed by syringe aspiration-dispensation with deuterium buffer which can be prepared to best suit the protein system such as matching storage buffer, salt, and pH conditions. Labeling proceeds for duration, *t*, at constant temperature (e.g., 20 °C). As per section 1.4.1, it is incredibly important that pH and temperature are constant, and identical for all samples because these factors impact the HDX rate. For completeness of the study, the duration and number of timepoints is encouraged to span several orders of magnitude (i.e., 10-10^5 secs; secs-days) in order to capture the least and most structured backbone amide HDX.[85]

Quenching is achieved by the addition of acid, ideally at 0 °C. The quench step occurs at amide pH_{min} (2.5-3) to slow the reaction to the minimal intrinsic rate constant, k_{ex} (**Fig. 12a**). For the remainder of the solution-phase workflow, the solvent is held at pH_{min} to reduce continued exchange in proton rich conditions (i.e., back-exchange, discussed below). In many protein systems, acid helps unfold the protein by disrupting the hydrogen bond network, increasing backbone access to subsequent proteolysis. Strategies including chemical denaturant addition (e.g., Guanidine-HCl, acetonitrile, etc.), chemical reduction (e.g., TCEP), and on-line electrochemical reduction have been employed to reduce disulfide linkages and encourage disruption of extremely rigid structures.[86] While protein unfolding efficiency is time and temperature dependent, this poses a practical limitation in HDX-MS due to the need to minimize potential back-exchange.

Sidechain heteroatom-bound hydrogens exchange to deuterium extremely rapidly at physiological conditions compared to protected amides and are therefore considered to be non-informative about structure. Also, since the pH_{min} necessary for quenching/slowing the HDX reaction is distinct for amide hydrogens and sidechains (**Fig. 1.12c**), LC-coupled experiments conducted at pH 2.5, 0 °C for 5–15-minute gradients assume that only amide backbone deuteration is preserved while sidechains continue to exchange back to H. Sidechains are also less quenchable; their V-shaped curve minima are higher/faster than that of amide hydrogen.

On the topic of back-exchange, isotope effects are an interesting consideration. Since D is twice the mass of H, the vibrational freedom of H is larger.[87] In other words, it takes less energy to break an H bond compared to a D bond. This isotope effect is observed in the lower self-ionization coefficient of D_2O compared to H_2O ($pK_{H2O}(20°C) = 14.17$ vs. $pK_{D2O}(20°C) = 15.05$).[70,75] Covalent D bonds are also shorter and less polarized than H.[88] This, and the larger size of D, [89] makes D a poorer H-bond donor than H.[82] This has been evidenced by increased melting and phase transition temperatures for deuterium labeled samples.[88] Rarely, this may be observed in

longer HDX labeling timepoints; at a certain threshold of labeling, proteins crash out due to loss of structure. On the other hand, in chemical synthesis, D can be used as a protecting group to modulate regioselectivity.[90] In general, when a reaction rate changes due to the introduction of an isotope, this is called the Kinetic Isotope Effect (KIE). Indeed, there is evidence from Connelly et al. (1993) of the Englander lab that at 20°C, PDLA peptide N-D/H_2O exchange ($k_{acid}=10^{1.4}$, $k_{base}=10^{9.87}$) is slower than N-H/D_2O ($k_{acid}=10^{1.62}$, $k_{base}=10^{10.05}$).[91] During bottom-up HDX, it is likely that back-exchange (from D-to-H) is slowed due to the KIE; during quench, the stability of the deuterium label effectively reduces k_{ex} from what is predicted by **Eq. 1.11** for H-to-D exchange.

Proteases are selected based on acidic functionality, with the most common being pepsin, protease XIII, and more recently commercialized nepenthesin I&II (Nep-1 and -2). They generally catalyze cleavage via two aspartic acid residues and retain structure at low pH thanks to stabilizing disulfide bonds and rigid β-motifs.[92,93] Unlike pepsin, Nep-1 and Nep-2 have significantly reduced activity in the presence of denaturants, such as urea and Gdn-HCl at pH 2 and 25 °C, which can impact versatility in HDX-MS experiments.[93] Yang et al. (2015) observed that immobilization is correlated with increased stability, and that Nep-2 actually retained activity in denaturants.[94] These aspartyl proteases are considered broad or non-specific in their cleavage patterns due to their preference for hydrophobic residues N- (P1) or C-terminal (P1') to the peptide bond. Notably, Nep-2 cleaves efficiently at basic residues H, K, and R, whereas pepsin does not, and neither protease cleaves at P.[93,94] Pepsin and Nep-2 were observed to have unique cleavage products, suggesting utility in mixing them (1:1) for bottom-up HDX-MS.[94] In practice, the first step before conducting any bottom-up HDX-MS is the optimization of peptide-level primary sequence coverage.

The ability to extract information from the peptide mixture is greatly improved if a separation stage precedes detection in MS. In a complex peptide mixture, there is a likelihood of encountering mass spectral overlap due to peptides with similar *m/z* and, more so, overlapping isotopic distributions from deuterium labeling. Separation can be achieved using RP-LC or IMS, although the best results are obtained when using both strategies in tandem.[50] RP-LC utilizes a hydrophobic stationary phase and a gradient mobile phase which increases in percent organic over time. This allows for the gradual elution of increasingly hydrophobic species off the column. Typically, stationary phases are composed of hydrocarbon backbones, such as C18 or C4, immobilized onto ~1-2 μm diam. porous beads packed into narrow (~2 mm diam.) high pressure columns with volumes of ~150-350 mm^3 (= μL).[95] Porosity increases functionalizable surface area which compensates for the small diameter beads needed for dense packing and better resolution.

Additionally, on-line desalting (i.e., short C18 column) dramatically improves the applicability of MS. On-line desalting enables a sample to be in ideal storage conditions prior to LC, such as native protein buffer (high salt, buffer, reducing agent, glycerol, etc.) and then during LC, it undergoes on-line buffer exchange into MS compatible solvent. RP-LC has the added sensitivity benefit of separating the mixture over time so that peptides undergoing ESI compete less for charge (i.e., reduced ion suppression), while also concentrating dilute like-peptides together (i.e., chromatographic focusing). On the other hand, IMS enables size separation (i.e., mobility, collisional cross-section) of structural isomers and ions of similar m/z. Use of both separation methods can greatly improve primary sequence coverage in HDX-MS, although 100% coverage is seldom achieved.

Other important considerations regarding instrumentation include the use of LC and valve-based fluid delivery system, particularly for higher throughput. In a conventional experiment, the HDX-labeling step is conducted off-line in a vial, often mediated by an autosampler. The quenched sample is injected onto the immobilized protease column followed by desalting on a short reverse-phase column using aqueous solvent. In a separate phase, mediated by a valve turn and binary solvent flow (aqueous and organic), the desalted peptides are separated on a longer reverse-phase column via solvent gradient. Ultimately, if all these steps occurred in-line with each other, without valves, it would compromise LC separation or inactivate the protease. Moreover, this compartmentalization enables each of the columns to be sufficiently flushed and cleaned, which not only improves lifespan of research consumables, but also inherently improves the HDX data by reducing carryover between samples and reducing the likelihood of clogging.

Finally, once peptides undergo ESI, gaseous peptide ions (5-15 residues) are formed, with a net charge at or close to the number of basic sites.[96] When CID is used to identify peptide sequence in MS/MS, the produced fragments must be charged to continue their flight through the MS to the detector. With that in mind, not all proteolytic peptides will carry a net positive charge, and some peptides will not undergo sufficient fragmentation to be identified. Both are limitations of using MS for HDX detection. On the other hand, bottom-up HDX-MS provides a means to extract information-rich segment-averaged data that can be pieced together in the order of the primary sequence, to reconstruct even massive proteins. From acquired mass spectra, segment-averaged deuterium incorporation can be quantified by taking the centroid of a peptide's isotopic envelope at each timepoint, relative to an undeuterated reference (**Fig. 1.15**).

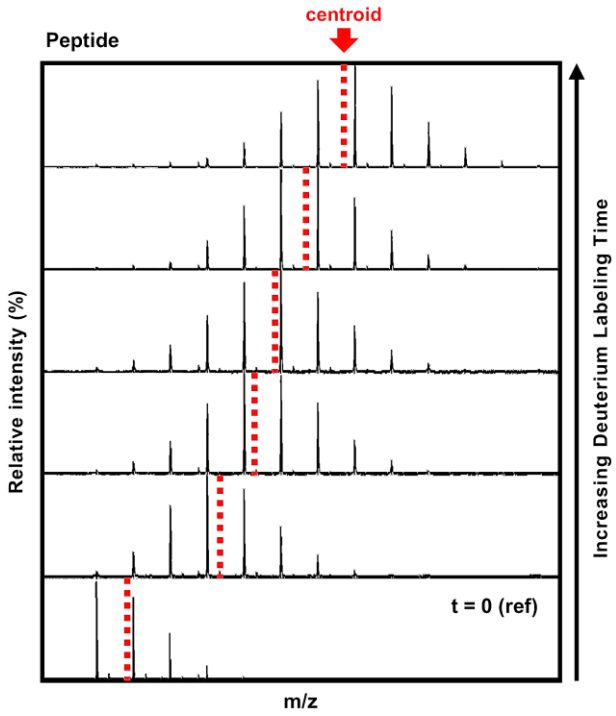


Figure 1.15 | **A peptide's mass spectra during HDX.** The centroid quantifies the average m/z of an isotopic distribution and shifts as more deuterium is incorporated into the peptide.

Since the peptide sequence is known, the number of exchangeable amides *N* can be determined (**Eq. 1.24**) where *T* is the total number of backbone amides, *Pro* is the number of proline (which has no amide hydrogen), and 1 is subtracted to account for the rapid back-exchange of the N-terminus in LC-coupled HDX-MS.

$$N = T - Pro - 1 \quad (1.24)$$

Deuterium uptake can then be expressed as either relative uptake (Da) or relative fractional uptake (RFU, %) (**Eq. 1.25**). RFU or relative uptake may be multiplied by a factor inversely proportional to the D_2O concentration (v/v) during the labeling step (e.g., 50% D_2O in labeling step, multiply **Eq. 1.25** by 2) to account for HDX catalyst dilution.

$$RFU = \frac{D}{N} \times 100\% \quad (1.25)$$

Deuterium uptake can be visualized as a kinetic plot *per peptide* which is ideal for extracting the HDX rate constant k_{obs}. On the other hand, bar plots and heatmaps are often used to aggregate exchange of all the peptides in a data set. Finally, HDX is best interpreted by mapping the peptide exchange on the 3D protein structure, which has become increasingly more available through open-access structural prediction via AlphaFold.[97]

Although HDX-MS is informative regarding the extent of a protein's structure and dynamics (**Fig. 1.16a-c**), differential HDX (ΔHDX) is commonly used to observe perturbations to structure by comparing one or more states (e.g., ligand-bound, mutant, etc.) to a reference protein state (**Fig. 1.16d-f**).

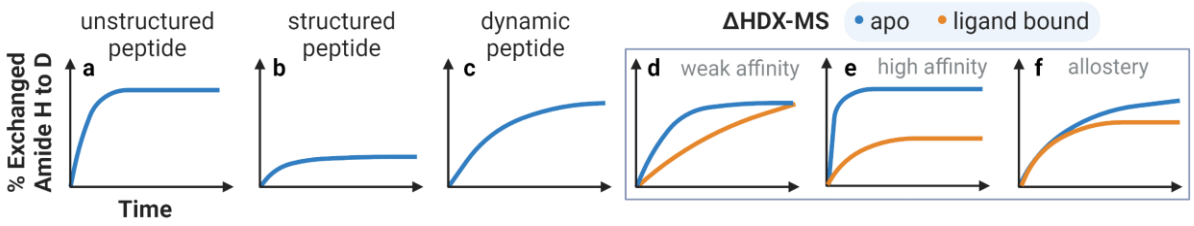


Figure 1.16 | **Examples of HDX kinetic plots.** (a-c) depict rate of uptake observed in disordered, rigid, and dynamic peptides, (d-f) examples of deuteration protection observed by ΔHDX due to weak and high affinity ligands, or subtle allosteric effects.[98] Created using BioRender.com.

If two extremes of affinity are considered (e.g., K_D of 1 nM vs. 50 μM), the low affinity ligand will rapidly cycle through association/dissociation whereas the high affinity ligand will take longer to dissociate. At the binding site, in ΔHDX, this manifests as quickly converging rates (**Fig. 1.16d**) or persistent divergence (**Fig. 1.16e**).[98] Depending on the extent of conformational adjustment to accommodate a ligand, the duration of the unbound state in the binding cycle may sufficiently enable the protein to sample its native/apo confirmational space for HDX labeling. The longer and more abundantly the unbound state/native conformation is populated, the quicker the HDX rates will converge. This same reasoning reveals that, provided adequate labeling time to distinguish between the two states, relatively shorter timepoints are better at capturing weaker interactions.

On the other hand, there are instances in ΔHDX where rates begin to diverge later, on the order of minutes or higher. These kinetics are less intuitive in their interpretation; why would protection accumulate over time in a (near) steady-state system with binding rate constants (k_{on} and k_{off}) on the order of seconds? This behaviour is consistent with the protein population sampling a subtly reduced conformational ensemble that becomes more evident as a deuterium incorporation or time increases. Only after a certain (i) deuteration or (ii) duration threshold has been reached can the change in structure be revealed. In the case of a deuteration threshold, a potential explanation involves the relationship between hydrogen bond 'unzipping' and isotope effects. Since N-D hydrogen bonds can be marginally weaker than N-H, there is a scenario (i) where a certain deuterium labeling threshold increases access to adjacent amide hydrogen for exchange by destabilizing the local hydrogen bond network. Alternatively, in the case of a (ii) particular native conformational excursion which is sampled *extremely rarely* by a peptide/region,

its labeling will be diluted across the ensemble HDX measurement (negligible signal vs. noise). However, as time goes by, more of these subtle excursions are sampled and labeled amongst the population. Then, (i) and (ii) may undergo reduced exchange because of stabilization imparted by ligand binding. If these cases happen at/near the binding site, they are convoluted by other structural events such as intermolecular bonding and interfacing, resulting in ΔHDX plots similar to **Fig. 1.16e**. On the other hand, a case like **Fig. 1.16f** is more likely to be observed in allostery: away from the direct binding site where there are little-to-no other confounding effects.

It is important to note that protein systems exhibit various kinetic profiles, often a blend of those depicted in **Fig. 1.16**. Every peptide's uptake rate is the culmination of various structural forms, breathing motions, or perturbations. This can substantially complicate interpretation and is often aided by corroborating evidence from 3D structures (crystals, cryo-EM, in silico models, NMR) and biophysical assays. Also, HDX-MS is a solution phase, steady-state, averaged measurement of a sample ensemble, which means that the resolution of ΔHDX is largely influenced by occupancy of bound state. Moreover, insufficient timescale coverage may miss distinctions in rate constant for relatively weak affinity binders, leading to false negatives in ΔHDX.[85] Generally, the resolution of ΔHDX is minimized or eliminated in prolonged labeling of fast-exchanging amides (i.e., deuterium saturation) and in overly brief labeling of slowly exchanging amides.[85]

1.3.3 Time-Resolved Hydrogen Deuterium Exchange for Rapid Labeling

Random coil or intrinsically disordered regions/proteins (IDRs, IDPs) represent physiologically important systems which are quite difficult to structurally characterize due to the broad conformational space they sample. Time-dependent, motion-permitting methods such as NMR, HDX-MS, and molecular dynamics are useful in studying structures that would otherwise have weak electron density. In HDX terms, disordered structures with little-to-no amide hydrogen bonding will undergo labeling at or near the intrinsic rate of exchange, which is incredibly fast (i.e., low-to-subsecond).[99] A 2023 mini-review by myself and co-first author, Vimanda Chow, explored the developments of rapid-HDX instrumentation and methodology for studying macromolecular systems with intrinsic rates of exchange.[100] Briefly, to achieve rapid labeling, reaction/mixing chambers are often on the micro-to-nano scale instead of mixing by syringe dispense-aspiration. This is commonly achieved by using small diameter capillaries and microfluidic chips with continuous flow, or by conducting HDX such that an analyte (e.g., gas phase) flies over/through a deuterium source (e.g., reaction gas, evaporating droplet).

Up until recently, there were no feasible options for commercialization of fully automated rapid HDX-MS.[101] Previously, labs interested in employing the method were restricted to in-house built apparatuses.[99] Particularly, a millisecond-to-second HDX-MS approach, termed TRESI-HDX, has been at the core of many Wilson Lab studies. This approach was mediated by a time-resolved capillary mixer[102] with syringe pump fluid delivery, and subsequent on-line proteolysis conducted using a laser-engraved polymethyl methacrylate (PMMA) chip (**Fig. 1.17**).

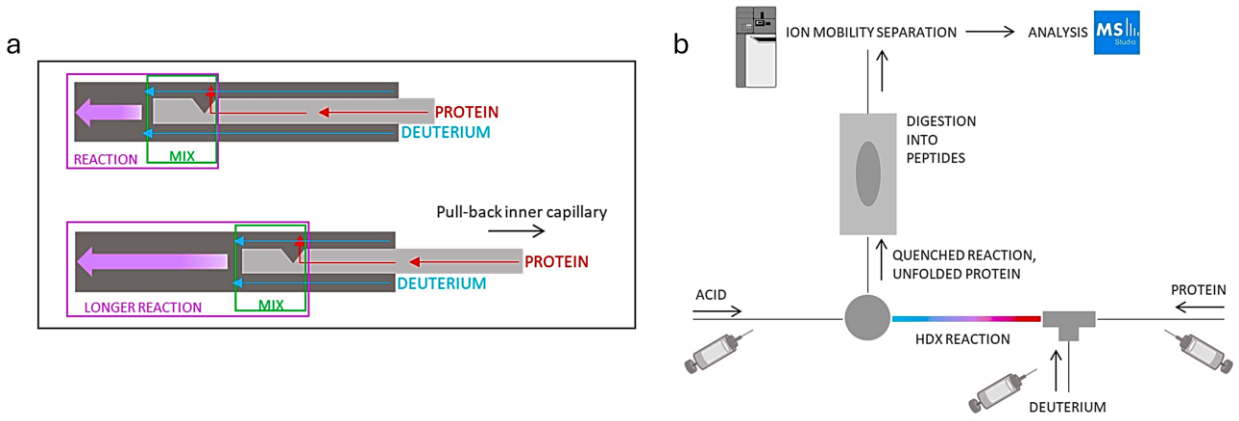


Figure 1.17 | **TRESI-HDX-MS apparatus circa 2022.** (a) time-resolved mixer and (b) on-line HDX and proteolysis system, driven by syringe-based fluid delivery.

The chip went through iterative modifications, across generations of Wilson lab members, including reversible and irreversible annealing (e.g., clamps vs. chemical fixing), as well as variations of the kinetic/HDX-mixer engraved upstream (rather than union-linked) to reduce time, volume, etc.[103–108] Although cost-saving, syringe-based fluid delivery would sometimes leak back against the driving plunger when back-pressure was high enough, such as a clogging event. Due to the incompatibility between syringe-based fluid delivery and reverse-phase chromatography (which typically requires valve- and binary pump-mediated solvent delivery), gas-phase IMS was used to add a dimension of separation to reduce spectral overlap with the added benefit of limiting time for solution-phase back-exchange. Another limitation introduced in the absence of desalting LC was that protein samples had to be buffer exchanged into ammonium acetate salt (NH_4OAc 50-250 mM) rather than storage buffer conditions (e.g., NaCl, Tris, HEPES, DTT, lipids, etc.). Though NH_4OAc is great for ESI, it is not a buffer at physiological pH because the associated conjugate acid (NH_4^+ and CH_3COOH) pK_as are 9.25 and 4.75.[109] This introduced some difficulties in finding TRESI-HDX compatible protein targets to study, which, in instances of incompatibility, would precipitate out in the syringe over the course of the experiment or clog the apparatuses very narrow diameter capillaries. With that said, after conceptualization by generations of various

members of the Wilson Group, TRESI-HDX was retrofitted to a conventional LC system by Anacleto et al. (2023) via autosampler plug injection.[110] The newly afforded higher pressure capacity enabled proteolysis using a short column made from small diameter tubing (<1/16 inch) packed with immobilized protease rather than the PMMA chip, which substantially improved robustness.

1.4 Research Objectives

Given the expense of HDX-MS instrumentation and some expertise associated with conducting experiments and data interpretation, few have direct access to this tool. However, through the power of academic collaboration, it is becoming increasingly applied to address the structural biology questions of complex targets in fast-paced environments, which the following chapters aim to convey.

Chapter 2 explores how two near-superimposable proteins, Bcl-2 and Mcl-1, enable accommodation of shared and distinct binding partners through their different conformational dynamics. This study was, in part, conceptualized by the Bryan Dickinson Lab (University of Chicago)

Chapter 3 delves into how binding affinity information can be extracted during HDX-MS structural characterization of single target ligand libraries. The study follows WDR5, an epigenetic reader protein, and a focused small molecule collection evolved by the Structural Genomics Consortium.

Chapter 4 describes a broad timescale-spanning HDX-MS characterization of binary and ternary complexation of FBXO22 and NSD2-PWWP1 enabled by a PROTAC, UNC10088. This work contributed to a longstanding collaboration between The Structural Genomics Consortium and the lab of Dr. Lindsey Ingerman James (North Carolina University at Chapel Hill).

Chapter 5 is a brief binding characterization study of a novel chemical probe against USP3-ZnF-UBD, spearheaded by members of the Structural Genomics Consortium. It is a showcase of how HDX-MS kinetics can reveal intricacies of a binding interactions.

Chapter 6 is an epitope mapping study of a pan-neutralizing antibody targeting the SARS-CoV-2 Spike Protein, in collaboration with Icosagen, an Estonia-based Contract Research, Development, Manufacturing Organization (CRDMO).

CHAPTER 2

Innate Conformational Dynamics Drive Binding Specificity in Anti-Apoptotic Proteins Mcl-1 and Bcl-2

Esther Wolf[1], Cristina Lento[1], Jinyue Pu[2], Bryan C. Dickinson[2], and Derek J. Wilson[1]

[1]*Department of Chemistry, York University, Toronto, ON, M3J 1P3*

[2]*Department of Chemistry, University of Chicago, Chicago, IL 60637*

This chapter has been published in *Biochemistry* 2023, 62 (11), 1619-1630, 10.1021/acs.biochem.2c00709,

E.W. and D.J.W. contributed equally to writing the manuscript. **E.W.** conducted protein purification, mass spectrometry and data processing. J.P. performed luciferase assay. D.J.W., C.L., and B.C.D. provided supervision.

2.1 Summary

The structurally conserved B-cell lymphoma 2 (Bcl-2) family of proteins function to promote or inhibit apoptosis through an exceedingly complex web of specific, intrafamilial protein–protein interactions. The critical role of these proteins in lymphomas and other cancers has motivated a widespread interest in understanding the molecular mechanisms that drive specificity in Bcl-2 family interactions. However, the high degree of structural similarity among Bcl-2 homologues has made it difficult to rationalize the highly specific (and often divergent) binding behavior exhibited by these proteins using conventional structural arguments. In this work, we use time-resolved hydrogen deuterium exchange mass spectrometry to explore shifts in conformational dynamics associated with binding partner engagement in the Bcl-2 family proteins Bcl-2 and Mcl-1. Using this approach combined with homology modeling, we reveal that Mcl-1 binding is driven by a large-scale shift in conformational dynamics, while Bcl-2 complexation occurs primarily through a classical charge compensation mechanism. This work has implications for understanding the evolution of internally regulated biological systems composed of structurally similar proteins and for the development of drugs targeting Bcl-2 family proteins for promotion of apoptosis in cancer.

2.2 Introduction

The mitochondrial outer membrane permeabilization pathway is an essential mechanism of programmed cell death (apoptosis) and is regulated through a complex web of interactions among Bcl-2 family proteins. The Bcl-2 family encompasses many members that share Bcl-2 homology regions BH1–4: pro-apoptotic multidomain proteins (Bak and Bax), pro-apoptotic BH3-only sensitizers (Bad, Noxa, etc.), pro-apoptotic BH3-only activators (tBid, Bim, etc.), and anti-apoptotic multi-domain proteins (Bcl-2, Mcl-1, Bcl-X_L, etc.).[111]

A brief snapshot of this intricate pathway begins when truncated Bid (tBid) interacts with Bak and Bax to facilitate their oligomerization within the mitochondrial membrane.[112,113] This initiates an irreversible commitment to apoptosis as cytochrome *c* effuses from the intermembrane space, apoptosomes (Apaf-l/cytochrome *c*/caspase-9) are formed, and subsequent caspase cascades lead to cellular destruction.[114] However, anti-apoptotic proteins Bcl-2 and Mcl-1 can sequester tBid from this pathway and prevent it altogether. As such, the “dance towards death,” artfully named by Kale, Osterlund, and Andrews, involves exchanging pro- and anti-apoptotic partners via heterodimerization at a conserved hydrophobic BH3-binding groove.[111]

The BH3-binding groove exists on the surface of multidomain family members including Bak, Bax, Bcl-2, and Mcl-1. Notably, despite only sharing 33% sequence identity in their binding grooves,[115,116] anti-apoptotic members Bcl-2 and Mcl-1 have nearly superimposable 3D structures: PyMOL backbone C_α alignment of Bcl-2 (PDB 1G5M) and Mcl-1 (PDB 2MHS) for 106 core atoms yields an RMSD of 2.275 Å.[117–119] Bcl-2 and Mcl-1 strongly interact with BH3-only proteins tBid and Bim; however, at physiological concentrations, only Mcl-1 can bind Noxa and only Bcl-2 can bind Bad. This poses an intriguing question about what molecular mechanisms can enable such unique binding selectivities in such structurally similar proteins.

Mass spectrometry provides a robust toolbox for elucidating molecular mechanisms including native electrospray ionization mass spectrometry (ESI-MS) to observe ligand binding, complexation, and protein folding behavior; ion mobility separation mass spectrometry (IMS-MS) to analyze protein size/shape; and hydrogen-deuterium exchange (HDX-MS), covalent labeling, and cross-linking (XL-MS) to monitor structural dynamics and binding footprints with sub-molecular structural resolution.[120] In particular, “time-resolved” ESI coupled to HDX-MS (TRESI-HDX), which uses short (ms─tens of seconds) deuterium labeling times, can characterize subtle shifts in conformational dynamics that accompany protein complexation.[121] These data can reveal binding sites and allostery and provide structural/dynamic rationales for functional properties including binding specificity.[122]

In this work, soluble Bcl-2 and Mcl-1 were examined using the mass spectrometry/hydrogen deuterium exchange toolbox to understand the molecular mechanisms that underly their distinct binding selectivities.[118,123] Ion mobility spectrometry (IMS) showed that despite being a smaller protein construct by mass, Mcl-1 (18.05 kDa) exhibited a longer drift time than Bcl-2 (20.25 kDa), agreeing with previous classical structural studies indicating that Mcl-1 has a broader binding pocket compared to other Bcl-2 family members.[119,124] Binding studies using time-resolved HDX revealed that Bcl-2 and Mcl-1 undergo distinctive shifts in their conformational ensembles that are unique to each protein and occur regardless of whether they are interacting with a partner that binds one protein or both. From this, we conclude that binding specificity in these Bcl-2 family proteins is driven not only by the charge compensation mechanisms proposed previously but also, in some cases (e.g., Mcl-1), by the ability of the ligand to induce a thermodynamically favorable shift in conformational dynamics upon complexation.

2.3 Materials and Methods

General. Plasmids encoding Bcl-2 and Mcl-1 were purchased from Biobasic Inc. (Toronto, Canada) and constructed based on solution stable sequences optimized by the labs of Fesik and Walensky, respectively.[118,123] Briefly, Bcl-2 was expressed in BL21(DE3) at 16 °C for 18 h using pET28a (1–34 Bcl-2, 35–50 Bcl-XL, 93–207 Bcl-2) with a C-terminal 6His tag and purified using Ni^{2+} IMAC (GE Healthcare, Fastflow). Similarly, Mcl-1 was expressed in BL21(DE3) at 16 °C for 18 h using pGEX-4T.1 encoding Mcl-1 (172–327) with an N-terminal GST-thrombin tag. GST tagged Mcl-1 was purified using GST-affinity chromatography (glutathione-sepharose resin by GE Healthcare, Fastflow). Subsequent GST-tag cleavage was carried out under rotation overnight at 4 °C (thrombin from bovine plasma, Sigma-Aldrich) followed by secondary GST-purification to remove free GST tag (untagged Mcl-1 collected in flow-through). Both Bcl-2-6His and untagged Mcl-1 were gently concentrated/buffer exchanged into a solution containing 50 mM Tris, 150 mM NaCl, and 1 mM EDTA (pH 7.0) by spin-sized centrifugal filtration (Amicon, 10 kDa MWCO) at 1200*g*, 4 °C, 12 min cycles, and resuspended by pipetting up/down between each cycle. For long-term storage at −80 °C, glycerol was added to a final concentration of 25% and the sample was aliquoted such that no tube was ever thawed more than once. Protein identity was verified by 15% SDS-PAGE and intact-MS.

BH3 Peptides. The following BH3-only peptides were synthesized and purchased from BioBasic Inc (Toronto, Canada): hBID (78–104) SQEDIIRNIARHLAQVGDSMDRSIPPG, hNOXA (19–43) AELEVECATQLRRFGDKLNFRQKLL, hBIM (141–166) DMRPEIWIAQELRRIGDEF-NAYYARR, and hBAD (103–127) NLWAAQRYGRELRRMSDEFVDSFKK. For long-term storage at −80 °C, peptides were resuspended in water and aliquoted such that no tube was thawed more than once.

Native MS. 5 μM Bcl-2 or Mcl-1 was ionized by electrospray into the Waters G2-S Synapt Quadrupole-Ion Mobility Separation-Time of Flight Mass Spectrometer using a modified nanospray stage at a flow rate of 6 μL/min. To prepare for ESI-MS, the protein was buffer exchanged using Slide-A-Lyzer MINI Dialysis Devices (2 mL, 10 kDa MWCO, Thermo Scientific) into HPLC-grade 100 mM NH_4CH_3COO pH 7.

TRESI–HDX. 5 μM Bcl-2/Mcl-1, or 5 μM:30 μM Bcl-2/Mcl-1 to BH3 peptide in 100 mM NH_4CH_3COO, pH 7 (incubated on ice for 1 h), underwent rapid HDX using a time-resolved, adjustable mixer, as discussed previously.[125,126] This enabled time-resolved HDX of 1, 2, 4, and 18 s corresponding to inner-capillary pullback of 2, 5, 10, and 50 mm at protein and 100% D_2O flowrates of 2 μL/min 10% CH_3COOH was injected at 16 μL/min to maintain a constant HDX quenching pH of 2.4 after the reaction and during proteolysis. Pepsin (porcine gastric mucosa,

Sigma-Aldrich) was cross-linked in-house onto NHS-activated agarose (Pierce, Thermo Fisher). The proteolytic chamber was constructed in-house using poly(methyl methacrylate) etched with a CO_2 laser (VersaLaser) and affixed with a 0.2 μm pore-size frit upstream of the ESI emitter. Pepsin-generated peptides were identified using Proteome Discoverer 3.0 (Thermo Scientific) after LC–MS/MS analysis with the Orbitrap Elite hybrid ion trap-Orbitrap mass spectrometer. The deuterium uptake of peptides was analyzed using the G2-S Synapt (Waters) ion mobility separation mass spectrometer and processed by MS Studio.[127] A maximal sequence coverage of 92 and 75% was obtained for Mcl-1 and Bcl-2, respectively. However, due to spectral overlap from digestion products of BH3 peptides, some sequence coverage was lost, resulting in 61% for Bcl2/Bim, 69% for Bcl2/Bid, 62% for Bcl2/Bad, 55% for Bcl2/Noxa, 84% for Mcl1/Bim, 77% Mcl1/Bad, 85% for Mcl1/Bid, and 92% for Mcl1/Noxa. Where possible, Expasy FindPept was used to identify peptides (by MS1) to make up for the loss of redundancy (e.g., 137–150 in Bcl-2/Bad).[128] For a given data set to be accepted, Bradykinin 2–9 (PPGFSPFR, Sigma-Aldrich) was spiked in as an HDX timepoint control peptide. If Bradykinin uptake differences between the bound and unbound samples fell within the error (a statistically insignificant difference), this meant that there was no significant variation between how the two states were prepared and analyzed, and thus, any changes that did occur were accurate.

Homology Modeling with SWISS-MODEL. Homology models were constructed using the AutoModel function in SWISS-MODEL[129] of the following sequences: Bcl-2 Chimera MA-HAGRTGYDNREIVMKYIHYKLSQRGYEWDAGDDVEENRTEAPEGTESEPVVHLTLRQAGDDF SRRYRRDFAEMSSQLHLTPFTARGRFATVVEELFRDGVNWGRIVAFFEFGGVMCVESVNREM SPLVDNIALWMTEYLNRHLHTWIQDNGGWDAFVELYGPSMRHHHHHH; Mcl-1 GSGSDELY-RQSLEIISRYLREQATGAKDTKPMGRSGATSRKALETLRRVGDGVQRNHETAFQGMLRKLDIKN EDDVKSLSRVMIHVFSDGVTNWGRIVTLISFGAFVAKHLKTINQESCIEPLAESITDVLVRTKRDW LVKQRGWDGFVEFFHVEDLEGG; tBid DSESQEDIIRNIARHLAQVGDSMDRSIPPGLV; Bim(EL) AEPADMRPEIWIAQELRRIGDEFNAYYARRVFL; Bad APPNLWAAQRYGRELRRMSDEFVDSF-KKGLP; Noxa ARAPAELEVECATQLRRFGDKLNFRQKLLNLI. The structures were analyzed using PyMOL to identify protein–peptide intermolecular interactions. The quality attributes of these models are listed in **Table 2.1** and are detailed in the publication by Studer et al. (2020).[130]

Table 2.1. SWISS-MODEL Results and Quality

Homology model	**Bcl-2 & Bid**	**Bcl-2 & Bim**	**Bcl-2 & Bad**	**Mcl-1 & Noxa**
Template	PDB 4QVE	PDB 4QVF	PDB 2BZW	PDB 2NLA
Sequence identity (%)	72.68	71.91	64.17	91.26

Global model quality estimate (GMQE)	0.7	0.68	0.73	0.78
QMEAN where 0.0 is native-like	–0.67	–1.58	–1.50	–1.65
Qmeandisco global	0.70 ± 0.06	0.71 ± 0.07	0.72 ± 0.06	0.79 ± 0.07

2.4 Results

2.4.1 Expression and Purification of Solution-Stable Bcl-2 and Mcl-1 Constructs

Wild-type Bcl-2 family proteins are often prone to aggregation and are difficult to express and purify, resulting in a need for truncation or loop deletion to obtain crystallographic or NMR structures. To express solution-stable recombinant Bcl-2 and Mcl-1 for this *in vitro* study, optimized sequences representing the Bcl-2 homology core and purification strategies published by Petros et al. (2001) and Lee et al. (2016) were used.[118,123] Petros et al. (2001) were the first to obtain a solution structure of Bcl-2. They based their solution-stable construct on Bcl-X_L, which had been the first structure solved from the Bcl-2 family in 1997.[131] By replacing the Bcl-2 ~60 amino acid loop with the shorter loop of Bcl-X_L, they were able to reduce disorder and lower the isoelectric point of the protein from near-neutral pH to pH 5.0 while retaining its biological antiapoptotic function upon electroporation into staurosporine-treated Jurkat cells.[118]

Although attempts were made to work with purified wild-type Bcl-2 and Mcl-1, both displayed a propensity for aggregation and precipitation in MS-compatible solutions and subsequent clogging of sub-millimeter diameter capillaries used for delivery of sample to the mass spectrometer. This can be attributed to their transmembrane tails; large, disordered regions; and hydrophobic binding grooves. Consequently, all experiments discussed here use Mcl-1 (172–327) with an N-terminal SGS artifact after thrombin cleavage for GST tag removal and Bcl-2 (1–34, 35–50 Bcl-X_L, 93–207) with a C-terminal His tag. The sequences used can be found in **Table B1** in addition to a split RNA-polymerase luciferase assay testing Bcl-2 and Mcl-1 Bid/Noxa binding activity compared to chimeric Bcl-2 and truncated Mcl-1 in **Fig. B1**.[132,133]

2.4.2 Native Mass Spectrometry Demonstrates the Expected Binding Specificities for Bcl-2 and Mcl-1 Constructs

Previously, a review by Kale, Osterlund, and Andrews (2018) compiled over 30 publications, which reported on the binding affinities of Bcl-2 family members using various biological assays including isothermal titration calorimetry, surface plasmon resonance, and fluorescence polarization.[111,118,134–147] **Table 2.2** is adapted from their work and lists the Bcl-2 and Mcl-1 dissociation constants (K_D) with BH3-only peptides Bim, Bid, Bad, and Noxa. Based on these measurements, Bcl-2 was expected to interact with Bid (50–10,000 nM K_D), Bad (10–50 nM K_D), and Bim (<10 nM K_D), whereas Mcl-1 was expected to interact with Bid (<10 nM K_D), Bim (<10 nM K_D), and Noxa (10–100 nM K_D). To confirm this and obtain the optimal molarity ratio of protein and BH3 peptide for the saturated "bound" protein state, native ESI-MS was used.

Table 2.2. **Dissociation Coefficients Reported for Select BH3 Peptides with Bcl-2 or Mcl-1**

Protein	Peptide (Start-Stop)	Sequence	Length (aa)	Bcl-2 K_D (nM)	Mcl-1 K_D (nM)
hBIM	141-166	DMRPEIWIAQELRRIGDEFNAYYARR	26	<10	<10
hBID	78-104	SQEDIIRNIARHLAQVGDSMDRSIPPG	27	50-10,000 *(20 aa)	<10 *(26 aa)
hNOXA	19-43	AELEVECATQLRRFGDKLNFRQKLL	26	>10,000	10-100
hBAD	103-127	NLWAAQRYGRELRRMSDEFVDSFKK	25	10-50	>10,000

Native mass spectra are shown in **Fig. 2.1**, starting with the unbound Bcl-2 and Mcl-1 spectra in panels A and B, respectively. Both proteins exhibited a similar, narrow charge distribution dominated by the 9+ and 8+ charged state and minimally populated by 10+ and 7+. For Bcl-2, the four *m*/*z* charge peaks correspond to 2013.19, 2236.84, 2516.24, and 2875.60, whereas for Mcl-1, the four peaks correspond to *m*/*z* of 1806.41, 2006.93, 2257.64, and 2579.86. Using ESIprot deconvolution, the average masses of Bcl-2 and Mcl-1 were calculated to be 20,121.69 ± 0.85 and 18,051.88 ± 1.04 Da, respectively (**Fig. 2.2**).[148] Although Mcl-1 matched its theoretical mass (18,053.47 Da) based on its primary sequence, Bcl-2 (20,253.53 Da) was off by −131.84 Da, which could be attributed to N-terminal methionine loss, a hypothesis supported by the N-terminal peptide later observed in HDX-MS experiments (shown in **Fig. 2.3**, peptide 2–13).

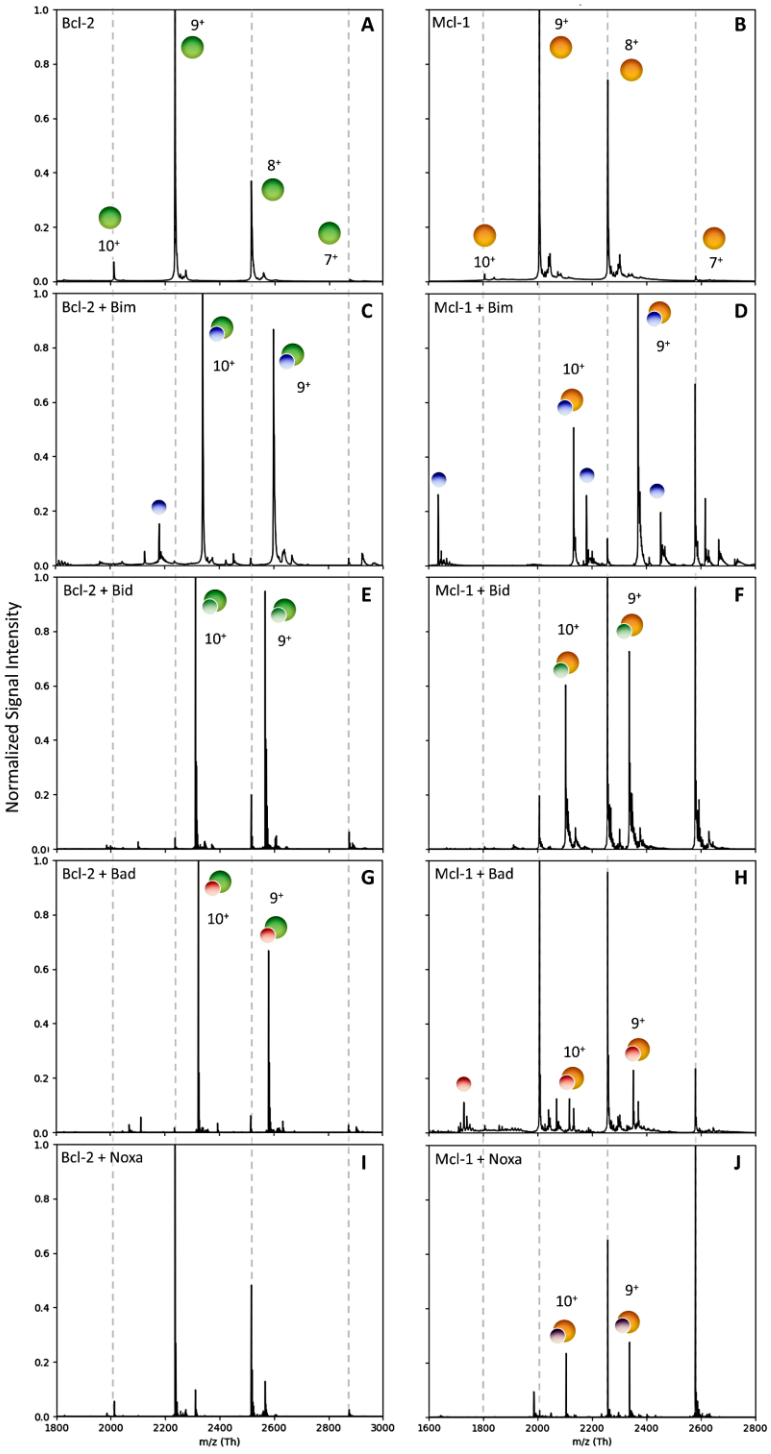


Figure 2.1 | **Native mass spectra of Bcl-2 and Mcl-1 bound to BH3 peptides.** (A) Bcl-2, (B) Mcl-1, (C) Bcl-2 + Bim, (D) Mcl-1 + Bim, (E) Bcl-2 + Bid, (F) Mcl-1 + Bid, (G) Bcl-2 + Bad, (H) Mcl-1 + Noxa, (I) Bcl-2 + Noxa, and (J) Mcl-1 + Bad. Dotted lines track the position of the unbound peak in all mass spectra. The protein concentration was 5 μM and, where applicable, the peptide concentration was 30 μM.

Next, 5 μM Bcl-2 or Mcl-1 was incubated with each BH3 peptide at a 1:1, 1:2, 1:4, and 1:6 molar ratio. As shown in **Fig. 2.1C–H**, a ratio of 1:6 generated saturated "bound" native spectra for known interactors, and this was used as the optimal ratio for all subsequent HDX experiments (**Fig. 2.3**). Dotted lines from **Fig. 2.1AB** indicate the unbound peaks for Bcl-2 and Mcl-1. For Bcl-2 + Bim (23,392.68 Da), the bound m/z peaks correspond to 2340.27 (10+) and 2600.20 (9+); for Bcl-2 + Bad (23,226.26 Da), the bound m/z peaks correspond to 2323.60 (10+) and 2581.74 (9+); and for Bcl-2 + Bid (23,098.53 Da), the bound m/z peaks correspond to 2310.87 (10+) and 2567.50 (9+). The experimental masses of the complexes fell within 2.6 Da of the expected masses of 23,391.38, 23,225.20, and 23,096.01 Da, respectively. On the other hand, for the Mcl-1 + Bim complex (21,324.35 Da), the m/z peaks correspond to 2133.45 (10+) and 2370.37 (9+); for Mcl-1 + Bid (21,029.79 Da), the bound m/z peaks correspond to 2103.96 (10+) and 2337.68 (9+); and for Mcl-1 + Noxa (21,033.33 Da), the bound peaks correspond to 2104.39 (10+) and 2337.99 (9+).

Interestingly, in the Bcl-2 "bound" spectra, peaks corresponding to the complex are dominant, whereas in the Mcl-1 spectra, binding appears to induce a shift to lower charge, even for *m*/*z* peaks corresponding to the unbound protein. This may arise from loss of Mcl-1/BH3 peptide complexes in the gas phase, accompanied by charge stripping by the departing peptide, resulting in a significant fraction of the lower-charge "unbound" peak intensity being attributable to protein that was originally "bound"' in solution.[149,150]

No interaction was detected for Bcl-2 + Noxa (**Fig. 2.1I**), which is consistent with what is known about Bcl-2 specificity (Noxa is an Mcl-1-specific binder). Minor peaks corresponding to Mcl-1 + Bad (a Bcl-2 specific binder) were observed (**Fig. 2.1J**), suggesting the possibility of weak binding. However, Bad did not induce the charge reduction effect observed for all other binders of Mcl-1, indicating that the maximum "bound" fraction can be estimated, based on the intensities of the "bound" peaks relative to the unbound peaks, to be no more than 25%. Note that this estimate excludes the very real possibility that some or all of the observed "bound" peaks arise from non-specific complexation or adduction, which is a common phenomenon in ESI-MS.[151]

2.4.3 Ion Mobility Mass Spectrometry Reveals Divergent Binding Modes of Mcl-1 and Bcl-2 in Binding the Bid BH3 Peptide

IMS-MS was used to examine the gas phase conformations that populated the native ESI-MS spectra of Bcl-2 and Mcl-1. In **Fig. 2.2**, the normalized intensity of signal (ion count) is recorded as a function of the ion mobility drift time in milliseconds, with lower drift times corresponding

generally to smaller globular size (via the collisional cross section) for a given charge state. The 9+ charge state was selected because it is well populated in both the unbound and bound state spectra for both proteins (and comparing structures of the same charge reduces the complexity of interpreting IMS data).

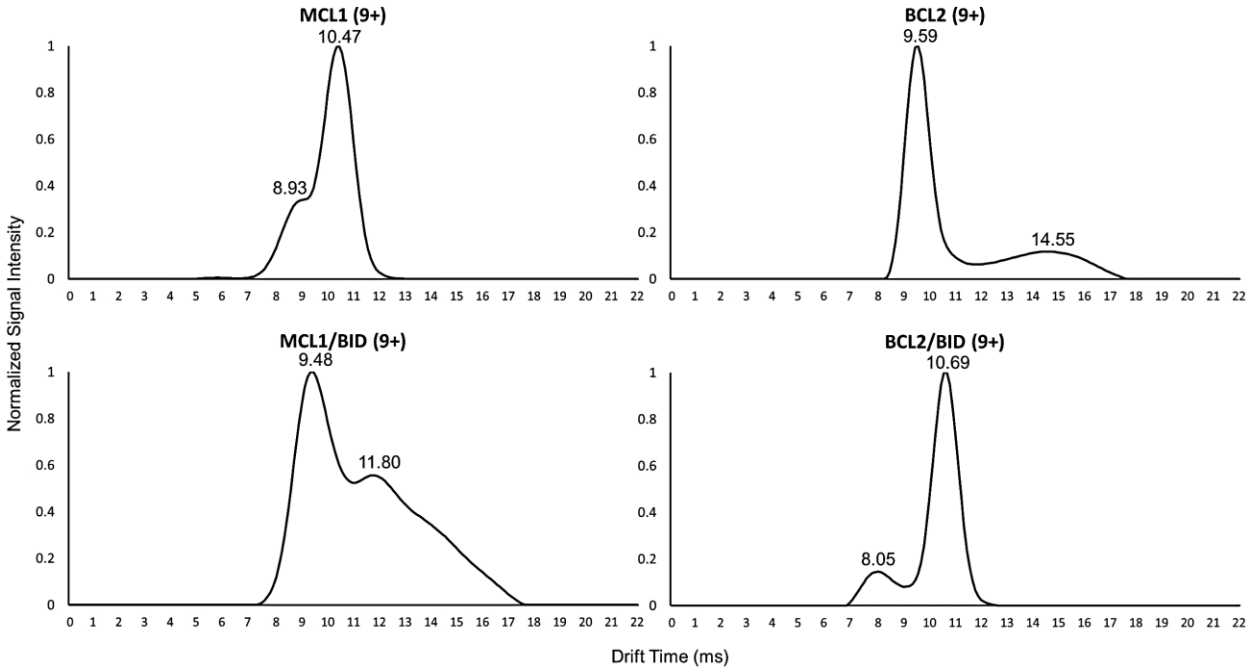


Figure 2.2 | **Ion mobility chromatograms of bound and unbound Bcl-2 and Mcl-1.** The relative signal intensity as a function of drift time (milliseconds) for Mcl-1 (top left), Bcl-2 (top right), Mcl-1/Bid (bottom left), and Bcl-2/Bid (bottom right) is shown for their 9+ charge state.

The dominant drift time peak for unbound Mcl-1 was centered on 10.47 ms, whereas the dominant peak for Bcl-2 was 9.59 ms (**Fig. 2.2**, top row). This is the opposite of what might be expected intuitively, since the Bcl-2 construct used in this study is 15 residues longer and thus somewhat larger than the Mcl-1 construct. However, the impact of this additional sequence on the collisional cross section could easily be subsumed by differences in how the protein is packed overall. In any event, based on the dominant mobility peaks, it appears that Mcl-1 has a somewhat larger cross section than Bcl-2. It is also worth mentioning that a larger peak width is typically linked to variability of structure (cross section) in the gas phase, and the broader larger dominant peak width of Mcl-1 (R_{FWHM} of 2.07 ms) is suggestive of higher conformational heterogeneity compared to Bcl-2 (R_{FWHM} of 1.10 ms). Both proteins also exhibit minor gas phase configurations in the unbound state, suggesting a “compact” structure in the case of Mcl-1 (shoulder at 8.93 ms) and an “extended/disordered” structure for Bcl-2 (broad low peak at 14.55 ms). Additional IMS plots for the 10+ charge state can be found in **Fig. B2**; however, the 8+ charge state was not

present in bound samples (see **Fig. 2.1**) and, therefore, could not be used for bound/unbound comparisons.

Upon complexation with the Bid BH3 peptide, Mcl-1 and Bcl-2 undergo distinct changes in their gas phase configurations. For relatively small proteins like Mcl-1 and Bcl-2, complexation with a large peptide such as Bid BH3 (2976.32 Da) could cause an increase or decrease in the ion mobility, depending on how the incoming peptide packs on to the structure and the extent to which the conformation "tightens" as new inter- and intra-molecular bonds are formed. In this case, complexation appears to have had the opposite effects on Mcl-1 and Bcl-2 drift times, resulting in a "tightening" of the structure for Mcl-1 (10.47 ms → 9.48 ms) and a larger cross section for Bcl-2 (9.59 ms → 10.59 ms). At the same time, the minor peaks have switched places relative to the main peaks, so that bound Mcl-1 is exhibiting a new extended configuration (11.80 ms) and Bcl-2 is exhibiting a new compact state (8.05 ms) upon binding. Additional IMS chromatograms of Bcl-2 and Mcl-1 complexed with BH3 peptides can be found in **Fig. B3** and are consistent with the dominant drift time peaks observed in Bid-bound mobiligrams. IMS of Bcl2/Noxa and Mcl1/Bad were compared to protein-only spectra, shown in **Fig. B4-B5**, as a negative control.

Taken together, these IMS data suggest significantly different binding modes for Mcl-1 and Bcl-2 in their interaction with Bid; however, any characterization of these differences from IMS alone would be, at best, highly speculative, and there is no guarantee that these gas phase observations directly reflect the process in solution. Our next step was, therefore, to undertake a TRESI-HDX analysis, which reflects structure and conformational dynamics in solution.

2.4.4 Time-Resolved HDX-MS Shows Dynamics-Driven Complexation Uniquely in Mcl-1, Regardless of the Binding Partner

To explore whether the IMS observations were consistent with global structure and conformational dynamics in solution, we first carried out "global" hydrogen deuterium exchange measurements for the intact proteins (**Fig. B6**). These data agree with the broad conclusions drawn from the IMS measurements, specifically that Mcl-1 is more dynamic in the unbound state than Bcl-2; however, to examine this system with sufficient spatial resolution to draw specific conclusions about the binding mechanism, we require segment-level data of the type provided by the "bottom–up" TRESI-HDX workflow.

The TRESI-HDX method for bottom–up millisecond hydrogen deuterium exchange has been described previously.[121,125,126] This approach enables "segment-averaged" (peptide-level)

measurements of deuterium uptake with millisecond–second labeling times, which can probe subtle shifts in conformational dynamics resulting from complexation.

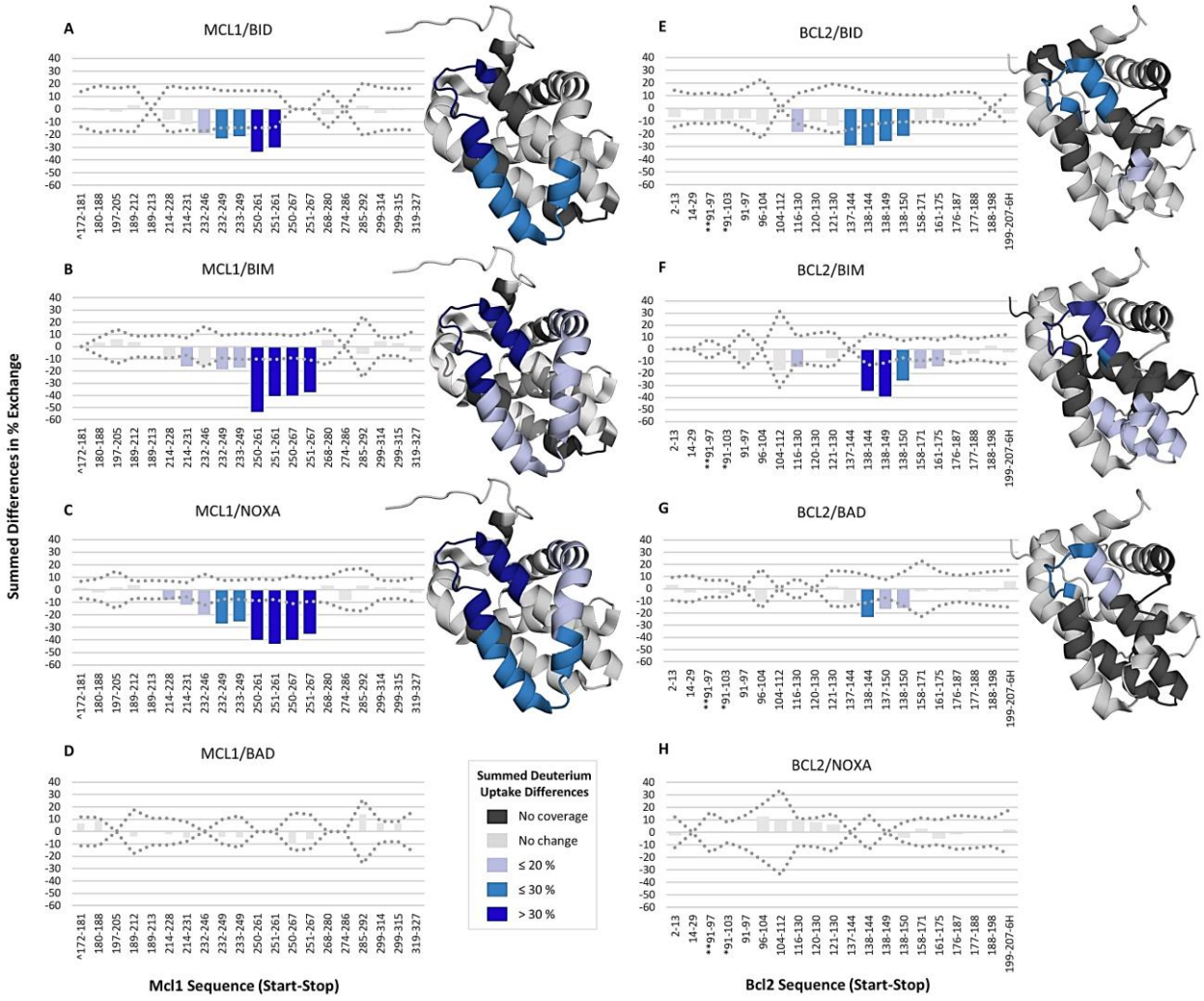


Figure 2.3 | **Summed Difference in % deuterium exchange of complexed versus unbound Bcl-2 and Mcl-1.** HDX difference plots comparing deuterium uptake in Mcl-1 and Bcl-2 in the presence and absence of BH3 peptides from (A,E) Bid, (B,F) Bim, (C,H) Noxa, and (D,G) Bad. The differences in magnitudes are obtained by subtracting raw uptake values from the kinetic plots shown in **Fig. B7–B14**. Colored bars indicate statistically significant decreases in uptake upon complexation with the significance threshold for each peptide (difference magnitude > 2σ, n = 6), represented by the dotted line. To the right of each plot, HDX difference magnitudes are mapped onto the corresponding PDB structures.

In **Fig. 2.3**, the local TRESI-HDX data are presented differentially, meaning that deuterium uptake in unbound protein is subtracted from uptake in the peptide-bound protein, and then the differences at each timepoint were summed. As a result, bars with negative values indicate that

deuterium uptake has decreased in the corresponding region because of complexation. HDX difference profiles were acquired for Bcl-2 and Mcl-1 upon complexation with peptides corresponding to the BH3 peptides of Bim, Bid, Bad, and Noxa.

One of the disadvantages of online TRESI-HDX is that without liquid chromatographic separation, it often provides lower sequence coverage than the corresponding conventional timescale experiment. Sequence coverage varied considerably depending on the system in question; the highest coverage corresponded to Mcl-1/Noxa (92%), and the lowest coverage corresponded to Bcl-2/Noxa (55%), with an average of 73% across the eight difference plots (sequence coverage and redundancy data are provided in **Table B2**, and the peptide list for each protein is provided in **Tables B3** and **B4**). Peptides in Bcl-2 involving the chimeric sequence are indicated using an asterisk (*) to denote 39–50 of Bcl-X_L or a double asterisk (**) to denote 36–50 of Bcl-X_L. In Mcl-1, the N-terminal peptide contains a thrombin cleavage site artifact (an extra S) denoted by a circumflex (^). In every case, peptides located within the binding grooves of Bcl-2 and Mcl-1 were observed.

Time-resolved HDX was used to measure the differences in deuterium uptake between Mcl-1 and Bcl-2 in the presence and absence of BH3 peptides, shown in **Fig. 2.3**. The x-axes for these plots include all peptides detected for the corresponding protein. The dashed line represents the propagated error and using two-fold standard deviation (2σ) of two repeats of three technical replicates per state (n = 6). Bars that did not exceed this magnitude by more than 1% on the "difference in % exchange" scale are colored light gray to indicate that they were not considered statistically significant. Regions with significant changes in deuterium uptake are mapped onto NMR-derived structures of Mcl-1 (PDB 2MHS) and Bcl-2 (PDB 1G5M).[118,119] Light-, sky-, and royal blue were used to illustrate the summed differences of ≤20, ≤30, and >30%, respectively. Regions for which no peptides were obtained are denoted in black, and regions for which there was no significant change are represented in gray.

For Mcl-1/Noxa, large magnitude decreases in deuterium uptake (compared to the unbound state) were observed in peptides spanning 232–267 and late-appearing (statistically significant after 4 s) decreases were observed from 214 to 246. A nearly identical profile was observed for Mcl-1/Bim, as shown in **Fig. 2.3B**. The Mcl-1/Bid profile (**Fig. 2.3A**) also showed substantial decreases at 250–261 and late decreases at 232–249; however, some peptides corresponding to the 250–267 region could not be analyzed due to extensive signal overlap in the raw uptake data (see **Fig. B15** for an example of raw mass spectra for protected peptide 250–261). For all three interactions with Mcl-1 (i.e., Noxa, Bid, and Bim), all significant changes occurred across

the conserved binding groove at helices α3−α5. In the case of Mcl-1/Bad (**Fig. 2.3D**), essentially no significant changes in deuterium uptake were observed across the entire protein, consistent with the expected lack of interaction.

For Bcl-2, the interaction with Bid resulted in changes in a much narrower region corresponding to residues 137–150 (**Fig. 2.3E**), which maps to an area within the Bcl-2 binding groove (see **Fig. B15** for an example of raw mass spectra for protected peptide 137–144). A weak signal is observed in the 116–130 region but is not observed in broadly overlapping peptides 120–130 and 121–130. The signal is, therefore, only mapped from 116-119 in the corresponding NMR structure. Similarly, for Bcl-2/Bim (**Fig. 2.3F**), uptake decreases were observed at 138–150 together with a weak, inconsistent signal at 116–130, and a slowly developing decrease was observed in the 158–175 region. The Bcl-2/Bad complex shown in **Fig. 2.3G** exhibited the weakest decreases in uptake, but these were persistent over all timepoints and occurred in the 138–150 region, in agreement with the other Bcl-2 complexes. Notably, in all three binding scenarios, the persistent decreases occurred at a localized region within the binding groove at the junction of helices α4−α5, corresponding to the BH1 domain that facilitates complexation in Bcl-2 and Mcl-1 via a conserved intermolecular salt-bridge. Bcl-2 did not exhibit any significant changes in deuterium uptake in the presence of Noxa (**Fig. 2.3H**), which is consistent with the expected lack of interaction.

2.4.5 Static Structure-Based Homology Modeling Provides an Incomplete Picture of Complexation that can be Complemented with Dynamics Information from HDX

To further explore the chemical interactions between the human BH3 peptides and Bcl-2 or Mcl-1, we used SWISS-MODEL[129], an open-access protein homology-modeling software package, to visualize 3D structures of Bcl-2 or Mcl-1 with “bound” BH3 peptides for which there are currently no co-crystal or NMR structures (**Fig. 2.4**). In Bcl-2 homology modeling, Bcl-XL was used as the template whereas the Mcl-1 Noxa model was based on a template of Mcl-1 bound to mouse NoxaB.

Overall, Bcl-2 was observed to form a higher number of protein–peptide interactions (i.e., salt-bridges and hydrogen bonds) compared to Mcl-1. However, this general observation provides an example of the limits of a purely static structural analysis of these protein interactions because it offers no basis for specificity differences between Bcl-2 and Mcl-1 and also (incorrectly) suggests that Bcl-2 should bind all BH3 targets tested more tightly than Mcl-1.

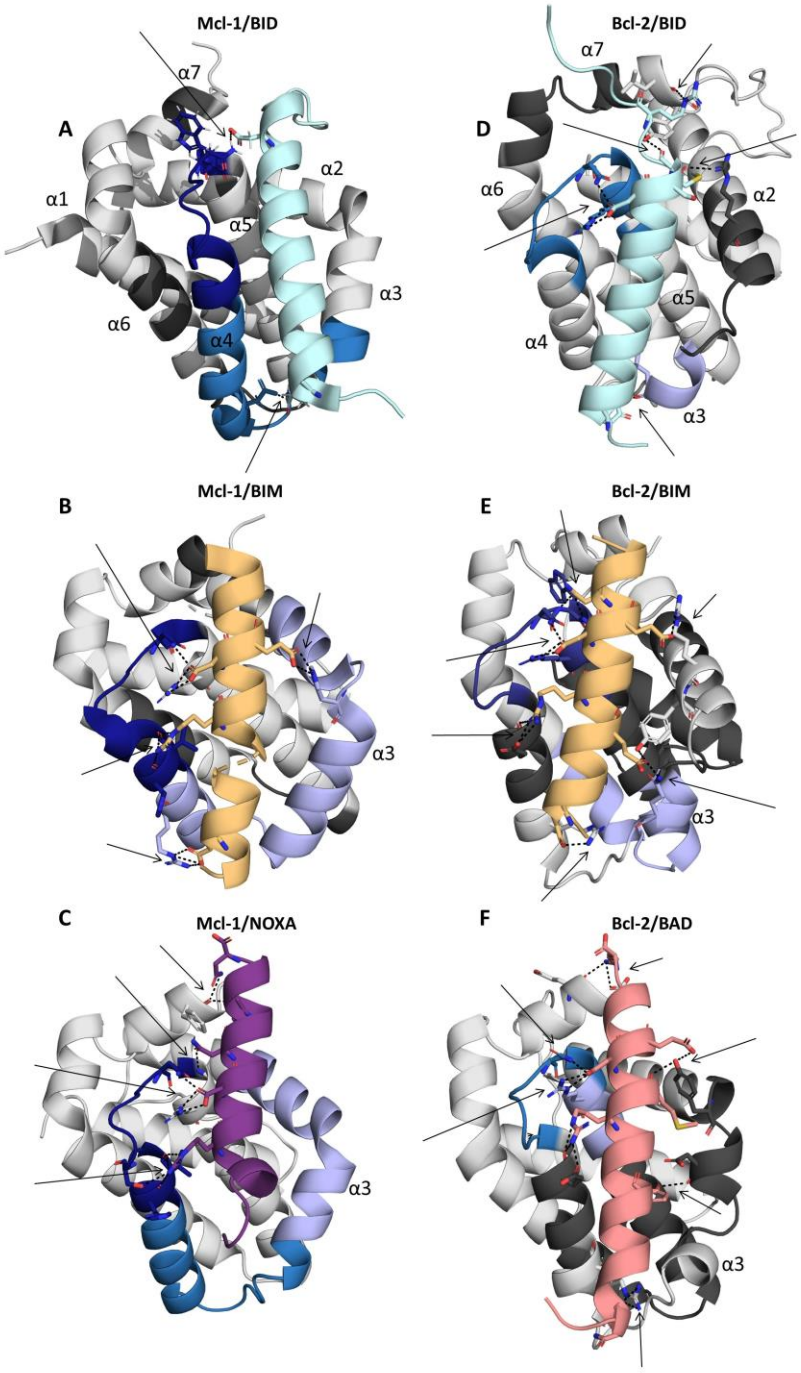


Figure 2.4 | **Homology models and PDB structures displaying intermolecular bonding.** Intermolecular bonds are denoted in black dashed lines and emphasized with black arrows. Mcl-1 is on the left (A–C), and Bcl-2 is on the right (D–F). For the BH3 peptides, Bid is shown in cyan, Bim is shown in yellow, Bad is shown in salmon, and Noxa is shown in violet. The structures are colored based on the summed HDX differences from **Fig. 2.3**: no sequence coverage (black), no change (light grey), ≤20% (light blue), ≤30% (sky blue), and >30% (royal blue).

Nonetheless, the homology modeling approach was able to capture important interactions that were reflected in the HDX data. Specifically, the conserved salt bridge between R146 (Bcl-2) or R263 (Mcl-1) in the BH1 domain at the α4−α5 loop and DBH3 (i+9) was detected in all interactions (see **Table B5** for BH3 residues). In the case of Bcl-2/Bid, salt-bridge and hydrogen bond interactions were observed at the C-termini of α2 (R107), α3 (L119), α4-α5 loop/α5 (N143, R146), and α7 (L201, Y202). For Bcl-2/Bim, interactions were observed at α2 (R107, Y108), α3 (Q118), α4 (R129, E136), and α4−α5 loop/5 (N143, W144, G145, R146), many of which were reflected in the HDX data. For Bcl-2/Bad, there appears to be a restructuring at the α2−α3 region to enable two binding interactions (Y108, D111) in addition to salt bridges at the C- and N-termini of α4 (R129, E136) and α4−α5 loop/5 (R146) and the C-terminus of α7 (E200, G203). These α4 N-terminal and α7 interactions were not detected by HDX. It is not clear if this is a result of a lack of sensitivity of HDX to these particular changes, the presence of a His tag, or an incorrect prediction from the homology modeling approach we are using here.

One unique difference for Mcl-1 interactions predicted by homology modeling, NMR, and crystal structures was that none of the heterodimerizations involved α3. For Mcl-1/Bim (PDB 2PQK), interactions were formed at the α2−α3 loop (H224), as well as the middle and C-terminal region of α4 (R248, H252, V253, N260, and R263). For Mcl-1/Bid (PDB 2KBW), the interaction did not occur at the conserved Arg–Asp salt bridge, but rather with a neighboring residue (N260), as well as a single salt-bridge at the N-terminus of α4 (D242). As for Mcl-1/Noxa, homology modeling revealed interactions at the C-termini of α4 (H252, V253, and S247), α4−α5/α5 (N257, G262, R263), and α7 (F318). An aspect of the experimental data that is supported by the homology models is the observation of an IMS drift-time increase for Bcl-2 and a drift-time decrease for Mcl-1 upon complexation. This is reflected by the fact that in the Bcl-2 models, the BH3 peptide lies “flat” across the binding grove, with little penetration into it. Conversely, in Mcl-1 models, the BH3 peptide is partially enclosed in the binding groove by α3 and α4.

2.5 Discussion

In this work, we have explored the binding of Mcl-1 and Bcl-2 each to three target and one non-target BH3 peptides with the aim of shedding light on how binding specificity can lead to the formation of protein families with a high degree of structural similarity. To emphasize this structural similarity, a backbone-aligned structural overlay of these proteins is shown in **Fig. 2.5**. The basis of heterodimerization between pro- and anti-apoptotic family members has long been thought to be driven primarily by hydrophobic interactions.[111] The BH3 groove in anti-apoptotic members is composed of α3−α5, which has four hydrophobic “pockets,” denoted as P1–P4, and a conserved

Arg residue at in the BH1 domain (α4−α5 loop).[152] The BH3-only proapoptotic members have a conserved Asp (*i* + 9) and four hydrophobic residues: H1 (i), H2 (*i* + 4), H3 (*i* + 7), and H4 (*i* + 11), which line up to form the hydrophobic face of the amphipathic BH3 helix. The conserved Asp and Arg form an intermolecular salt bridge at the BH1 domain of Bcl-2 and Mcl-1 (α4−α5 hinge DGV(T_{Mcl1})NWGR).[153] While these conserved interactions can be used to rationalize the strongly divergent specificity between pro- and anti-apoptotic Bcl-2 proteins, they cannot explain specificity within the anti-apoptotic group of Bcl-2 family proteins like Bcl-2 and Mcl-1, whose specificity profiles are distinct, yet significantly overlapping.

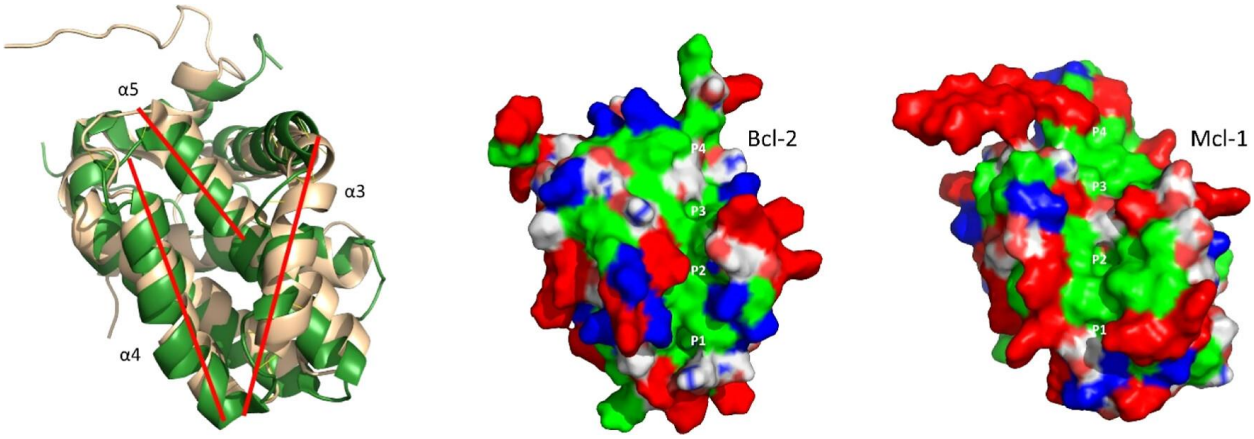


Figure 2.5 | **Backbone aligned Mcl-1 and Bcl-2 structure and electrostatics.** Left: Mcl-1 (yellow) and Bcl-2 (green) share the Bcl-2 homology core, which includes the conserved BH3 binding groove composed of helices α3−α5 (red lines). The solution structures used here were 2MHS (Mcl-1) and 1G5M (Bcl-2).[118,119] Right: Bcl-2 and Mcl-1 color-coded with acidic amino acids in blue, basic amino acids in red, and hydrophobic pockets denoted as P1–P4 in green regions.

The Bcl-2 and Mcl-1 constructs used in this work both consist of 7 α-helices: the predominantly hydrophobic helix α5 makes up the core of both proteins, which is surrounded by amphipathic helices α1–4 and α6–7.[154] Distinctly, the α3 in Bcl-2 is considered a 3_{10}-helix. Native MS revealed that despite their different masses, both recombinant Bcl-2 and Mcl-1 exhibit highly similar, narrow, low charge magnitude charge-state envelopes. This suggests that these proteins were ionized with the broadly comparable structural topology that is conserved within the anti-apoptotic Bcl-2 family members and is consistent with their known structures (**Fig. 2.5**).[118,119,124] A sequence alignment of Bcl-2 and Mcl-1 and their superimposable BH domains is provided in **Fig. B16**, as well as publications where sequence alignment of other Bcl-2 family members and species orthologues can be found.

2.5.1 Gas Phase Measurements Suggest Distinctive Unbound Conformational Ensembles and Divergent Conformational Shifts in Bcl-2 and Mcl-1 upon Complexation

Despite their similarity in NMR structures, IMS data show a distinct difference in drift time between unbound Bcl-2 and Mcl-1, with Mcl-1 exhibiting longer drift times than Bcl-2. This observation is consistent with previous work suggesting that Mcl-1 has a more flexible binding groove than other antiapoptotic members of the Bcl-2 family.[119] The multimodal distributions in the IMS data also point to coexisting conformations in the gas phase that may reflect low-abundance conformational configurations in solution. The presence of these configurations in solution cannot be directly confirmed by global HDX (which is almost always conformationally averaged under “native-like” conditions), but global HDX data nonetheless support the general conclusion that unbound Mcl-1 is more dynamic in solution than unbound Bcl-2.

When in complex with Bid BH3, the dominant configuration of Mcl-1 exhibits a decreased IMS drift time, which would be consistent with substantial penetration of the BH3 peptide into the binding groove coupled with an overall “tightening” of the protein structure (**Fig. 2.2**, left column). However, the dominant IMS peak is accompanied by a broad, substantially higher drift time peak, which indicates that a significant fraction of the Mcl-1 population has undergone a conformational shift, making the protein structure more extended and heterogeneous (or more susceptible to unfolding in the gas phase) upon complexation. While the dominant Mcl-1/Bid IMS peak shift is easily rationalized in the context of complexation, it is not clear what binding effects might generate the minor peak corresponding to a less ordered configuration. Conformational disruption associated with binding is technically possible and has been observed previously with myoglobin[155] and lipocalin.[156]

Conversely, Bcl-2 bound to Bid BH3 has a narrowly distributed, longer drift time peak that is indicative of a larger collisional cross section compared to unbound Bcl-2 (**Fig. 2.2**, right column). This could imply that Bid binding does not involve deep penetration of the BH3 peptide into the Bcl-2 binding groove and does not require substantial rearrangement of the protein structure, which would be consistent with most observations from conventional structural studies[153] and our homology modeling analysis. However, it is as always difficult to draw unambiguous conclusions about processes occurring in solution from IMS data alone, and the appearance of a highly compact minor peak upon binding (**Fig. 2.3**, bottom right) does not fit with this narrative. Despite the inherent ambiguities in interpreting ion mobility data, the IMS analysis clearly suggests that Mcl-1 and Bcl-2 undergo different conformational transitions when binding the same BH3 peptide.

2.5.2 Bcl-2 and Mcl-1 Undergo Substantially Different Conformational Rearrangements upon Complexation

Differential peptide-level HDX unambiguously showed that Bcl-2 and Mcl-1 each exhibit a unique, protein-specific structural and dynamic behavior during binding, regardless of the BH3 peptide that is bound. Mcl-1 undergoes a large, broadly distributed conformational change within the binding groove at α3−α5. To the best of our knowledge, Bcl-2 has never been studied by HDX-MS; however, in agreement with our findings, previously published work by Lee et al. (2016) reported broadly distributed decreases in deuterium uptake for Mcl-1/Bid at a 10 s HDX timepoint.[123] Using LC-HDX-MS, they detected high magnitude decreases at α3, α4, the C-terminal of α2, and the N-terminal of α5, which is a more spatially constrained effect than we observe here, likely a result of the improved sensitivity of rapid HDX measurements to subtle changes in dynamics. In contrast to Mcl-1, Bcl-2 exhibits a decrease in uptake that is localized specifically to the α4−α5 hinge when in complex with Bim and a small decrease at the C-terminal end of α3 for two out of three binders (Bim and Bid).

Comprehensive studies *in vitro* and *in silico* have identified critical residues necessary in BH3 binding for Bcl-2 and Mcl-1: F104, Y108, L119, E136, L137, N143, and R146, and V220, V253, N260, R263, and F319, respectively.[137,157–160] For Bcl-2, most of these critical residues are represented in, or nearby, regions of deuterium decrease upon complexation (**Fig. 2.3**) and salt-bridges/hydrogen bonds observed in homology models (**Fig. 2.4**). Critical residues F104 and Y108 were not observed because they fall within a peptide region that lacks coverage in Bcl2/Bid and Bcl2/Bad (104–112), and a higher propagated error/variation in Bcl2/Bim and Bcl2/Noxa was observed. An *in silico* study by Raghav et al. (2012) found that residues 108–135 of Bcl-2 (α2−α3 loop, α3, and α3−α4 loop) are associated with a “large dynamic motion” and that the flexibility of 108–116 (α2−α3 loop) may enable the plasticity of the BH3 groove.[161] Peptide 116–130 showed weak HDX differences, but they may have been averaged out across this relatively larger peptide (16 residues), and these changes are further supported by salt-bridges in homology models of Bcl-2/Bid and Bcl-2/Bim.

For Mcl-1, the critical residues all fell within regions experiencing reduced deuterium uptake upon complexation and agreed with the BH3-bound homology model, NMR, and crystal structure. Campbell et al. (2015) have suggested that promiscuous binding is enabled by hotspots that are not conserved between Bcl-2 family members: Mcl-1 has consistent hotspots to enable BH3 binding (V220, N260, R263, and F319), but Bcl-2 hotspots are dispersed across the binding pocket.[157] Although this agrees with the homology models of Bcl-2 and Mcl-1, this critical

residue/hotspot model falls short of explaining the decreased deuterium uptake spread across the Mcl-1 BH3 groove and particularly α3 where no hotspots existed. However, studies conducted *in silico* have noted the importance of flexibility in the α3−α4 loop of Mcl-1 to enable adjustment of the α4 helix.[158,159]

Cumulatively, these studies, along with the HDX and homology modeling data, may support a conformational change in Mcl-1 that occurs within α3 to accommodate BH3 peptides, whereas dispersed Bcl-2 hotspots form contacts without the need for extensive conformational adjustment. This is quite contrary to *in silico* observations that Mcl-1 did not undergo backbone conformational changes to accommodate BH3 peptides whereas Bcl-2 type proteins did.[160]

Overall, our TRESI-HDX data indicate that Mcl-1 must undergo a substantial change in structure and/or dynamics to drive heterodimerization, whereas in Bcl-2, binding is driven less by dynamic shifts and more by specific charge–charge interactions at the BH1/BH3 interface as indicated by homology modeling. A cursory explanation based on this observation could be that higher flexibility in the Mcl-1 binding pocket results in “specificity” simply because it allows for the accommodation of larger BH3 ligands. However, the BH3 ligands used in this study are all similar in size. An alternative explanation arises from a recent study on the N-terminal domain of p53 whose binding specificity is modulated by phosphorylation at specific sites.[162] In that case, the authors demonstrated that conformational instability of the phosphorylated state can be a thermodynamic driver of complexation when the phosphorylated and unphosphorylated bound states had a similar conformational stability. Translated to the current system, this rationale would argue that the higher degree of conformational instability in unbound Mcl-1 (as observed in IMS and global HDX), together with the fact that BH3-bound Mcl-1 and Bcl-2 have similar conformational stabilities, provides a driving force for complexation that is unique to Mcl-1. In such a model, Mcl-1 specificity would arise primarily from the extent to which a “potential binder” BH3 domain is able induce the necessary shift in dynamics upon complexation.

2.6 Conclusion

From the perspective of molecular evolution, these results provide a potential hypothesis as to how the specificity of interactions between proteins can be altered through a difficult-to-predict set of mutations. In recent work, it was demonstrated that the evolution of Mcl-1 and Bcl-2 specificity was “path-dependent”, implying that divergent evolutionary pathways generate unique and functionally inequivalent solutions (i.e., sets of mutations) to subsequent evolutionary challenges.[163] The results of the current work reinforce the view that specificity is not always

dictated by specific amino acid substitutions at the interface but can also be driven by widely dispersed mutations that impact protein-wide conformational dynamics. Future efforts examining how alterations in protein dynamics impacted the evolution of functions within Mcl-1 and Bcl-2 orthologues, and specifically whether there is an identifiable “branch-point” ancestor whose evolution diverged into “charge-dominated” and “dynamics dominated” branches, could shed light on how native protein functions emerged, and could also provide new approaches for engineering novel binding specificity profiles into biomolecules.

From the perspective of practical outcomes from this type of analysis, a more complete understanding of binding specificity among structurally similar proteins is a foundation for the development of targeted therapeutics in many “challenging” protein families, including the Bcl-2, GST, LXR, and GPCR families among many others. Our hope is that further explorations of the dynamic drivers of binding specificity will encourage drug development targeting not just structures but function-critical dynamic transitions.

CHAPTER 3

Quantitative HDX-MS for Simultaneous Structural Characterization and Affinity Indexing of Single Target Drug Candidate Libraries

Esther Wolf[1], Oleksandra Herasymenko[2], Maria Kutera[2,3,4], Cristina Lento[1], Cheryl Arrowsmith [2,3,4], Suzanne Ackloo [2], Derek Wilson[1]

[1]Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada.
[2]Structural Genomics Consortium, University of Toronto, Toronto, ON, Canada.
[3]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada.
[4]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

E.W. designed and conducted HDX-MS experiments. O.H. conducted SPR. M.K. purified protein. S.A. selected a representative subset of the WDR5 ligand library. D.W. and C.H.A. supervised the project.

This chapter has been submitted for publication.

3.1 Summary

Hydrogen Deuterium eXchange Mass Spectrometry (HDX-MS) is increasingly used in drug development to locate binding sites and to identify allosteric effects in drug/target interactions. However, the potential of this technique as a quantitative tool remains largely unexplored. Here, a pool of 13 WDR5-targeting small molecules, with Surface Plasmon Resonance (SPR) dissociation coefficients (K_D) ranging from 20 nM to ~116 μM, were characterized using differential HDX-MS (ΔHDX-MS). Qualitative analyses of the ΔHDX-MS revealed that a small number of known contacts observed in co-crystals were uniformly "HDX-silent", and that several compounds induced dynamic shifts outside of the binding interface. We then demonstrated that ΔHDX-MS Signal-to-Noise (S/N) can quantify library-relative affinity and can additionally predict be K_D from a calibration curve. Three methods of S/N calculation are explored, each suitable for a particular scenario. These results demonstrate the potential of ΔHDX-MS use in drug candidate library affinity validation and/or determination while simultaneously characterizing structure.

3.2 Introduction

Small molecule drug development, particularly for protein-protein interaction domains, includes a primary screen of compound libraries to identify putative hits, followed by an orthogonal screen to confirm the hits.[164] SPR is well-suited to primary screening because of its throughput and the depth of information obtained. However, given that the protein target is immobilized, there may be false positives. In a typical candidate library SPR dataset, for example, only a fraction of the fits used to determine K_D kinetically will be high enough quality to provide a high-confidence result. Consequently, drug developers will seek secondary screening as a means to validate affinity.

Protein Hydrogen-Deuterium eXchange Mass Spectrometry (HDX-MS) characterizes the exchange of backbone amide hydrogens with solvent deuterium, typically at peptide-level spatial resolution, and is largely influenced by the extent of hydrogen bonding and solvent access in various regions of the protein.[165–167] Differential HDX-MS (ΔHDX-MS) compares a protein's deuterium uptake in different states, most commonly an 'unbound' versus 'bound' state, to observe perturbations to the protein structural ensemble (e.g., binding). Accordingly, ΔHDX-MS can reveal the binding sites of small molecules, peptides, other proteins, *etc*., on protein targets due to intermolecular contacts and interfacial structural changes.[168] Recent advances in automation, liquid-handling, liquid chromatography, and MS have greatly improved sample

throughput, enabling increasingly widespread application of HDX-MS in earlier stages of drug discovery.[169]

Previously, Marciano *et al.* (2014) and Chalmers *et al.* (2011) compiled reviews emphasizing the utility of HDX-MS in small molecule and biopharmaceutical drug discovery.[170,171] Several groups have linked ΔHDX-MS data to dissociation coefficient (K_D) using titration[172,173], including in the development of the PLIMSTEX approach[174,175], and others have demonstrated the determination of off-rates for protein interactions through sophisticated analysis of deuterium uptake kinetics in ΔHDX-MS experiments.[176,177] Moreover, Hamuro and Coales have shared great insight regarding optimization of HDX-MS for ligand binding studies, particularly when K_D is known and theoretical protection factors can be calculated.[178,179] Notably, Marciano *et al.* (2014) predicted the pharmacological class of small molecules (e.g., agonist, antagonist, etc.) targeting peroxisome proliferator-activated receptor gamma (PPARγ) by correlating the difference in percent exchanged deuterium of helix-12 peptide (SLHPLLQEIYKDLY) to PPARγ transactivation.[170] In our work, we build on this idea by taking a quantitative approach to ΔHDX-MS using Signal-to-Noise (S/N), exploring how it can be defined to optimally suit different scenarios of ΔHDX-MS data set collections of single target ligand libraries. Our study uses a straightforward protein-to-analyte stoichiometry (1:10) and three timepoint ΔHDX-MS workflow to simultaneously monitor structure and determine affinity.

To do this, we conducted a ΔHDX-based hit screening process by characterizing a small molecule library of WDR5 binders. WDR5 is a 36.6 kDa 7-bladed β-propellor protein that notably contains a tryptophan-aspartate (WD40) repeat (WDR). WDR5 has various functions including acting as a scaffold adaptor between Retinoblastoma-binding Protein 5 (RbBP5) and Histone-lysine N-methyltransferase 2A (MLL1) for epigenetic regulation and it is additionally known to be implicated in various cancers such as acute myeloid leukemia.[180–184] Due to this protein's adaptor behavior, two distinct pockets can be targeted with small molecule hits *via* the WIN (WDR5-interacting motif) and less commonly WBM (WDR5 binding motif).[185] Modulating WD40 domains is a growing topic of interest in the computational community, particularly highlighted in work by Schapira *et al.* (2017) as a protein domain which fulfills the requirements of druggability.[186] Thus, we have aggregated a library of small molecules that bind to WDR5 at the Structural Genomics Consortium.[182,187–189] Given this resource and the wealth of 3D structures (>100 deposited to the PDB since 2006)[182,188,190], WDR5 provides an ideal system to explore quantitative ΔHDX-MS for small molecule screening.

We observed that, when compared to co-crystals, ΔHDX-MS correctly revealed the WIN motif binding interfaces and an allosteric pocket while other WDR5 regions were "HDX-silent". Amongst the library, the ΔHDX-S/N offered a semi-quantitative ranking of relative binding affinity, useful in the absence of K_D values. On the other hand, and expectedly,[172,173,191,192] a correlation was observed between the ΔHDX-MS S/N and K_D measured by Surface Plasmon Resonance (SPR). This indicated that once a calibration curve was established, K_D could be predicted. Our focus here is to demonstrate the proportionality between affinity and ΔHDX-MS signal intensity in real-world measurements, which can be used alternatively- (i.e., library-relative ranking where no K_D available), orthogonally- (i.e., validation), or in addition to (i.e., prediction from calibration curve) K_D values obtained from other biophysical methods.

3.3 Materials and Methods

WDR5 Purification, and WDR5-targeting molecules. WDR5 protein and WDR5-targeting molecules were provided by the University of Toronto Structural Genomics Consortium (see **Table C1** and **Fig. C1** for SMILES and structures). All small molecules and peptides were resuspended in 100% DMSO to yield 10 mM stocks. Human WDR5 (2-334), with an N-terminal His-tag and C-terminal Avi-tag (vector pNicBio), was transformed into *E. coli* BL21/Bir A cells and grown in Terrific Broth (Sigma) at 37°C until OD600 ~ 0.8. Expression was induced by 0.5 mM IPTG at 16°C overnight. Cells were pelleted (3000 x g) and resuspended into lysis buffer (50mM Tris pH 8, 500 mM NaCl, 5% glycerol and 0.5 mM TCEP). Cells were sonicated and WDR5 was purified using Ni^{2+} IMAC using elution buffer (200mM Imidazole pH 8) and further purified by SEC. The protein was flash frozen in liquid nitrogen and stored at -80°C in storage buffer (50mM Tris pH 8, 150 mM NaCl, 0.5 mM TCEP). Biotinylation for SPR was assessed using a streptavidin-band shift assay.

Bottom-up HDX-MS. Stocks (WDR5 7.5 μM or WDR5 7.5 μM + ligand 75 μM) were prepared in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5 mM TCEP, 0.75 % DMSO (v/v) and stored at 0.2 °C. HDX labeling was conducted at 0, 15, 60, and 600 seconds in independent, triplicate liquid handling steps. 8 μL of sample stock was mixed with 52 μL Deuteration Buffer (10 mM Phosphate Buffer, 150 mM NaCl, pD 7.5) at 20 °C, resulting in 87% D_2O (v/v) during labeling. After the HDX reaction, 50 μL of the labeled sample was mixed with 50 μL of Quench Buffer (100mM Phosphate Buffer, pH 2.5) at 0.2 °C for 1 minute prior to proteolysis (1:1 Nepenthesin-2 Pepsin, Affipro), desalting (ACQUITY UPLC BEH C18 VanGuard Pre-column, Waters Corp.), and reverse-phase separation (ACQUITY UPLC BEH C18 Column, Waters Corp.). Liquid handling was conducted using a PAL3 Robotic Tool Change system and Chronos Software (Leap Technologies/CTC Analytics). Liquid chromatography was conducted using the ACQUITY UPLC M-Class System

and HDX Manager (Waters Corp.), and ESI-MS was performed using the SELECT SERIES Cyclic IMS (Waters Corp.) in HDMSe mode. Peptide identification and HDX mass spectra were analyzed using ProteinLynx Global Server 3.0.3. and DynamX 3.0 (Waters Corp.), respectively. Data presentation and visualization was done using Excel (Microsoft) and PyMOL 2.5.0 (Schrodinger, LLC).

SPR. Binding affinity of the compounds was tested by SPR using Biacore T200 instruments at 20˚C. Biotinylated human WDR5 (2-334) protein was immobilized on a SA sensor chip using two separate active flow cells, while the remaining active flow cell was charged with a negative control protein, NSD2-PWWP1 (WHSC1-32, Addgene plasmid 101262). After conditioning the flow cells with 50 mM NaOH (3 injections for 60 s with a flow rate of 10 µL/min), each protein solution (0.03 mg/mL) was injected through the flow cells for 30-60s at 5 µL/min reaching an immobilization level of 7000-7100 RU (active) and 4500 RU (reference). Following protein immobilization, 50 µL/min running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% Tween 20 and 2% DMSO) was injected until a stable base line was observed. The tested compounds were injected over the reference and active flow cells using single-cycle kinetics with a five successive injection at 40 µL/min with 60s association and 120s dissociation. For the analysis run, blanks and 50% DMSO wash were included before and after each cycle. The top concentration for each compound was estimated by multiplying 5 times previously obtained K displacement values (data not shown). 3-fold serial dilutions of compound were performed using the running buffer while maintaining the final DMSO concentration of 2% (v/v). Additionally, an eight-point solvent correction was included for each run to adjust high bulk responses from the solvent. Double referencing of the data was performed by subtraction of the reference flow cell and then the respective blank cycles. The data analysis was performed using Biacore Insight evaluation software. Steady-state fitting, in addition to kinetic fitting for compounds 10 (DS0413), 11 (DS0415) and 12 (OICR-9429), were performed by applying 1:1 binding model provided with the software. The final figures were prepared in the GraphPad Prism version 9.4.1.

3.4 Results & Discussion

3.4.1 Qualitative characterization of WDR5-small molecule interactions using HDX-MS

∆HDX-MS was performed for 13 small molecule binders of WDR5 with SPR K_Ds ranging from 20 nM to ~116 µM (**Table 3.1, Fig. 3.1**). To suit high-throughput needs, protein was immobilized for SPR and K_D was determined mainly through the steady-state response for compounds that showed fast turnover and weak-moderate binding affinity. Meanwhile, kinetic analysis was applied

to estimate the K_Ds for compounds with nanomolar binding affinity and slow association and dissociation profile (**Fig. C2**). This library comprised of ligands developed from structure-guided optimization of the WDR5 chemical probe OICR-9429.[182,189] The OICR-9429 chemotype was originally discovered using a WIN-peptide displacement assay by Senisterra *et al.* (2013), and therefore the related small molecules were expected to share the same one-site binding mode.[189]

Table 3.1. **List of WDR5 small molecule binders and their respective K_D measured by SPR**.

#	ChemiReg ID / Common Name	SPR K_D (μM)
1	XS060584a/**DS0227**	10±0
2	XS060479a/**DS0233**	1±0
3	XS060480a/**DS0234**	~116±7
4	XS068123a/**DS0271**	77±2
5	XS068124a/**DS0273**	5±0
6	XS068128a/**DS0278**	2±0
7	XS083358a/**DS0324**	22±2
8	XS083452b/**DS0334**	60±0
9	XS083453a/**DS0335**	37±1
10	MT000951b/**DS0413**	0.02±0.001
11	MT000953b/**DS0415**	0.03±0.002
12	MT005481d/**OICR-9429**	0.02±0.001
13	MT000954a/**OICR-0547**	46±1

The timescale (15 sec, 60 sec, and 600 sec) was chosen for its experimental efficiency as instrument times can quickly become prohibitive for large libraries and labeling times can vary in effectively capturing differing strengths of interactions and shifts in dynamics. Short labeling times (15 and 60 sec) enabled by conventional, automated HDX-MS systems can capture dynamic, labile, weaker affinity binding interactions while moderate-duration labeling of 600 sec (10 min) can catch conformational divergence (e.g., allostery) that becomes pronounced in the HDX over time.[193,194] With efficient scheduling, this enabled our HDX-MS screen to span, cumulatively, under one week (<168 hrs) of experimentation, including data processing, sufficient chromatographic separation, clean blanks between each injection, and triplicate injections per timepoint.

For each ΔHDX-MS profile, a total of 120 WDR5 peptides were analyzed, corresponding to 86.5% coverage of the primary sequence, and 3.62 peptide-per-residue redundancy (**Fig. C3**). Amongst the small molecules tested, OICR-9429 and its inactive analogue OICR-0547 acted as positive and negative controls, respectively.[182]

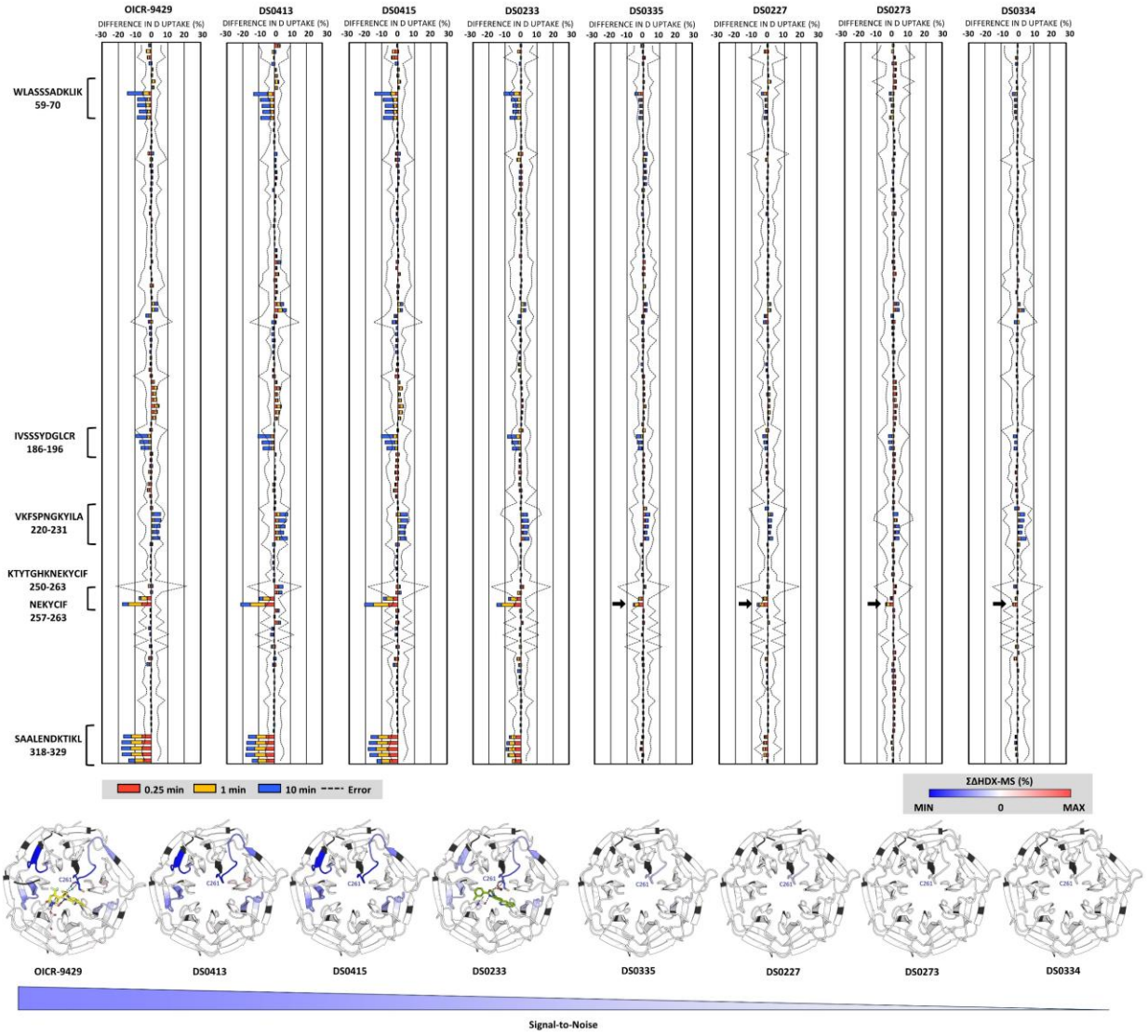


Figure 3.1 | **Summed Differences in Deuterium Uptake of WDR5 Small Molecule Binders.** WDR5 peptide sequences plotted against the extent of D uptake perturbation (%) for each small molecule. To be statistically significant, the summed signal of HDX timepoints 0.25 (red), 1 (yellow), and 10 min (blue) must exceed the summed propagated error (dashed line, 3σ), and show up as shades of blue (ΣΔHDX-MS < 0) or red (ΣΔHDX-MS > 0) in the WDR5 heatmaps. White corresponds to regions which are not statistically significant and therefore exhibit no change (ΣΔHDX-MS = 0), black indicates no peptide sequence coverage. The heatmaps highlight the relationship between relative binding affinity and signal-to-noise. PDB 4QL1 and 5EAL depict OICR-9429 and DS0233 ΣΔHDX-MS, whereas 2GNQ (apo structure) was used for remaining molecules. For clarity, barely statistically significant differences were emphasized with a black arrow in the bar plots.

The difference in deuterium uptake for each molecule was obtained by subtracting deuteration in the unbound state from deuteration in the bound state at each timepoint. Then, the ΔHDX-MS

for each timepoint was summed (ΣΔHDX-MS) and considered statistically significant if it exceeded triple the summed propagated error for that peptide. This approach of summed propagated error enabled clear visualization of the ΔHDX profiles based on a strict significance threshold, both of which are necessary factors for the analysis of larger libraries of ligands (**Fig. 3.1**).

Notably, the positive control, OICR-9429, had the highest magnitude ΔHDX-MS signals, alongside the other two small molecules with sub-micromolar SPR K_Ds of 20-30 nM (DS0415 and DS0413). The remaining small molecules with HDX-detectable binding were DS0227 (K_D of 10 μM), DS0233 (1 μM), DS0273 (5 μM), DS0335 (37 μM), and DS0334 (59 μM). On the other hand, no binding was observed by HDX-MS for the negative control, OICR-0547 (K_D of 46 μM), and four others with a wide range of SPR-measured K_D's: DS0278 (2 μM), DS0324 (22 μM), DS0271 (77 μM), and DS0234 (~116 μM) (**Fig. C4**).

Eight of the 13 small molecules that generated significant differences are shown in the **Fig. 3.1** ΣΔHDX-MS heatmaps (see **Fig. C5** for HDX aligned to detailed peptide list). Of the molecules tested, only OICR-9429 (4QL1) and DS0233 (5EAL) had WDR5-bound structures deposited in the Protein Data Bank (PDB), and so the remaining heatmaps were illustrated on apo WDR5 (2GNQ).[182,195,196] The absence of this handful of co-crystals highlights that HDX-MS can be an excellent orthogonal tool to quickly characterize binding in systems which evade or have yet to undergo crystallization. Importantly, for all molecules where ΔHDX signals were observed, a clear correlation was evident between ΔHDX-MS signal-to-noise and SPR K_D, suggesting that HDX could be used to rank relative binding affinity. This was accentuated by a common set of WDR5 regions that generated ΔHDX-MS responses to small molecule binding: residues 59-69, 186-196, 250-263, and 318-329 experienced decreased dynamics (negative bars) while 220-230 became slightly more flexible (positive bars).

Upon comparison of the ΔHDX-MS plots, it became clear that NEKYCIF (257-263) was the *characteristic* peptide, which we define here as *reliably showing an intense signal upon complexation for all binding-positive small molecules in the library*. This was supported by the WDR5-bound OICR-9429 (4QL1) and DS0233 (5EAL) crystal structures, as Cys261 forms water-mediated hydrogen bonds with both molecules and was likely to similarly involved with the rest of the HDX-detectable small molecules (**Fig. C6**). [182,195] In fact, another structural study noted that the loop bearing C261 in the WIN site is the only significant, contiguous sequence of residues to undergo conformational rearrangement for ligand accommodation.[197]

Based on deuterium uptake kinetics between bound and unbound states, we would have been inclined to categorize 250-263 and 318-329 as regions involved in direct intermolecular contacts to the small molecules given that they both exhibit significant divergence at the earliest timepoint (0.25 min, red bar). However, upon referencing WDR5-bound OICR-9429 and DS0233 crystal structures, we observed that dynamics decreases at 318-329 were in fact a consequence of pronounced allostery further from the WIN site (**Fig. C6**). Given the proximity of 318-329 to the N-terminus of the protein, it is highly probable that WIN site binding stabilized this region within its beta-propellor motif through allostery. Interestingly, point mutagenesis of WDR5 L321A and E322A have shown reduction or elimination of binding to histone H3.1 while not impacting interaction of MLL1.[198–200] L321 and E322 fall within van der Waals range with the threonine-3 methyl group and dimethylated lysine-4 of H3 and the latter makes a water-mediated hydrogen bond via it's methyl-ammonium group to E322.[196,199] HDX-MS suggests that binding at the WIN site allosterically activates region 318-329, which, in the case of H3 tail WIN site binding, can then serve to accommodate the dimethylated portion of H3, facilitating presentation of lysine-4 for di- to tri-methylation conversion.[180,197,201]

Comparison of the WDR5 co-crystals to the HDX also highlighted that some WDR5 residues consistently involved in binding interactions were "HDX-silent". This phenomenon was explored by Scrosati *et al.* (2021) who observed that dynamic events can occur which leave the local backbone hydrogen bond network unchanged and are therefore "HDX-silent", leaving an incomplete picture of the system being studied.[202] In the case of WDR5, to accommodate WIN site binding, the sidechains of Ser91 and Asp107 form hydrogen bonds, Phe133 π stacks, and Phe149 forms part of a small hydrophobic pocket.[198] Mutation of these sites reduced both Histone H3 and MLL1 binding capacity.[198–200] We believe that the HDX-silence of these interactions could be, at least in part, a consequence of the rigidity of the WDR5 structure, which has been observed in the near-identical backbone (main chain) similarity of peptide-bound and unbound WDR5 crystal structures.[197] This agrees with what we observed: during HDX labeling, the WDR5 β-propellor structure likely allowed the sidechains to undergo dynamic changes without a detectable change in amide backbone conformation, manifesting as a lack of differences in amide backbone deuterium uptake from the unbound state. It is possible that modifications to the workflow, including HDX labeling times outside those explored, would capture these missing ΔHDX-MS signals. In the absence of backbone adjustment, changes in amide exchange exclusively due to intermolecular contacts may be captured at shorter timepoints (given the lifetime of a non-covalent bond), but the 15 sec timepoint is at the lower limit for the automated system used. We have also

previously observed HDX-silence in the context of millisecond HDX measurements of the human La protein with UUU-3′OH–containing substrate (U10), in that case by making the comparison to NMR $\Delta\delta$ measurements.[203] On a more general note, a binding event will likely not occur without impacting the overall protein dynamics in some capacity. Therefore, it is extremely unusual for there to be a false negative in ΔHDX-MS as long as a sufficient HDX-labeling time window is sampled, sequence coverage is high, and the affinity is reasonable. Generally, very weak affinities (e.g., K_D >50-100 μM) are indicators of poor drug candidates or nonspecific binding events during affinity measurement and will perform poorly in any structural characterization approach.

The theoretical bound fractions (θ_{eq}) of DS0334 and DS0335, the two molecules with the weakest K_D's to have generated statistically significant, HDX-detectable binding, can be calculated by using the formulas below. For derivation of formula 1, please see the supplementary information (**Formula C1**) and work by Mandell, Falick, and Komives (1998).[191]

$$PL_{eq} = \frac{P_0 + L_0 + K_D \pm \sqrt{(P_0 + L_0 + K_D)^2 - 4P_0 L_0}}{2} \quad (3.1)$$

$$\theta_{eq} = \frac{PL_{eq}}{P_0} \times 100\% \quad (3.2)$$

In this formula, $PL_{eq}$is the ligand-bound protein at equilibrium, K_D is the small molecule's SPR measured dissociation coefficient, and P_0 and L_0 are the initial concentrations of protein and ligand upon deuterium dilution of the sample stock concentration. Given the concentrations of WDR5 (1 μM) and ligands (10 μM) during the HDX labeling step, the minimum bound fractions of DS0334 and DS0335 which were sufficient to detect binding by ΔHDX-MS were 20% and 15%, respectively. This sensitivity to binding is especially exciting given its potential to reduce sample consumption in ΔHDX experiments, where the assumption is usually that binding must be close to saturating. While we would expect similar occupancy sensitivity for most systems, there is likely some degree of protein, binder, and workflow dependence in this parameter, which can only be explored by using this approach for multiple libraries and targets.

Another consideration is that dilution into deuteration buffer pushes the mixed sample to a new equilibrium that is established during the evolution of HDX. The above minimum occupancy calculations are determined using post-dilution concentrations of protein and ligand, which could introduce some error because dilution impact may vary in a ligand-dependent manner. However, the expected sensitivity is maintained by two factors: (i) the 7.5-fold dilution we use does not substantially change the occupancy for ligands with high affinity, and (ii) weak affinity ligands are more likely to undergo fast cycling which allows occupancy to approach equilibrium well within

the first uptake measurement. Having said that, alternative HDX workflows have been described to reduce equilibrium disruption from dilution (*e.g.*, ligand can be dissolved in the deuterium buffer, smaller dilutions can be used etc.)[179]; however, these approaches have drawbacks that can impact sample consumption and sensitivity.

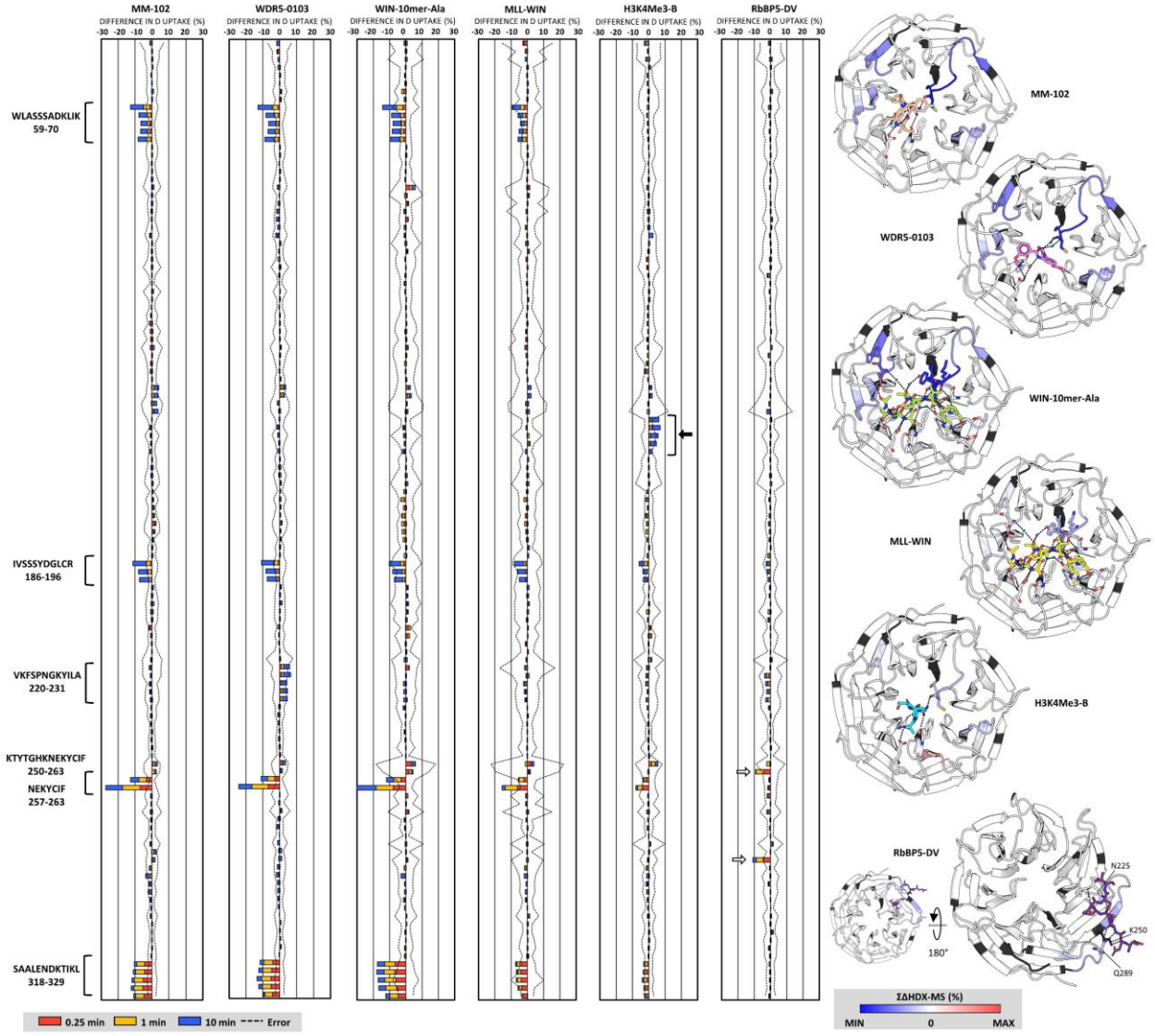


Figure 3.2 | **Summed Differences in Deuterium Uptake of Additional WDR5-Targeting Molecules and Peptides.** WDR5 peptide sequences plotted against the extent of D uptake perturbation (%) for each binder. The HDX timepoints are 0.25 (red), 1 (yellow), and 10 min (blue). The black arrow indicates WDR5 region LIVSGSFDESVR (143-154) and white arrows indicate peptides WDYSKGKCL (241-249) and IWNLQTKE (285-292). On the heatmaps, blue (ΣΔHDX-MS < 0) or red (ΣΔHDX-MS > 0) correspond to statistically significant summed signals which are those that exceed the summed propagated error (dashed line, 3σ) in the plot. White regions represent no change in deuterium exchange between the bound and unbound state

(ΣΔHDX-MS = 0), black indicates no sequence coverage. The PDBs used to depict ΣΔHDX-MS were 4GM8 (MM-102), 3UR4 (WDR5-0103), 3EG6 (WIN-10mer-Ala and MLL-WIN), 2H9P (H3K4Me3-B), and 3P4F (RbBP5-DV).

Having characterized binding in a set of molecules with known K_Ds, we proceeded to examine the binding of additional WDR5 targeting molecules for which K_D was unknown (**Fig. 3.2**). To do this, we tested two commercially available small molecules, MM-102 and WDR5-0103, and four peptides MLL-Win, WIN-10mer-Ala, H3K4Me3-B, and RbBP5-DV (**Table C1**). To illustrate HDX-MS heatmaps, the structures of human WDR5-bound MM-102 (4GM8), WDR-0103 (3UR4), MLL1-Win and WIN-10mer-Ala (3EG6), H3K4Me3 (2H9P), and RbBP5 (3P4F) were used.[189,196,204–206]

WIN-10mer-Ala, MLL-WIN, MM-102, and WDR-0103 exhibited HDX-MS behaviour consistent with the WIN binding site they share with the small molecules previously tested. Again, residues 59-69, 186-196, 250-263, and 318-329 experienced decreased dynamics; however, the slight increases previously observed at 220-230 were only observed for WDR5-0103. These increases at 220-230 may be biased toward smaller sized binders, given their absence in the samples containing peptides (≥ 1310.5 Da) and MM-102 (669.8 Da). As noted earlier, residues Ser91, Asp107, Phe133, and Phe149 were HDX-silent despite being implicated in crystal structures and alanine-scanning mutagenesis.[204,189,198,205] Between WIN-10mer-Ala and MLL-WIN, the former induced higher magnitude ΔHDX-MS signals likely due to increased hydrophobicity from its 6-amino hexanoic acid modifications.[207]

Moreover, amongst this group, H3K4Me3-B generated the weakest signal at the characteristic peptide NEKYCIF (257-263), no decreases at 59-69, and, distinctly, increased in dynamics at 143-151. These differences in ΔHDX activity compared to other WIN site binders may be due to the presence of the biotin tag or a consequence of the H3K4Me3-B length (this peptide was more than 10 residues longer than the other peptide probes). Despite H3 and WDR5 sharing many multimeric complexes, their full-length structures have not been shown to directly interface with each other *in vitro*; complexation always appears to be mediated through RbBP5, MLL1, and other histone subunits in 3D structures.[184] Nonetheless, H3K4 peptide WDR5 co-crystals have been solved (2H9P and 2H6Q), previous studies have shown that mono- and di-methylated H3K4 peptides outcompete MLL peptides for the WIN site, and notably, WDR5 was first discovered as a Histone H3 associated protein.[180,196–198] In any event, these ΔHDX data add to the evidence that there may be direct interactions between histone H3 and WDR5 that remain undetected.

On the other hand, RbBP5-DV bound WDR5 did not undergo significant ΔHDX changes at the WIN site characteristic peptide, rather its impact was strongest at WDYSKGKCL (241-249) and IWNLQTKE (285-292). This is consistent with RbBP5 interacting with a different known binding site, the WBM region.[183,206] The WDR5 peptides 241-249 and 285-292 co-localize and exhibit ΔHDX at earlier timepoints (red and yellow, 0.15 and 1 min) which begin to diminish at/by 10 minutes. Thus, RbBP5 peptide binding does not induce persistent conformational effects at the timescale sampled, suggestive of weak or high-turnover interaction. Moreover, the only co-crystal structure (3P4F) for the *truncated* RbBP5 peptide 373-380 (EEVDVTSV) implicates WDR5 residues Asn225 (with RbBP5 D376, V377, and S379), Lys250 (with RbBP5 E371), and Gln289 (with RbBP5 E374 and V375).[206] Again it appears that we identify an HDX-silent region about N225. Furthermore, in full-length Cryo-EM structures of the nucleosome complex (e.g., 7MBM, 7MBN, 8DU4), the region encompassing the RbBP5 peptide is not resolved, while RbBP5 makes additional contacts to WDR5 *via* regions 54-59 and 327-335.[184,208] Other full-length complexes (6KIV and 6KIW) clearly exhibit noncontiguous regions of contact for RbBP5 to WDR5, implicating both the peptide region we studied and those observed in other complexes.[183] This supports the apparent weak ΔHDX signals observed for peptide RbBP5-DV given its lack of electron density in some structures and the additional RbBP5 contacts made in physiologically relevant contexts.

3.4.2 Quantitative K_D determinations from ΔHDX data

As demonstrated previously, there is an intuitive, theoretical link between the magnitude of the ΔHDX signal and ligand affinity (i.e., K_D).[173,179,191] Binding interactions that dissociate relatively quickly (i.e., higher magnitude K_Ds) would be consistent with a protein conformational ensemble that has not changed much from its unbound state and would manifest as minimal ΔHDX-MS signal.[173] Thus, holding all other experimental variables constant (e.g., concentration, temperature, buffer, etc.), ΔHDX-MS signal intensity will correlate to the affinity of a set of ligands. To translate this relationship to real-world measurements, we illustrate three methods of ΔHDX S/N calculation which yield intensity values that can be used to rank affinity or link to K_D using a calibration curve. The suitability of each S/N method is determined by the nature of the candidate library. For example, in the above study, the first 13 ligands were all developed from the same scaffold, and later, new molecules and peptides were introduced, including one targeting a completely different binding site. Here we outline three S/N methods to cover possibilities of consistent binding mode, distinct ligand size, and distinct ligand binding site.

We began our exploration with a Cumulative S/N ($S/N_{cumulative}$) approach, best applied to libraries of structurally similar ligands with a similar binding mode and footprint, such as small molecules evolved from the same scaffold to target the same protein site. $S/N_{cumulative}$ is outlined by the formula below:

$$\frac{S}{N_{Cumulative}} = \sum_{p=a}^{b}\left(\sum_{t=i}^{f}|S_t|\right)_p \div \sum_{p=a}^{b}\left(\sum_{t=i}^{f}N_t\right)_p \qquad (3.3)$$

where S_t is the sum of differences in deuterium uptake (relative fractional uptake %, as per DynamX 3.0) across all timepoints evaluated (*t*, initial *i* to final *f*), and is then summed across all peptides *p* in the sample (*p*, initial *a* to final *b*). Similarly, the sum is taken of propagated error *N* across all timepoints evaluated (*t*, initial *i* to final *f*), and all peptides *p* in the sample (*p*, initial *a* to final *b*). Triple the propagated error was calculated (see supplementary Formula 2), obtained from a technical triplicate (n=3). The output of this calculation may be less than 1 because the majority of the ΔHDX signals in these peptide lists are well below error (i.e., do not meet or exceed the 3-sigma dashed line in **Fig. 3.1-3.2**).

Using $S/N_{cumulative}$, the affinity ranking for the 19 ligands is (highest to lowest): WDR-0103 (0.67) > MM-102 (0.64) > DS0413 (0.60) > DS0415 (0.51) > OICR-9429 (0.50) > Win-10mer-Ala = DS0233 (0.44) > H3K4Me3-B (0.32) > MLL-Win (0.22) > DS0335 = RbBP5-DV (0.20) > DS0227 (0.18) > DS0273 = DS0278 (0.17) > DS0334 (0.16) > DS0324 = OICR-0547 (0.11) > DS0234 (0.10) > DS0271 (0.09). The bar plots in figure 2 can be referenced to visualize the S/N cut-off for HDX-detectability, which falls around DS0334 (0.16).

$S/N_{cumulative}$ magnitude will be impacted significantly by peptide redundancy and/or sequence coverage of a region undergoing perturbation. To prevent erroneous application of this approach, it is imperative that the ΔHDX-MS signal and noise are calculated from identical peptide lists for each ligand (i.e., composed of the exact same number and identity of peptides). In the event that the peptide lists were to differ between samples, S/N based on average S and N could be used; however, S/N differences between samples would be less distinct as the signals would be diluted during averaging, particularly for larger peptide lists.

Alternatively, similar to the Griffin lab's approach[170], the S/N of a single, identical *characteristic peptide* in each data set could be compared and this would effectively normalize the S/N to a constant across all the aggregated data sets. Here, we take the quotient of the sum of absolute ΔHDX signal S_t and the sum of noise N_t across all timepoints evaluated (*t*, initial *i* to final *f*) of the *characteristic peptide*.

$$\frac{S}{N}_{Characteristic} = \left(\sum_{t=i}^{f} |S_t|\right) \div \left(\sum_{t=i}^{f} N_t\right) \qquad (3.4)$$

This approach can be particularly useful for studying libraries with diverse classes of drug candidates that target the same binding site but with potentially different interfacial areas. For example, in a mixed library of peptides and small molecules, peptides might engage across a larger interface, causing their total summed ΔHDX signals to be greater than that of a small molecule with a similar K_D. This would make the cumulative S/N approach unsuitable for such a library, but analysis using a *characteristic peptide* would remain unaffected.

The following affinity ranking of the 19 ligands was obtained using Characteristic S/N (highest to lowest): MM-102 (9.57) > Win-10mer-Ala (8.13) > WDR5-0103 (7.71) > OICR-9429 (7.22) > DS0415 (6.24) > DS0413 (4.49) > DS0233 (4.34) > MLL-Win (3.95) > DS0335 (2.29) > H3K4Me3-B (1.99) > DS0277 (1.63) > DS0273 (1.12) > DS0334 (1.01) > S/N = 1 > DS0278 (0.61) > DS0324 (0.53) > RbBP5-DV (0.49) > OICR-0547 (0.30) > DS0234 (0.26) > DS0271 (0.14). The S/N of 1 indicates the statistical significance threshold for the characteristic peptide NEKYCIF (257-263). As an indication of the limitations of this approach, this method incorrectly assigns RbBP5-DV as a non-binder, which instead binds to the WDR5's WBM site in an HDX-detectable manner (Figure 3).

The same formula can be applied using other peptides, such as the allosteric region, SAALENDKTIKL (318-329), to rank signal intensity (**Fig. C7**). However, in this case, the S/N will be strongly influenced by the strength of the allostery. This excludes DS0335, DS0273, and DS0334 (and RbBP5-DV), below the S/N = 1 cut-off, despite statistical significance in other peptides (**Fig. 3.1-3.2**). When this formula is applied, ideally the characteristic peptide chosen is most commonly significant for all ligands, otherwise this threshold should be monitored closely.

Finally, to expand this ranking tool to libraries of molecules which may bind to distinct regions of the protein or the same region with a distinct binding mode, and to therefore avoid false outcomes using the characteristic peptide S/N approach, we explored use of the peptide with Maximum S/N ($\frac{S}{N}_{max}$) in as the 'characteristic peptide'.

$$\frac{S}{N}_{Maximum} = \left[\left(\sum_{t=i}^{f} |S_t|\right) \div \left(\sum_{t=i}^{f} N_t\right)\right] \div \left[\left(\sum_{t=i}^{f} D_t\right) \div \left(\sum_{t=i}^{f} E_t\right)\right] \qquad (3.5)$$

Here, we monitor the peptide with the highest statistically significant difference signal, the *maximum peptide*, and it may differ depending on the ligands. We take the quotient of this peptide's summed absolute difference signal S_t and the summed noise N_t across all timepoints

evaluated (*t*, initial *i* to final *f*). We then divide this by the quotient of the apo state peptide's summed relative fractional deuterium uptake D_t and the sum of triple its standard deviation (3σ), denoted as E_t, across all timepoints evaluated (*t*, initial *i* to final *f*). Since this calculation may involve distinct peptides between the data sets, we normalize using the apo state peptide uptake to account for flexibility that can influence the difference signal magnitude.

The affinity ranking obtained here is (highest to lowest): Win-10mer-Ala (0.78) > OICR-9429 (0.53) > DS0415 (0.46) > MM-102 (0.44) > MLL-Win = DS0233 (0.38) > WDR5-0103 (0.35) > DS0413 (0.33) > H3K4Me3-B (0.28) > RbBP5-DV (0.21) > DS0335 (0.17) > DS0227 (0.14) > DS0273 (0.08) > DS0334 (0.07).

An advantage of this approach is that by definition, it is simpler because it excludes compounds without significant ΔHDX signals. On the other hand, there is a risk that the selected peptide may correspond to, for example, allostery beyond the direct binding interface which will not necessarily reflect K_D of the binder (rather the conformational 'lag' of the protein).[209,210] Direct examination of the HDX kinetics can sometimes provide evidence as to the direct (binding) or indirect (allostery) nature of the ΔHDX signal and this can help ascertain whether Maximum S/N should be used for quantitation.[211] With respect to WDR5 residues 318-329, we have also shown the importance of corroborating the HDX kinetics with other structural methods to discern between allostery and direct binding interactions.

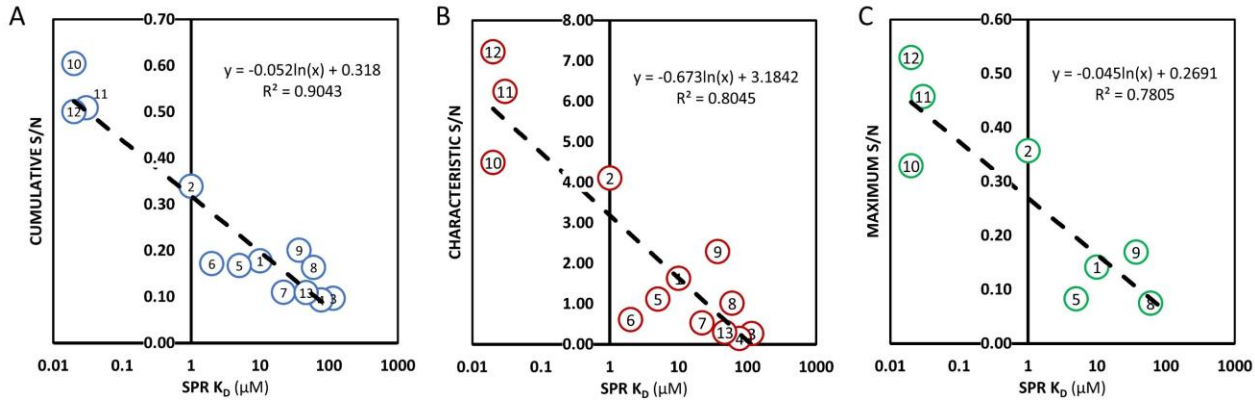


Figure 3.3 | **Correlation of SPR K_D and ΔHDX-MS S/N.** (A) Cumulative S/N (formula 3.3), (B) Characteristic S/N (formula 3.4), and (C) Maximum S/N (formula 3.5). The numbering of each data point corresponds to compounds listed in **Table 3.1**.

The ΔHDX-MS S/N of the first 13 molecules with known SPR K_D values were used to form calibration curves (**Fig. 3.3, Table S1**). Each plot was generated with a logarithmic x-axis for K_D (μM) to yield a linear regression. This calibration curve offers a quantitative means to predict K_D

for the additional WDR5-targeting molecules not yet assessed in-house using SPR. It is necessary to note that our Coefficients of Determination (R^2) are about 0.90, 0.80, and 0.78 for each of the S/N methods, respectively, indicative of variability in the data, and therefore using this line-of-best-fit to extrapolate the K_D will not offer an exact value, but certainly a reasonable estimate. Also, using an alternate peptide, such as one involved in allostery, for the characteristic S/N approach can generate a reasonable correlation (e.g., R^2 = 0.88 in **Fig. C7**), as long as the choice is statistically significant for the majority of the samples tested. As mentioned earlier, this strategy excludes more compounds from the affinity ranking, which, in this example conveniently improves R^2. In general, as more points are added to these any of these lines, and non-specific SPR interactions are ruled out, the accuracy of K_D prediction will undoubtedly increase.

A good example of variability in our regression likely from SPR uncertainty is compound 6 (DS0278), whose ΔHDX S/N was weak despite an SPR K_D (2 μM) indicating substantially higher affinity than most molecules tested. This could point to a poor theoretical fit for K_D determination and/or a confounded measurement during SPR due to, for example, non-specific binding to the chip surface (previously observed by Ruthenburg *et al.* 2006 to occur for WDR5 with immobilized ligands).[197] Indeed, consistent with ΔHDX-MS, extended SPR results (Figure S2) revealed that DS0278 had low WDR5 binding response (27%), which would not be observed in a true high affinity binder. In any event, to ensure the ΔHDX results were rigorously reproducible, DS0278 was rerun using experimentally identical conditions and again did not yield any significant differences in ΔHDX-MS (**Figure C8**). In practice, instances where binding does not induce assay response, such as the absence of thermal shift in Differential Scanning Fluorimetry (DSF) or absence of signal in ΔHDX-MS, will benefit from further corroborative screening. On the other hand, orthogonal validation using ΔHDX-MS S/N can highlight outliers in a focused collection of small molecules while simultaneously characterizing structure.

Finally, using the Cumulative, Characteristic, and Maximum S/N calibration curves, we predicted K_D for novel library additions. Cumulative S/N for WDR5-0103 and MM-102 yielded K_D values of 0.001 and 0.002 μM, respectively. Similarly, Characteristic S/N generated for WDR5-0103 and MM-102 yielded K_D values of 0.0012 and 0.0001 μM, respectively. These S/N values fell outside of the data range of the plots A-C (**Fig. 3.3**); however, they compared moderately well to reported high affinity K_D values of 0.45 μM (measured by isothermal calorimetry, ITC) for WDR5-0103 and an IC_{50} of 0.0024 μM for MM-102.[189,204] Note that K_D prediction for S/N values such as these which fall outside the linear range of the calibration curve should be considered only qualitatively informative rather than accurate. Moreover, K_D measurements are notoriously

variable due to the conditions of a given biophysical method. These results demonstrate that ΔHDX data can provide utility similar to (and in some cases exceeding) SPR methods in the context of hit screening and should prove particularly attractive as an orthogonal method for K_D validation when ΔHDX data have already been acquired for qualitative characterization.

To extend predicting power to peptide binders with larger interfacing at the WIN site, Characteristic S/N can be used, while distinct binding sites can use Maximum S/N. For WIN-10mer-Ala, MLL-Win, H3K4Me3-B, Characteristic S/N predicted K_D's of 0.0064, 0.32, and 5.87. For RbBP5-DV Maximum S/N predicted 3.7 μM. Consulting the literature, although unavailable for WIN-10mer-Ala and RbBP5-DV, our HDX-MS predicted values were very similar to ITC-determined K_Ds obtained for MLL-WIN (1.7 μM) and H3K4Me3 (4.9 μM).[200,205] Other studies have reported various magnitude K_Ds of methylated H3 peptides dependent on concentration and assay; however, these are of substantially shorter length than ours (11 residues) and untagged, with the most similar study using a 14-residue long H3K4Me3 peptide and reporting an SPR-measured K_D of 7.8 μM.[196,197]

3.5 Conclusion

In this work, a WDR5-targeting library was structurally characterized by ΔHDX-MS. The magnitude of ΔHDX-MS intensity was quantified using 3 methods of S/N calculation (cumulative, characteristic, and maximum). The HDX S/N enabled the relative affinity ranking of the library and, when correlated to SPR-measured K_Ds, enabled the estimation of K_Ds for novel library additions. We note that this method was developed using an exemplary model system; WDR5 is well behaved, and most of the library of binders targeted the same binding site with a nearly identical binding mode. For K_D prediction, the calibration curve dynamic range was courtesy of a ligand pool aggregated and evolved over more than a decade. With that said, drug development often follows a similar framework, and as more structural characterization is driven by HDX-MS, S/N calculation is a highly amenable, simple tool to implement for affinity ranking, prediction, validation as the required data is already built-in.

CHAPTER 4

Recruitment of FBXO22 for Targeted Degradation of NSD2

David Y. Nie[1,2,3], John R. Tabor[4], Jianping Li[5,6], Maria Kutera[1,2,3], Jonathan St-Germain[2,3], Ronan P. Hanley[4], **Esther Wolf**[7], Ethan Paulakonis[8], Tristan M.G. Kenney[1,2,3], Shili Duan[1,2], Suman Shrestha[1,2], Dominic D.G. Owens[1], Ailing Pon[1], Magdalena Szewczyk[1], Anthony Joseph Lamberto[5], Michael Menes[5], Fengling Li[1], Dalia Barsyte-Lovejoy[1,9], Nicholas G. Brown[8,10], Anthony M. Barsotti[11], Andrew W. Stamford[11], Jon L. Collins[12], Derek J. Wilson[7], Brian Raught[2,3], Jonathan D. Licht[5], Lindsey I. James[4,10], Cheryl H. Arrowsmith[1,2,3]

[1]Structural Genomics Consortium, University of Toronto, Toronto, ON, Canada

[2]Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada

[3]Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada

[4]Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[5]University of Florida Health Cancer Center, Gainesville, FL, USA

[6]Department of Pharmacology, Physiology, and Cancer Biology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, USA

[7]Department of Chemistry, York University, Toronto, ON, Canada

[8]Department of Pharmacology, University of North Carolina, Chapel Hill, NC, USA

[9]Department of Pharmacology & Toxicology, University of Toronto, Toronto, ON, Canada

[10]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[11]Deerfield Discovery and Development, Deerfield Management, New York, NY, USA.

[12]Office of the Vice Chancellor for Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

A version of a portion of this chapter is *in press* at Nature Chemical Biology

A bioRxiv preprint is available: DOI 10.1101/2023.11.01.564830

E.W. conducted HDX-MS experiments and data analysis. Protein and PROTAC samples were prepared by M.K.

4.1 Summary

Proteasome targeting chimeras (PROTACs) are bifunctional molecules which enable degradation of a target protein via proximity induced ubiquitination. Here, a PROTAC, UNC10088, which targets cancer associated NSD2 for degradation, is revealed to recruit FBXO22, an E3 ligase subunit. They form a ternary complex whose structural characterization is quickly achieved using HDX-MS corroborated by AlphaFold models and crystals.

4.2 Introduction

Nuclear receptor binding SET Domain-containing 2 (NSD2) catalyzes histone 3 lysine-36 dimethylation (H3K36Me2), a transcription activating epigenetic marker, which, when overexpressed, has been associated with various cancers.[212,213] The PWWP1 domain forms an aromatic cage and is used by NSD2 to bind ("read") H3K36Me2 marks, through cation-pi interactions (**Fig. 4.1a**).[214,215] To date, despite efforts (e.g., UNC6934 in **Fig. 4.1b**),[216,217] antagonism of the NSD2-PWWP1 domain has not been successful at inhibiting catalysis of H3K36Me2 *in vivo*, suggesting the need for an alternative approach.

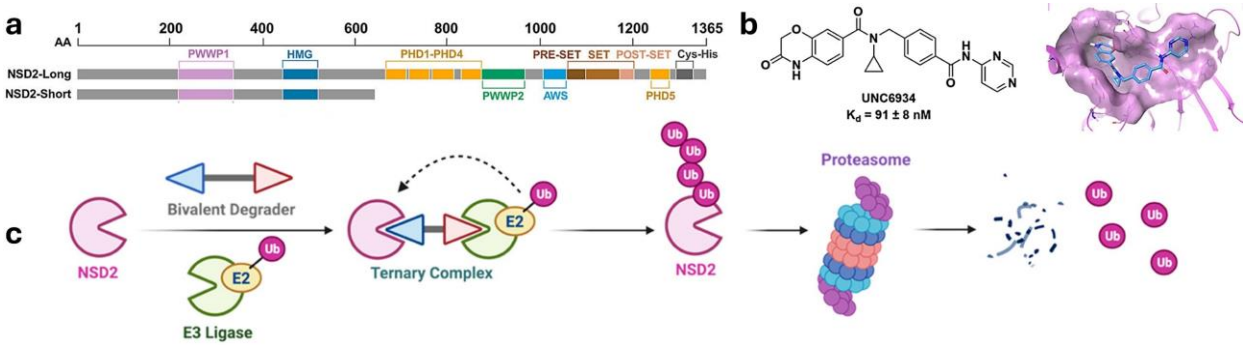


Figure 4.1 | **UNC6934 targets the PWWP1 domain of NSD2 and has potential as a PROTAC.** (a) linearized NSD2 sequence, (b) UNC6934 molecule and PWWP1-bound crystal structure (PDB 6XCG)[217], (c) general schematic of a PROTAC (bivalent degrader) for NSD2 degradation. Adapted with permission from Hanley et al. (2023), copyright © 2023 American Chemical Society.[218]

Proteasome targeting chimeras (PROTACs) are proximity induced pharmacological agents. They are heterobifunctional/chimeric molecules that bind a degradation signal source (e.g., E3 ligase complex) with one end and a degradation target on the other (**Fig. 4.1c**). By directing ubiquitination toward a particular target through ternary complex formation, dose-dependent therapy can be achieved via proteasomal degradation. E3 ligases target proteins for degradation by recognizing destabilizing N-terminal residues (N-recognin/N-degron system).[219] Particularly, Type 1 N-degrons (basic residues: R, K, etc.), have been crystalized with ubiquitin recognition

boxes (UBRs) of the UBR-dependent E3 ligases UBR1 and UBR2, and bind by charge complementarity (i.e., UBR acidic residues).

UNC8153, a PROTAC, was previously discovered by Hanley et al. (2023) to target NSD2 for degradation, correspondingly reducing H3K36Me2 abundance.[218] Bifunctional UNC8153 bound the PWWP1 domain via a chemical moiety derived from UNC6934 (**Fig. 4.2a**), highlighting the utility of PROTACS in repurposing chemical probes that may have failed as drugs. Extended by a 6-carbon linker, at the other terminus of UNC8153, a primary alkylamine served as the ubiquitin ligase binder. In that work, Structure-Activity Relationship (SAR) studies revealed that a primary alkylamine, rather than a basic residue, resulted in the highest affinity and degradation capacity of NSD2. With this in mind, and after observing no impact on NSD2 degradation from UBR-knockdown, UBR-dependent E3 ligases were ruled out as UNC8153-mediated degraders.

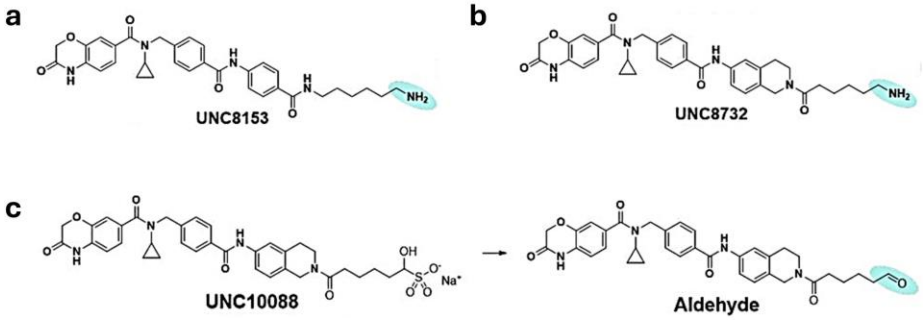


Figure 4.2 | **Evolution of NSD2 degrader.** Chemical structures of (a) UNC8153, (b) UNC8732, (c) UNC10088. UNC10088 undergoes hydrolysis to form an active aldehyde. Reproduced with permission from bioRxiv preprint Nie et al. (2023).[220]

Present work by Nie et al. (2023),[220] explores the development and characterization of a PROTAC, UNC8732, evolved from UNC8153 (**Fig. 4.2b**). Particularly, BioID, a proximity biotinylation identification strategy (~10 nm radius), of NSD2 in the presence of UNC8732 detected F-box only protein 22 (FBXO22) which is the substrate-recognition subunit of the SCF(SKP1-CUL1-F-box)-type E3 ubiquitin ligase complex. Also, upon LC-MS/MS confirmation that the primary alkylamine of UNC8732 was converting to an aldehyde in cells, which enabled the PROTAC to function *in vivo*. UNC10088, an aldehyde-bisulfite adduct (α-hydroxy sulfonate) was synthesized to enable *in vitro* study of the aldehyde moiety by simple hydrolysis conversion (**Fig. 4.2c**). The formation of the FBXO22-UNC10088-NSD2 ternary complex was confirmed by biolayer interferometry (BLI) and Size Exclusion Chromatography (SEC) co-elution, and *in vitro* polyubiquitination of NSD2 was confirmed by gel shift (SDS-PAGE).

This chapter will focus on the HDX-MS characterization of the structural interactions mediating ternary complex formation. Notably, only a predicted AlphaFold model of FBXO22 is available (AF-Q8NEZ5-F1). Co-purification (achieved by M.K.) of FBXO22 with SKP1 (of the SCF complex) was necessary for *in vitro* stabilization of FBXO22. Using HDX-MS, UNC10088 was observed to form binary complexes with both NSD2-PWWP1 and FBXO22, and no interactions with SKP1. The PWWP1/UNC10088 interface was consistent with the 3D structure of PWWP1 complexed with UNC6934 (PDB 6XCG).[217] The previously unknown FBXO22/UNC10088 interface was revealed to be narrowly localized to one face of FBXO22 and to contain Cys326, which was consistent with the possibility of covalent linkage to the activated aldehyde of UNC10088. This was further supported by the abrogated binding to the FBXO22 C326A mutant. Since the signal intensity of C326-bearing peptides was not impacted in HDX-MS, it spoke to the possibility of a *reversible* covalent linkage between FBXO22/UNC10088. HDX-MS of FBXO22/UNC10088/PWWP1 (1:1:1) revealed additional structural changes to FBXO22 and changes in binding dynamics of PWWP1, evidence of ternary complex formation (1:1:1).

4.3 Materials and Methods

HDX-MS. 6 protein samples were prepared in buffer (20mM Tris pH 7.5, 200mM NaCl, 5% glycerol, 1mM TCEP, 2mM Benzamidine): (i) 7.5 µM SKP1/FBXO22, (ii) 7.5 µM SKP1/FBXO22 + 7.5 µM PWWP1-NSD2, (iii) 7.5 µM SKP1/FBXO22 + 7.5 µM PWWP1-NSD2 + 75 µM UNC10088, (iv) 7.5 µM SKP1/FBXO22 + 75 µM UNC10088, (v) 7.5 µM PWWP1-NSD2, (vi) 7.5 µM PWWP1-NSD2 + 75 µM UNC10088. HDX labeling was performed in 87% D2O (10mM Phosphate Buffer pD 7.5, 150 mM NaCl) at 20ºC for 0, 1, 5, 10, 60, 180, 720 mins. Quenching took place for 2 mins using 100 mM Phosphate Buffer at a pH 2.5 at 0 ºC. The quenched sample was loaded onto a mixed Nepenthensin-2:Pepsin (1:1) column (15ºC) and desalted on a short C18 column (0 ºC) at 0.2 mL/min for 3 mins using a mobile phase A (0.1% formic acid in water). The trapped, desalted peptides were then reverse phase separated using a long C18 column at 0.04 mL/min with a 12 min gradient from 35% to 85% of mobile phase B (0.1% formic acid in acetonitrile). The eluting peptides were electrosprayed and analyzed using a single pass on the SELECT SERIES Cyclic IMS (Waters Corp., U.K.) in HDMSe mode. Leu-Enk was used for lock mass (m/z axis calibration). Peptides were identified using ProteinLynx Global Server 3.0.3. (Waters Corp., U.K.). HDX uptake was determined using DynamX 3.0 (Waters Corp., U.K.). Cumulative HDX differences (∆HDX-MS) were calculated using Microsoft Excel and mapped using PyMOL. For a cumulative ΔHDX-MS peptide signal to be considered statistically significant, it must have exceeded the cumulative error (three times propagated error, technical triplicate).

4.4 Results & Discussion

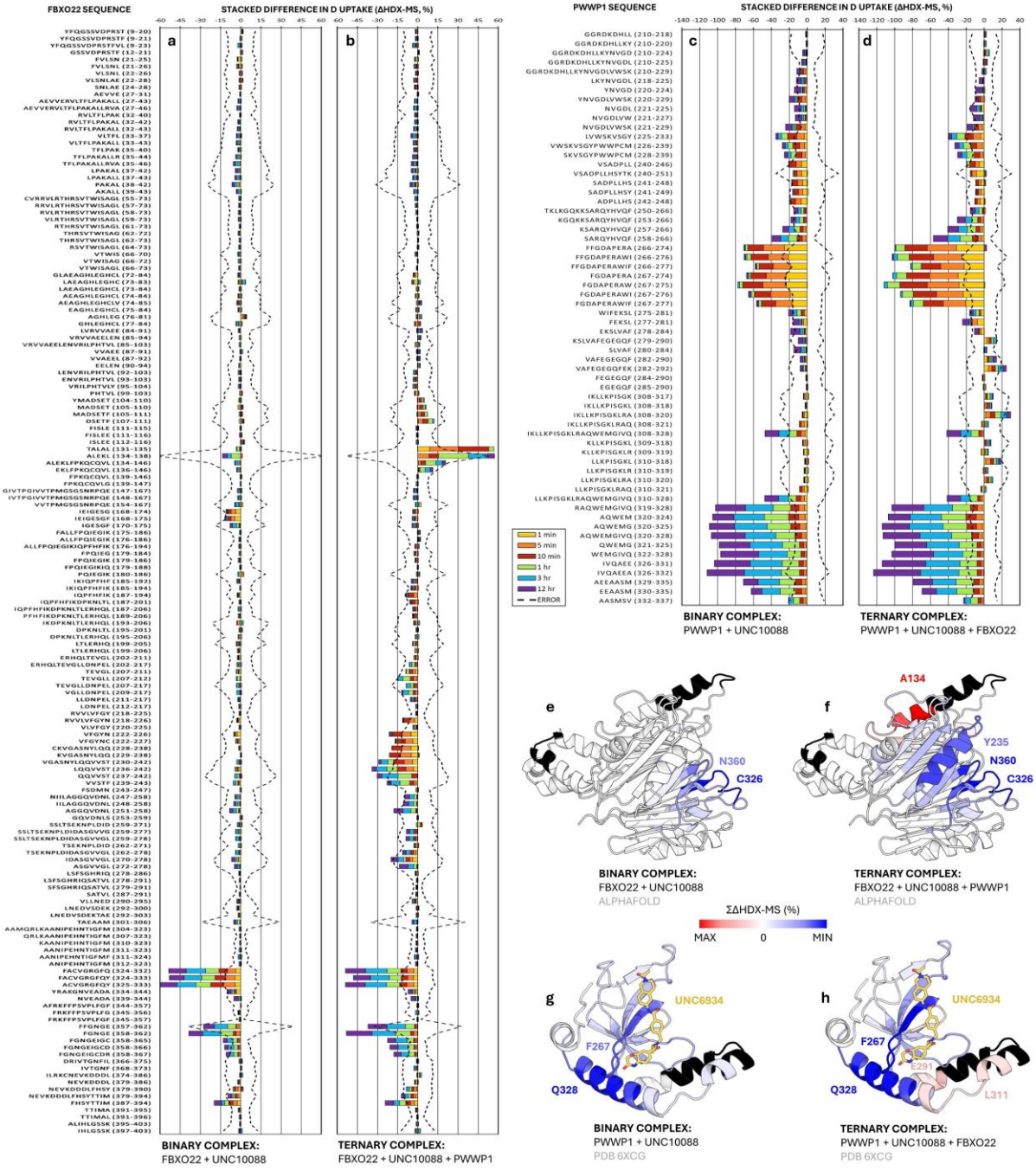


Figure 4.3 | **Structural Characterization of UNC10088 Binary and Ternary Complexes.** ΔHDX-MS of (a) FBXO22 + UNC10088, (b) FBXO22 + UNC10088 + PWWP1, (c) PWWP1 + UNC10088, (d) PWWP1 + UNC10088 + FBXO22 and their cumulative heatmaps (e-h). ΔHDX-MS was performed at 1, 5, 10 mins (yellow, orange, red) and 1, 3, 12 hrs (green, blue, purple). To be considered statistically significant, summed signal must exceed three times the summed propagated error (dashed line). The FBXO22 structure is an AlphaFold model and the NSD2-PWWP1 domain is a crystal structure with UNC6934 (PDB

6XCG). Structure colouration (e-g) corresponds to cumulative ΔHDX-MS decreases (blue) and increases (red), insignificant differences (white), and missing coverage (black).

UNC10088, an *in vitro* active PROTAC which enables NSD2 degradation, now known to be FBXO22-mediated, has been developed through years of collaboration (**Fig. 4.1** & **4.2**).[216–218,220] To structurally characterize the binary and ternary complex formation mediated by UNC10088, FBXO22, and the NDS2-PWWP1 domain, ΔHDX-MS was used, spanning a timescale of 1 min to 12 hrs. A co-purified complex of FBXO22/SKP1 was necessary to study FBXO22 *in vitro*. The sequence coverage of $FBXO22_{13\text{-}403}$, $SKP1_{1\text{-}163}$, and NSD2-$PWWP1_{211\text{-}350}$ was 91%, 96%, and 71%, respectively (**Fig. D1-D3**). As expected, no significant differences in deuterium uptake were observed between SKP1/UNC10088, or FBXO22-SKP1/PWWP1 (**Fig. D4**).

Upon addition of UNC10088 (1:1), FBXO22 exhibited decreases in deuterium uptake from 324-333, 358-367, and 387-394, all localized to one area of the AlphaFold model (**Fig. 4.3a**,**e**). This region makes up the FIST_C sensory domain (299-364), which is the first of its kind to be discovered in the human proteome.[221] Decreased HDX observed from 324-333 appeared significantly at the earliest timepoint (1 min, yellow) and persisted for the entire 12 hr time course, whereas regions 358-367 and 387-394 showed up slowly, significantly captured by the 1 hr timepoint (green) or later. Consideration of these kinetics reveal that residues spanning 324-333 are likely making intermolecular bonds with UNC10088, whereas 358-367 and 387-394 undergo decreased dynamics as a consequence of binding (i.e., “orthosteric effects”). Interestingly, the terminal aldehyde of UNC10088, expected to bind FBXO22, is considered reactive and this identified C326 as a candidate for covalent linkage. ΔHDX-MS of UNC10088 with a C326A mutant revealed total abrogated binding (**Fig. D5**). However, evidence of irreversible covalent linkage in the form of peptide mass shift was not observed in the wildtype and mutant FBXO22/UNC10088 mass spectra. In other words, signal peaks corresponding to peptides spanning 324-333 covalently linked to UNC10088 were not observed, and signal intensity of peptides spanning 324-333 was not reduced. These results were consistent with a *reversible* covalent linkage, such as a hemithioacetal,[222] between C326 thiol and UNC10088 terminal aldehyde. A handful of covalent PROTACs have been reported, and a reversible covalent linkage to FBXO22 may help in ubiquitination efficiency of the degrader without the expense of spatial isolation from the catalytic E3 subunit.[223]

The binary complex formed by UNC10088/PWWP1 (1:1) exhibited ΔHDX-MS consistent with interfacing observed in the crystal structure PDB CXCG of PWWP1 with UNC6934 (**Fig. 4.3c**,**g**). Incubation of UNC10088 with PWWP1 resulted in decreased deuterium uptake spanning mainly

266-277 and 319-335. The uptake kinetics of these regions are particularly striking because 266-277 exhibits decreased deuterium uptake from 1-10 minutes that becomes insignificant by 1 hr, whereas 319-335 appears on the longer timescale (≥ 10 mins). This reveals that, consistent with the UNC6834 crystal structure, the β-strand spanned by 266-277 makes direct contacts with UNC10088 and whereas the conformational freedom of the helix spanned by 319-335 is reduced due to the binding event.

The ternary complex (1:1:1) formation of FBXO22/UNC10088/PWWP1 was evidenced by newly populated ΔHDX-MS signals, distinct from those observed in binary complexes. To start, FBXO22 (**Fig. 4.3b**,**f**) decreased in deuterium uptake in peptides spanning 222-242, 251-258, and increased in deuterium uptake at 131-138. This may be due to new interfacing between FBXO22/PWWP1 stabilized by UNC10088, or UNC10088 populating a conformation that interfaces with FBXO22 when PWWP1 is bound. Moreover, 131-138 is distinct from the localized binding site (according to AlphaFold) which may point to an allosteric relationship. This is evidenced by the timescale at which newly formed decreased 222-242 appear (1-10 mins) which is the same as peptide 131-135, suggesting the structural changes are linked. Apparently, the 131-138 region (131-135, 134-138) is susceptible to destabilization given its high error relative to the rest of the data set.

On the other hand, PWWP1 did not experience any new involvement from regions other than some increased deuterium uptake around 282-292 and 308-320, indicative of subtly increased flexibility (**Fig. 4.3d**,**h**). This marginal impact on binding footprint size is understandable given the small size of the isolated PWWP1 domain. Instead, the same peptides which were impacted in the binary complex displayed slower converging kinetics; they have higher magnitude differences, and in the case of 266-277, the kinetics diverge for longer (1 hour, green).

4.5 Conclusion

This ΔHDX-MS enabled structural characterization helped corroborate ternary complex formation. The prolonged, high intensity changes in deuterium uptake suggest very tight interactions in both the binary and ternary complexes~~, even at equimolar stoichiometries~~. Moreover, ΔHDX-MS localized a putative covalent site (C326) which informed subsequent point mutagenesis to better understand the binding mechanism of UNC10088. These data can also provide a blueprint for future *in silico* modeling. To continue this study, since the isolated NSD2-PWWP1 domain was used for structural characterization, more may be revealed by using full-length NSD2. Given the capacity of bottom-up HDX to observe highly complex samples, in theory,

even the reconstituted ubiquitin ligase complex could be studied. Another strategy employed in the context of PROTACs is using ΔHDX-MS for SAR studies, such as observing how changing linker length impacts ternary structure.[224] Notably, at such high affinity, bifunctional molecules sometimes undergo the hook effect, where both binary complexes saturate in place of ternary complex formation.[223] This highlights a need for structural characterization at low concentrations as the hook effect becomes increasingly more significant with higher PROTAC concentrations. The above work is a testament to the utility of using ΔHDX-MS for rapid structural characterization of such dynamic structures and complexes.

CHAPTER 5

Small Molecule Screen Identifies Non-catalytic USP3 Chemical Handle

Mandeep K Mann[1,2], **Esther Wolf**[3], Madhushika Silva[1], Haejin Angela Kwak[1,2], Brian Wilson[4], Albina Bolotokova[1], Derek J. Wilson[3], Rachel J Harding[1,2], Matthieu Schapira[1,2]

[1]Structural Genomics Consortium, University of Toronto, 101 College Street, MaRS South Tower, Suite 700, Toronto, Ontario M5G 1L7, Canada
[2]Department of Pharmacology and Toxicology, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada
[3]Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada
[4]Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, Ontario M5G 0A3, Canada

This chapter is a version of a portion of the publication in ACS Omega 2023, 9 (1), 917-924, 10.1021/acsomega.3c07070,

M.K.M. expressed and purified proteins, designed and optimized biophysical assays and tested compounds., B.W conducted analytical chemistry. **E.W** conducted and D.J.W. supervised HDX experiments. H.K. conducted molecular modeling and simulations. A.B. dissolved, plated, and managed compounds. R.J.H. and M.S. advised throughout the project.

5.1 Summary

A focused small molecule screen targeting the Zinc-finger-ubiquitin binding domain (ZnF-UBD) of USP3 was conducted, identifying ligand **59**. Ubiquitin cleavage function of USP3 was observed to be retained at a high concentration of **59**, indicating that 59 does not inhibit catalytic function. This is particularly useful in developing bifunctional molecules for proximity induced pharmacology where the aim is to retain deubiquitylation capability. To complete characterization of **59**, HDX-MS was employed to determine the binding site. Parsing through the time-dependence of peptide HDX revealed which regions underwent direct contacts and binding-induced conformational change.

5.2 Introduction

Bifunctional molecules can facilitate proximity-induced pharmacology. For example, proximity to deubiquitylases (DUBs), which remove ubiquitin from targets to prevent their degradation, can increase target protein lifespan and abundance. Bifunctional molecules that have a DUB ligand, and a target protein ligand are called DUBTACS: deubiquitylase-targeting chimeras.

Importantly, to serve as a DUBTAC, the ligand targeting the DUB must not impact catalytic function. As such, the Zinc-finger ubiquitin binding domain (ZnF-UBD) can be targeted, which is a noncatalytic domain that binds the C-terminal end of ubiquitin (RLRGG motif) in 12 ubiquitin specific peptidases (USP). In an effort to find such DUB ligands, my coauthors conducted a screen against 11 members of the USP family using a ligand pool based on the chemical scaffold of C-terminus of ubiquitin.[225]

Ligand **59** was found to bind the USP3-ZnF-UBD with moderate selectivity compared to other DUBs (e.g., $K_{D,USP3}$=14 ± 4 μM while the next best affinity was $K_{D,USP16}$= 72 ± 16 μM). Using a fluorogenic ubiquitin rhodamine assay, USP3 was observed to retain DUB function in the presence of up to 1 mM **59**. Additionally, cleavage of a di-ubiquitin moiety was retained by full-length USP3 in the presence of **59**, whereas the isolated ZnF-UBD did not catalyze cleavage in the presence of **59**. This further confirmed that the ZnF-UBD was distinct from catalytic function, corroborating that binding of **59** would not inactivate catalysis.

Despite various efforts, crystallization of apo USP3 and co-crystallization of USP3 with ubiquitin or **59** was not achieved. Instead, this brief chapter explores how HDX-MS was used orthogonally to rapidly confirm binding and characterize the binding interface of **59**. Consistent with the chemical screen evolved from the ubiquitin C-terminus, **59** was observed to perturb the pocket occupied by ubiquitin in USP3 homologue USP5. The unique kinetic profiles of USP3 ZnF-

UBD peptides revealed which regions of the pocket likely made direct contacts (e.g., intramolecular bonds, interfacing) with **59**, and where binding-induced conformational changes occurred. This was further supported by comparing to USP5 cocrystal with ligand ZQ1, an analogue of **59**, which has pi-stacking in the region where the earliest kinetic differences were observed.

These results indicate that **59** can be a moderately selective chemical probe of USP3 which, given the preservation of catalytic function, can serve as a starting point for DUBTAC development.

5.3 Materials and Methods

HDX-MS. To begin, three samples were prepared and incubated for 30 min at 0 °C: 7.5 μM USP3 ZnF-UBD (apo), 7.5 μM USP3 ZnF-UBD and 75 μM **59** (1:10 complex), and 7.5 μM USP3 ZnF-UBD and 150 μM **59** (1:20 complex). All samples were then held at 0 °C prior to mixing 8 μL of the sample with 52 μL of buffered D_2O (10 mM phosphate buffer pD 7.5, 150 mM NaCl) to yield a final D_2O concentration of 87%. The HDX reaction was allowed to take place for 15, 60, or 600 s at 20 °C. Then, 50 μL of the HDX reaction was quenched for 1 min at 0 °C with 50 μL Quench buffer (100 mM phosphate buffer, pH 2.5) to stop the HDX reaction. Next, 50 μL (25 pmol) of the quenched sample was loaded onto a mixed Nepenthensin-2 Pepsin (1:1) column (Affipro) and desalted (ACQUITY UPLC BEH C18 VanGuard Precolumn, 130 Å, 1.7 μm, 2.1 mm × 5 mm, Waters) using Mobile Phase A (0.1% formic acid in water) at 200 μL/min for 3 min. The peptides were then reverse-phase separated (ACQUITY UPLC BEH C18 Column, 130 Å, 1.7 μm, 2.1 mm × 50 mm, Waters) at 40 μL/min using a gradient from 5 to 35% Mobile Phase B (0.1% formic acid in acetonitrile). Eluted peptides were electrosprayed into the Select Series Cyclic IMS (Q-IMS-TOF, Waters) while Leu-Enk was used for lockspray. Fragmentation was conducted using collision-induced dissociation in the $HDMS^e$ mode with a collisional energy ramp from 20 to 29 V (in the transfer cell). To obtain undeuterated peptides, the above steps were followed except in place of deuteration buffer, and an equilibration buffer was used (10 mM Phosphate Buffer pH 7.5, 150 mM NaCl). HDX was automated using the ACQUITY UPLC M-Class System with HDX Technology (Waters) and PAL3 liquid handling (CTC Analytics AG). Peptide identification was conducted using ProteinLynx Global Server 3.0.3 (PLGS, Waters). HDX analysis and visualization were conducted using DynamX 3.0 (Waters) and PyMOL 2.5.0.

5.4 Results

The structure of USP3 ZnF-UBD is not available from the PDB and, despite extensive crystallization screening efforts, solution by X-ray crystallography was not obtained, in either its apo form or in complex with ubiquitin or **59**.

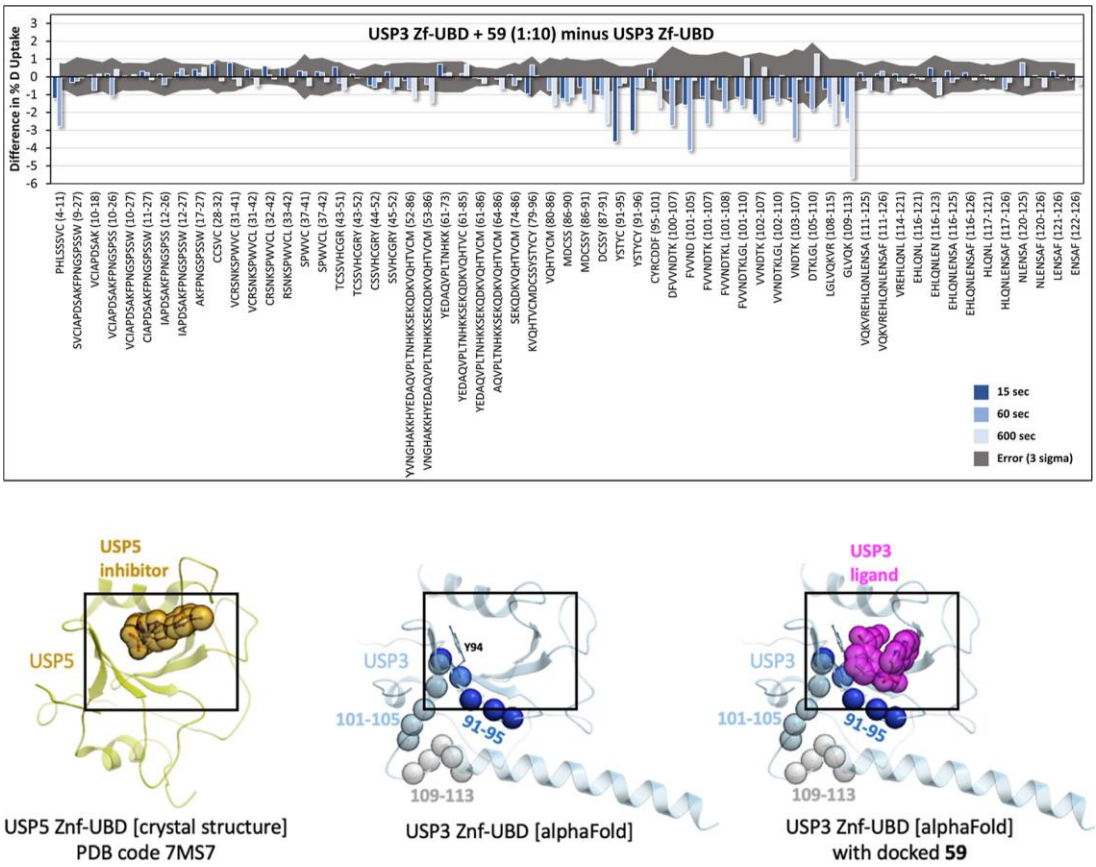


Figure 5.1 | **Mapping of the ligand binding site by HDX-MS.** Top: the difference in deuterium fractional uptake (%) between USP3 ZnF-UBD + **59** (1:10) and USP3 ZnF-UBD is plotted as a function of each time point per peptide. The time points are 15 s (dark blue), 60 s (light blue), and 600 s (light grey). The error (3 times the propagated error) is shown as a dark gray area. If the bars exceed this gray area, they are statistically significant. Bars pointing down indicate a decrease in the level of deuterium uptake in the presence of the ligand. Bottom: Maximum HDX difference signals observed at 15 s (blue), 60 s (light blue), and 600 s (light grey) are highlighted as spheres on an AlphaFold prediction of the USP3 ZnF-UBD structure (center). For reference, a crystal structure of USP5 ZnF-UBD bound to ZQ1, analogue of **59** (PDB 7MS7) (left) and a model of **59** docked to the USP3 ZnF-UBD AlphaFold model prediction (right). Reproduced with permission from Mann et al. (2023), open access under CC BY-NC-ND 4.0 DEED.[225]

To identify the binding site of **59** and characterize the binding event's corresponding allosteric effects, HDX-MS was used.[226,227] A sequence coverage of over 90% was obtained for USP3 ZnF-

UBD with a peptide per amino acid redundancy of 5.24 (**Table E1 and Fig. E1**). HDX-MS analysis was performed for USP3 ZnF-UBD alone and in complex with **59** (1:10 and 1:20) for 15, 60, and 600 s at 20 °C (**Fig. 5.1** & **E2**).

Given the 14 μM equilibrium dissociation constant (K_D), the bound fractions in the 1:10 and 1:20 states upon D_2O buffer addition were expected to be 40.1 and 58.1%, respectively. The HDX data described here are differential (deuterium uptake of bound state minus unbound state), in technical triplicate (*n=3*), and both the 1:10 and 1:20 ligand ratios show consistent results (effectively *n=6*, **Fig. E2**). To be considered statistically significant, differences must have exceeded triple the propagated standard error. Complexation resulted in significant attenuation of deuterium uptake rates in 20 of 59 peptides, all localized to the same region (**Fig. 5.1**).

At the earliest time point (15 s), a high-intensity, negative uptake difference was observed at residues 91–96 (YSTYCY), with the magnitude of the difference decreasing rapidly over the next two time points (1 and 10 min). Significant uptake differences appear for peptides 5–10 (HLSSSVC) and 102–107 (VVNDTK) at the 1 min time point. Finally, residues 110–113 (LVQK) exhibit a slowly evolving decreased deuterium uptake, whose magnitude becomes the dominant difference signal at the longest measured time point (10 min).

5.5 Discussion

All HDX differences localized to one region of USP3, the same pocket observed in an analogous USP5 structure, PDB 7MS7.[228] There, the USP5 ZnF-UBD is bound to compound ZQ1 (N-{5-[4-(4-chlorophenyl)piperidine-1-sulfonyl]pyridine-2-carbonyl}glycine), which shares a similar chemical scaffold to **59**. This same USP5 pocket recognizes the of the C-terminus of ubiquitin.[229]

When considering ΔHDX-MS peptide kinetics, reduced deuterium uptake in the presence of a ligand can arise from (i) intermolecular contacts or (ii) binding induced intramolecular changes (e.g., changes in dynamics at the binding site *interfacial*, or beyond, *allosteric*) in a time-dependent manor.[226,227,230,231] While intermolecular contacts and interfacing appear at earlier timepoints, relatively subtle conformational changes will appear later as they gradually build-up deuteration difference. The HDX differences at 15 sec for residues 91–96 quickly disappear by 60 sec, which is consistent with the high/fast k_{off} of **59** (as observed in the SPR sensorgrams, see **Fig. E3C**). Indeed, this agrees with intermolecular bonding through critical π-stacking in the docked model between the acetophenone group of **59** and Y94 (**Fig. 5.1,** bottom right). Analogous to the HDX-supported **59**/USP3 Y94 interaction, the 7MS7 USP5 cocrystal has a close π-stacking

interaction between the pyridine moiety of ZQ1 and Y259. All other significant decreases in uptake, corresponding to residues 5-10, 102–107, and 110–113, are near the predicted binding pocket, but exhibit slower appearing uptake differences, indicative of induced, interfacing effects.

5.6 Conclusion

Structural characterization can be quickly achieved using HDX-MS for dynamic, weakly populated systems, such as an enzyme domain or a compound with a moderately high K_D. Here, USP3 ZnF-UBD was found to interact with focused molecule screen hit **59** at the conserved ubiquitin C-terminus binding pocket. Although this was corroborated by USP5 crystal structures, these types of data are not always available, and in their absence, it may be difficult to interpret the HDX results. Therefore, it is useful to include shorter timepoints (15 sec), in addition to relatively longer ones (10 mins), to capture distinctions in peptide uptake kinetics which in turn provide clues about the putative binding site. With access to *in silico* predicted apo models, HDX can then guide and corroborate small molecule docking studies, a convenient, appropriate approach in the fast-paced early drug discovery environment.

CHAPTER 6

Broadly Neutralizing Humanized SARS-Cov-2 Antibody Binds to a Conserved Epitope on Spike and Provides Antiviral Protection Through Inhalation-Based Delivery in Non-Human Primates

Paule Hermet[1], Benoît Delache[2], Cecile Herate[2], **Esther Wolf**[3], Gaily Kivi[1], Erkki Juronen[1], Karl Mumm[1], Eva Žusinaite[4], Denis Kainov[5], Eve Sankovski[1], Kai Virumäe[1], Anu Planken[1], Andres Merits[4], Jessica E Besaw[6], Ai Woon Yee[6], Takefumi Morizumi[6], Kyumhyuk Kim[6], Anling Kuo[6], Asma Berriche[2], Nathalie Dereuddre-Bosquet[2], Quentin Sconosciuti[2], Thibaut Naninck[2], Francis Relouzat[2], Mariangela Cavarelli[2], Mart Ustav[1], Derek J. Wilson[3], Oliver P Ernst[6,7], Andres Männik[1], Roger LeGrand[2], Mart Ustav Jr[1]

[1]*Icosagen Cell Factory OU¨ ; Tartu, Estonia*
[2]*Universite´ Paris-Saclay, Inserm, CEA, Center for Immunology of Viral, Auto-immune, Hematological and Bacterial Diseases (IMVA-HB/IDMIT); Fontenay-aux-Roses, France*
[3]*York University; Toronto, Canada*
[4]*University of Tartu; Tartu, Estonia*
[5]*NTNU; Trondheim, Norway*
[6]*Department of Biochemistry, University of Toronto; Toronto, Canada*
[7]*Department of Molecular Genetics, University of Toronto; Toronto, Canada*

A portion of this chapter has been published in *PLoS Pathogens* 2023, *19* (8), e1011532, 10.1371/journal.ppat.1011532,

E.W. conducted all HDX-MS and data processing. D.J.W. provided supervision.

6.1 Summary

Icosagen recognized that antibody inhalation, in addition to being a non-invasive form of drug therapy, has a particularly enticing application given that severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) replicates in areas involved in inhalation and can be used for immunocompromised individuals who cannot be vaccinated.[232] Whilst conducting the necessary testing and structural characterization (X-Ray Crystallography, Cryo-EM) of monoclonal antibody (mAb) ICO-hu23, SARS-CoV-2 continued undergoing variant evolution, rendering the antibody unusable as a therapeutic. Upon restarting and discovering ICO-hu104, HDX-MS was chosen as a more rapid structural characterization method. Epitope comparison of ICO-hu23 and ICO-hu104 against the isolated Receptor Binding Domains (RBD) of Wuhan and Omicron revealed that the two mAb binding sites were distinct, and that ICO-hu104 was binding to an area undergoing minimal mutation. However, as we continue to see, the antigenic drift of SARS-CoV-2 has led to a loss of utility in antibody-based therapy against it.

6.2 Introduction

The global pandemic of coronavirus disease 2019 (COVID-19) was caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a lipid-enveloped, positive-sense, single stranded mRNA virus which is made up of several structural proteins: nucleocapsid (N), membrane (M), envelope (E), and spike (S) proteins (**Fig. 6.1a**). The remainder of the SARS-CoV-2 proteome includes 16 non-structural proteins (NSPs) and 6 accessory proteins.[233]

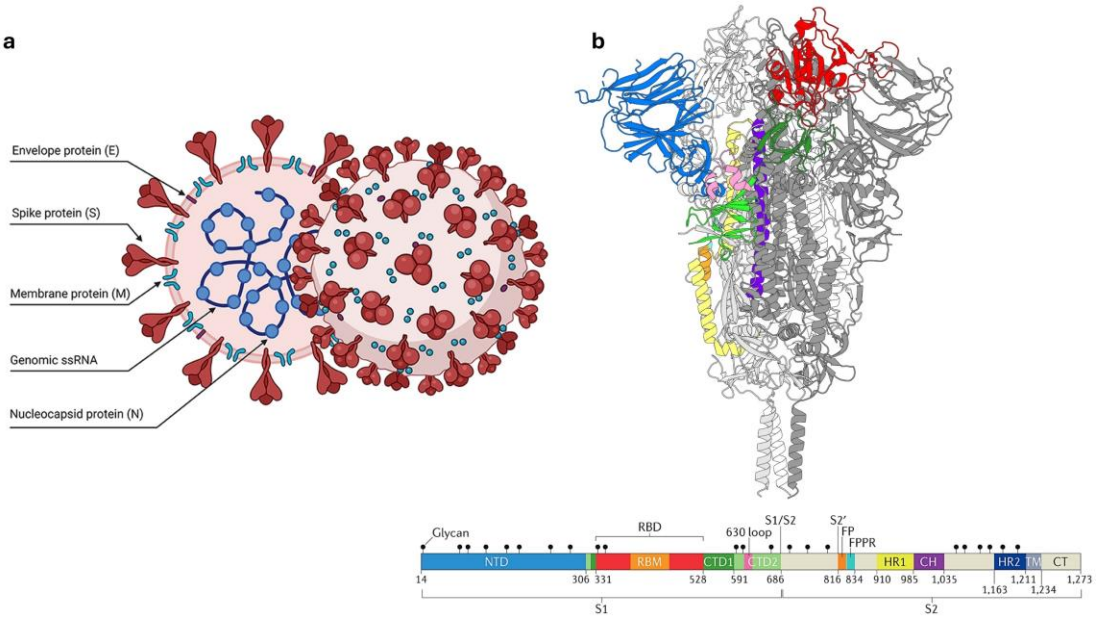


Figure 6.1 | **SARS-CoV-2 Virion and Spike Protein.** (a) The enveloped viral particle has several structural proteins which enclose genomic positive sense, single-stranded RNA: envelope (E), spike (S), membrane (M), and nucleocapsid (N). (b) The spike protein has distinct domains, including a receptor binding domain (RBD, green) and an N-terminal domain (NTD, blue). It is cleaved into S1 and S2 during virion formation at the S1/S2 site. S1 and S2 are noncovalently bound upon hACE2 recognition. (a) and (b) were reproduced with permission from Pizzato et al. (2022) (Open access under CC BY 4.0 DEED) and Jackson et al. (2021).[234,235]

The following discussion pertains to the spike glycoprotein, which is a homotrimer that protrudes from the viral particle. The monomeric spike is subdivided into 2 regions S1 and S2 (**Fig. 6.1b**), due to furin protease cleavage at the S1/S2 site (res. RRAR) during virus maturation in the Golgi apparatus of infected cells.[235] S1 contains the N-terminal domain (NTD), receptor binding domain (RBD), and 630 loop, while S2 contains an S2' cleavage site, fusion peptide (FP), heptad repeats 1 and 2 (HR1, HR2), and a transmembrane domain. For study *in vitro*, the transmembrane domain is removed and the S1/S2 RRAR site is often inactivated (e.g., res. GSAS), leaving the S1 and S2 covalently attached, whereas in physiological settings, non-covalent interactions between S1 and S2 are necessary for infection.[235]

The COVID-19 infection begins via human angiotensin-converting enzyme 2 (hACE2) receptor-mediated SARS-CoV-2 endocytosis.[235] On the host cell surface hACE2 binds the spike protein protruding from the viral particle at the RBD (**Fig. 6.2**). Particularly, the receptor binding motif (RBM, res. 437-508) primarily interacts with the N-terminal helix (res. ~30-45) of hACE2.[236]

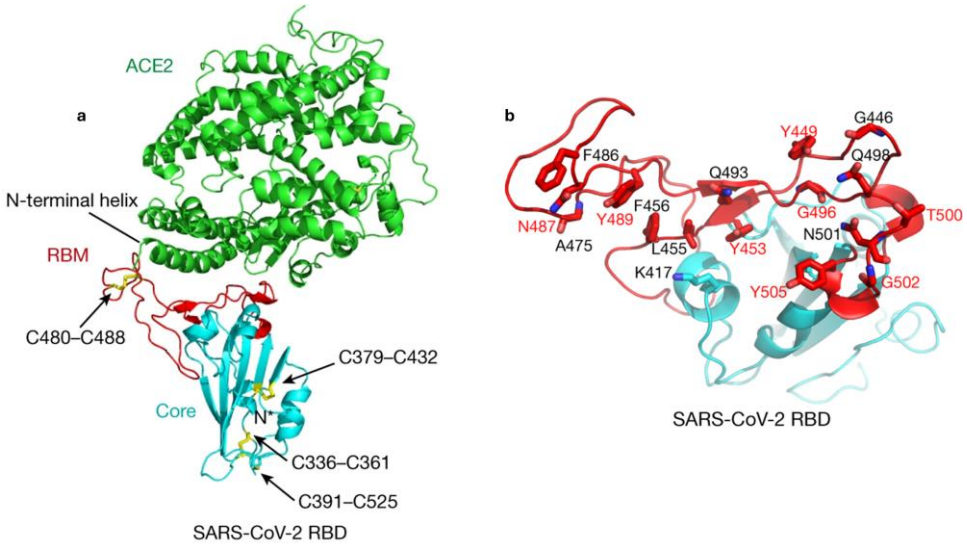


Figure 6.2 | **SARS-CoV-2 bound to hACE2.** Red residues of SARS-CoV-2 RBD (blue) interface with hACE2 N-terminal peptidase domain (green). Disulfide bonds (yellow) appear in several areas of the RBD. Reproduced with permission from Lan et al. (2020).[236]

At the plasma membrane, or once the viral particle has been engulfed into an endosome, the non-covalently attached S1 is shed and S2 is free to extend, exposing the S2' cleavage site. After S2' cleavage by host cell proteases including TMPRSS2 and cathepsins, the S2 inserts the FP into the host plasma/endosome membrane and facilitates host-virion membrane fusion so that the viral genome can enter the host cell.[235]

To prevent COVID-19, it was largely believed that the best therapeutic approach involved the disruption of Spike/hACE2 interaction. To do this, vaccine, drug, and antibody therapy development of was targeted at the spike S1/RBD. Icosagen, an Estonia-based Contract Research, Development, Manufacturing Organization (CRDMO), focused their efforts on developing an inhalable antibody therapy.[232] The inhalation of antibodies was particularly applicable with SARS-CoV-2 given the viruses replication in the respiratory tract, and could be an alternative medication for those who could not be vaccinated or those seeking non-invasive delivery. Icosagen screened patients who had recently overcome COVID-19, and isolated monoclonal antibody (mAb) ICO-hu23 for further product development. To structurally characterize the ICO-hu23 interaction with the spike protein, they employed Cryo-EM and Crystallography. Unfortunately, SARS-CoV-2 had other plans; the world watched as variants of concern (VoC) began to spread and disease symptoms worsened. The spike protein accumulated mutations, particularly at the RBD, which enabled its immune evasion from antibodies targeting

previously existing epitopes. This mutational phenomenon, termed antigenic drift, rendered many therapeutics less effective against the virus.

To restart, Icosagen newly identified ICO-hu104 and confirmed that it broadly neutralized VoCs Wuhan, Alpha, Beta, Delta, and Omicron.[232] They proceeded to quickly pursue structural characterization, this time using ΔHDX-MS. Epitope mapping using ΔHDX-MS has been a well-known strategy, dating back to 2002,[237–240] and, given the tightness of antibody-antigen interactions, it requires little sample and binding is readily observable. In fact, HDX-MS was quite widely used for rapid Spike protein structural studies during the COVID-19 pandemic.[241–246]

Using isolated RBDs of Wuhan and Omicron, the epitopes of ICO-hu23 and ICO-hu104 were revealed to be distinct. Then, although beyond the scope/needs of the manuscript, HDX-MS was further used to observe ICO-hu23 and ICO-hu104 impacts on the full-length Wuhan, Omicron, and Delta spike trimer structure. ICO-hu23 seemed to globally impact the trimer, causing a cascade of reduced conformational freedom along the length of the protein. On the other hand, ICO-hu104 seemed to clearly impact the predicted epitope while increasing the flexibility in some S2 regions. This increased S2 flexibility has been widely observed upon hACE2 binding[244] but it was interesting to see the distinction between ICO-hu104 and ICO-hu23, which did not clearly exhibit increased flexibility.

6.3 Materials and Methods

Protein Preparation. Proteins were expressed in CHO cell lines and purified using Ni-NTA affinity chromatography by our coauthors, Hermet et al. (2023),[232] including the Wuhan RBD (res. 319-541, Icosagen CAT#: P-307-100), Omicron RBD (B.1.1.529, res. 319-537, Icosagen CAT#: P-3607-100), Trimeric Spike Omicron (B.1.1.529, res. 14-1211, Icosagen CAT#: P-369-100), Trimeric Spike Delta (B.1.617.2, res. 14-1211, Icosagen CAT#: P-353-100), and Trimeric Spike Wuhan (res. 14-1211, Icosagen CAT#: P-309-100). The trimers had inactivated S1/S1 cleavage site (GSAS) and a C-terminal T4 fibritin trimerization sequence to improve *in vitro* stabilization of the intact trimer.

HDX-MS Epitope Mapping. Hydrogen-deuterium exchange was carried out using the commercial Waters HDX system. Briefly: samples were mixed and injected in a labeling time-randomized order by a PAL3 Autosampler followed by UPLC separation and mass spectrometry analysis using a Waters M-Class ACQUITY UPLC with peptides detected on a Waters Select Series Cyclic IMS Mass Spectrometer. Labeling times of 1, 10, and 60 min at RT were used, followed by quenching (7.5 M guanidine hydrochloride and 0.5M TCEP) at 0°C and digestion

using an Enzymate BEH Pepsin column. The peptides were reverse-phase separated using an ACQUITY UPLC BEH C18 column. Peptide identification was carried out using HDMSe with CID fragmentation in the transfer cell and ProteinLynx Global Server (PLGS), followed by HDX analysis using DynamX. Structures were visualized using PyMOL 2.5.

6.4 Results & Discussion

6.4.1 Determination of ICO-hu23 and ICO-hu104 epitopes using Spike RBDs.

Atomic resolution and explicit residue contacts were determined for ICO-hu23/Wuhan RBD using Crystallography conducted by our coauthors. HDX-MS of ICO-hu104 was chosen mainly for speed of characterization, at the expense of spatial resolution.

To characterize binding of ICO-hu23 and ICO-hu104 to the isolated RBD of Wuhan and Omicron, ΔHDX-MS was conducted for 1, 10, and 60 mins. The RBD sequence coverage may have been reduced due the presence of O/N-linked glycosylations (T323, S325, N331, N343)[247] or disulfide bonds (C336↔C361, C379↔C432, C391↔C525, C480↔C488)[236] which may interfere with protease accessibility. In any case, to enhance unfolding in the presence of disulfide bonds, a quench containing 7.5 M Gdn-HCl and 0.5M TCEP was used. This also ensured the downstream desalting and separation columns were kept clean by reducing accumulation of undigested protein.

The complexed state required only a 1:1 RBD-mAb stoichiometry given the tight sub-nanomolar K_D's previously determined.[232] The ICO-hu23/Wuhan RBD state had a sequence coverage of 78% and average redundancy of 4.59 (**Fig. F1**). Similarly, the ICO-hu104/Wuhan RBD and ICO-hu104/Omicron RBD had sequence coverage of 72% and 84%, respectively, and shared a peptide-per-residue redundancy of 2.97.

Upon binding ICO-hu23, the Wuhan RBD intensely decreased in deuterium uptake about regions 472–490 and 495-497 across the entire time course, and more moderate decreases were observed in regions 433–449 and 516–533 at 10 and 60 mins (**Fig. 6.2a**). These results indicated that the ICO-hu23/Wuhan epitope was situated on the RBM (res. 437-508), making ICO-hu23 an orthosteric SARS-CoV-2 neutralizing antibody by competing with hACE2 binding,[98,236] an abundantly reported mechanism of action.[248]

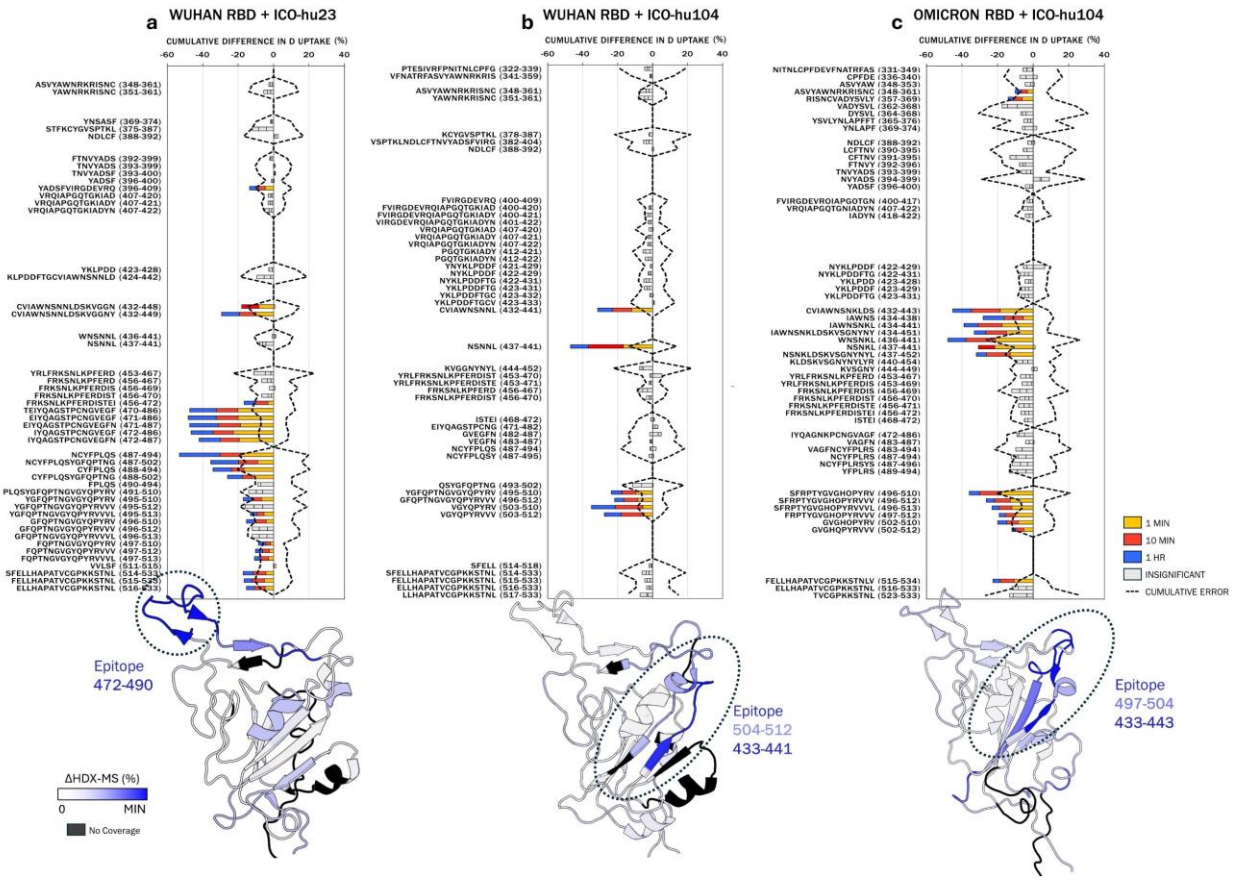


Figure 6.3 | **Localizing ICO-hu23 and ICO-hu104 epitopes.** The ΔHDX-MS at 1 (yellow), 10 (red), and 60 mins (blue) of Wuhan RBD in the presence of (a) ICO-hu23, (b) ICO-hu104, and (c) Omicron RBD in the presence of ICO-hu104. Although sequence coverage was above 75% for each system, the peptide lists were not identical, so the general regions were grouped together across (a-c). The structures below (Wuhan PDB 7DDN and Omicron PDB 7TGW) illustrate the 1 min ΔHDX-MS timepoint to reveal epitopes in dark blue while absence of coverage is coloured black (see **Fig. F2** for residue averaged maps). The gradient is set to a minimum of -15%.

The RBD contains 13 dispersed Omicron mutations, with five spanning ICO-hu23 epitope residues 472-497 (S477N, T478K, E484A, Q493R, and G496S). This is consistent with the lack of affinity of ICO-hu23 for the Omicron RBD. In the case of broadly neutralizing ICO-hu104, the Wuhan RBD underwent distinct decreases in deuterium uptake compared to ICO-hu23 (**Fig. 6.2b**). High intensity reduced uptake was observed in residues 433-441 at the earlier HDX timepoints and later appearing decreases were observed at 504-512. This positioned the ICO-hu104 epitope on a set of beta-strands farther from the ICO-hu23 epitope, but still orthosteric to the hACE2 binding site. Similar behaviour was observed in the Omicron RBD; ICO-hu104 induced high intensity deuterium uptake decreases around 433-443 and 497-502. Between the two RBDs

tested against ICO-hu104, residues 433-441 are similarly perturbed. Given this consistency, these residues are likely primarily responsible for ICO-hu104 broad neutralizing capability.

6.4.2 ICO-hu23 and ICO-hu104 impacts on full-length Spike.

Beyond the work of the manuscript, we were interested to see the impact of antibody binding on the full-length Spike trimer. Various classes of Spike-targeting antibodies have been identified, including those that bind the RBD in the up position (Class 1), both the up and down position (Class 2), and outside of the hACE2 binding interface (Class 3).[249] Our coauthors performed Cryo-EM with ICO-hu23 and full-length spike,[232] where they observed binding in the down RBD confirmation for all three monomers, suggesting ICO-hu23 is in Class 2. Accordingly, using HDX-MS, distinctions between Class 1 and 2 could then be revealed from long distance structural conformational changes of the full-length Spike. However, given the nature of HDX-MS, the impacts on the three spike monomers would be averaged, so we would not be able to observe variable monomer binding in the way that Cryo-EM or Crystallography could.

As in our RBD exploration, the quench step required denaturant and reducing agent (Gdn-HCl, TCEP). The timescale selected was 5, 10, 30 mins. All samples were treated identically: the apo state was 15 uM and the complexed state was equimolar (1:1). In the presence of ICO-hu23 and ICO-hu104, the sequence coverage (and peptide-per-residue redundancy) of the Wuhan Spike was 64% (2.74) and 68% (2.92), respectively (**Fig. F3**). These values were similar for Omicron and Delta, at 62% (2.07) and 65% (2.92), respectively. Relative to the isolated RBD experiments, the sequence coverage of the RBD dropped in the context of full-length spike: Delta (67%), Wuhan (63% and 68%), and Omicron (66%). The sequence coverage (and signal intensity) was likely reduced due to overloading of the pepsin column. Although the concentration is necessary for MS sensitivity and confident isotopic distribution fitting, the total mass of protein increases on column with the added spike sequence and mAb, reducing digestion efficiency.

Generally, ICO-hu23 induced the highest intensity ∆HDX-MS (**Fig. F4**) with a maximum absolute cumulative difference signal of >70%. To aid in visualization of the remaining spikes, the gradient in ∆HDX-MS **Fig. 6.4** was set to a maximum cumulative signal of 9% and a minimum of -30%. ∆HDX-MS revealed that, in addition to the RBD, ICO-hu23 induced decreased deuterium uptake in various regions along the Wuhan trimer including the linker regions of the RBD and NTD, and the S2 helices and helical hinges. Taken together, this suggests a globally reduced conformational freedom which agrees with the previously mentioned Cryo-EM results.

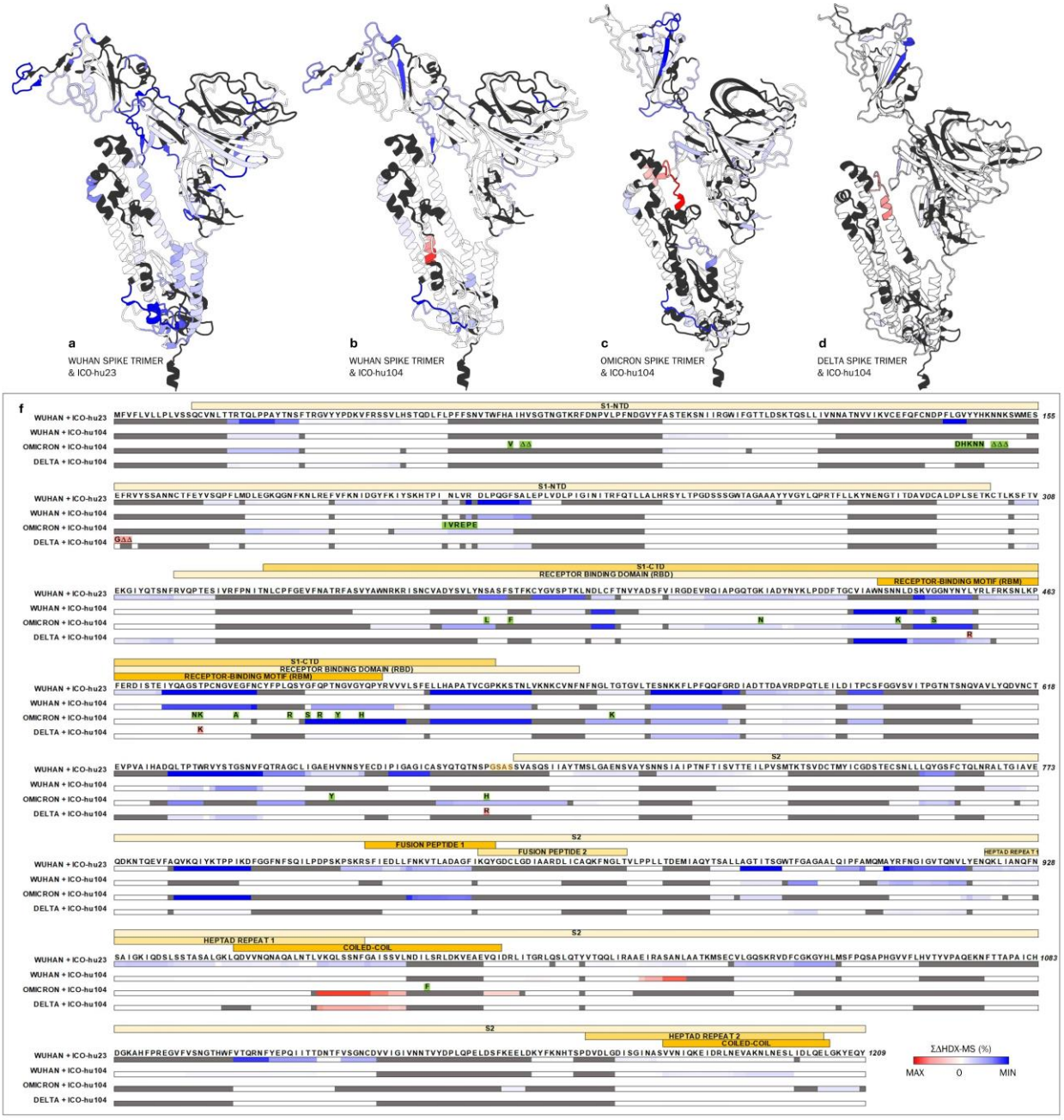


Figure 6.4 | **Impacts of ICO-hu23 and ICO-hu104 on Spike Trimer.** Cumulative ΔHDX-MS (f) of (a) ICO-hu23/Wuhan Spike, and ICO-hu104 with (b) Wuhan Spike, (c) Omicron Spike, and (d) Delta Spike. Structures of Wuhan Spike (PDB 7DDN), Omicron Spike (7TGW), Delta (7ZJL) coloured according to (f). Gradient set to cumulative ΔHDX-MS min. (blue) of -30% and max. (red) of 9%. White indicates no statistically significant cumulative signal (3 times the summed propagated error, **Fig. F4**) and black/grey denotes missing sequence coverage.

Conversely, outside of the RBD, ICO-hu104 seemingly had the highest intensity impacts on the coiled regions of the S2 stock, resulting in increased deuterium uptake. Given this unique

profile relative to ICO-hu23, and the distinct location of the ICO-hu104 epitope, perhaps this is indicative of the antibody preferentially binding the up RBD position. It is also possible that in solution, the spike trimer was not fully occupied by the mAb, diminishing more subtle structural changes, and leaving only the most significantly impacted regions to be observed.[250] Notably, as stated earlier, the Spike construct used for *in vitro* study is distinct from the physiologically active Spike, whose S1 and S2 domains are cleaved apart and noncovalently attached. With that said, it is unclear whether this experimentally observed phenomenon would translate to infection mechanism of action.

2.5 Conclusions

The development of therapeutics against SARS-CoV-2 was challenging given how quickly VoCs were emerging. Particularly, the VoCs carrying mutations dispersed along the RBD rendered many antibody-based therapies less effective due to changes in epitope. ΔHDX-MS was used to quickly characterize the epitope ICO-hu104 after the original mAb, ICO-hu23, was rendered ineffective due to antigenic drift. The epitope of ICO-hu104 was revealed to be positioned distal to the ICO-hu23 epitope, and ICO-hu104 binding to the full-length Spike seemingly increased flexibility in the S2 region, unlike ICO-hu23. Unfortunately, despite these structural distinctions that may make ICO-hu104 resistant to decreases in efficacy as VoCs continue to emerge, an alarming reduction in mAb neutralization of Omicron subvariants has been observed.[251–253] Simultaneously, the affinity for hACE2 is mostly unchanged.[251] Indeed, viral evolution will select for mutations that improve fitness, such as immune evasion and, fortunately, also reduced severity of disease symptoms. Perhaps, in this context, hACE2 would have been a potential alternative direction for inhalation-delivered mAb therapy.

CHAPTER 7

Conclusions and Future Work

7.1 Conclusions

This work aimed to convey the utility and applicability of using MS and HDX-MS as an orthogonal means to resolve pertinent structural biology questions. Access to a variety of protein systems was enabled through collaborators from across the world, including the lab of Bryan Dickinson at the University of Chicago, the lab of Lindsey Ingerman-James at the University of North Carolina at Chapel Hill, the Estonian research organization Icosagen, and the Structural Genomics Consortium. Chapter 1 focused on the basics of MS interpretation, mass analyzers, and how ions enter the gas phase through ESI. It went on to summarize the basics of protein HDX-MS as a structural biology tool as well as the experimental considerations of the bottom-up HDX workflow and TRESI.

Chapter 2 covered the structural study of superimposable proteins Bcl-2 and Mcl-1 using various MS techniques and short HDX labeling times. Both apo proteins had comparable native MS charge state distributions, consistent with their similar topology. Upon BH3 peptide addition, Mcl-1 appeared to have a mixture of bound and unbound peaks compared to Bcl-2, suggesting that Mcl-1 interacts with the peptides in a distinct manner. This was further supported by the IMS mobiligrams observed for each protein in their bound and unbound states. Bottom-up, segment averaged HDX-MS revealed that despite binding both distinct and shared peptide ligands, each protein had unique conformational dynamics. These conformational changes appeared at the same hotspots regardless of the peptide in complex with Bcl-2 or Mcl-1. In order to accommodate ligand binding, Mcl-1 displayed broader conformational changes. Finally, these results were compared to in silico derived homology models, revealing that perhaps Bcl-2 relies more on charge compensation than Mcl-1 for complexation.

Chapter 3 delved into higher throughput HDX-MS used to study a focused library of 13 small molecules targeting WDR5. The changes in deuterium uptake were largely comparable to the interfaces observed in several crystal structures; however, some areas that were expected to be perturbed remained HDX-silent or undetected. Further, HDX-MS revealed that the histone 3 methyl binding pocket was allosterically linked to intermolecular bond formation at the WIN site. The library was then broadened to include commercially available ligands and peptides. These data were aggregated to develop signal-to-noise affinity ranking methods: cumulative,

characteristic, and maximum, each with their own optimal application. Finally, the S/N was plotted against K_D to extract K_D of new additions to the library.

Chapter 4 described the characterization of a novel PROTAC, UNC10088, using an extended HDX-MS timescale. The bifunctional molecule was observed to form binary complexes with NSD2-PWWP1 and FBXO22. Although the interaction of PWWP1/UNC6934 was previously crystallized (UNC6934 being the target 'handle' of the PROTAC UNC10088), the peptide uptake kinetics of FBXO22/PROTAC elucidated persistent contacts formed between FBXO22 and UNC10088. This helped inform mutational analysis by identifying C326 as a potential covalent site for the PROTAC. Then, the ternary complex formation was both confirmed and characterized using increases in the binding footprint area, as well as changes in HDX uptake rates. Given that the only FBXO22 structure available was an *in silico* model, HDX-MS was an ideal candidate for quick, attainable structural studies of this complex, PROTAC-mediated system.

Chapter 5 was a brief but satisfying example of using peptide uptake kinetics to categorize structural components of binding between compound screen hit **59** and USP3. This was corroborated by crystal structures of a protein analogue, USP5; however, in the absence of these structures, the HDX-MS was sufficient to inform molecular docking. These results can pave the way for developing DUBTACs by making a bifunctional molecule based on **59**.

Chapter 6 explored the world of SARS-CoV-2 spike protein targeting mAbs that could be used in inhalation-based therapy. It is a concise example/microcosm of the antigenic drift experienced world-wide, both by those developing and using medicines. After identifying the epitopes of ICO-hu23 and ICO-hu104 using the isolated RBD, the impacts of the mAb binding were further explored using full-length Spike protein. ICO-hu23 was revealed to cause global decreases in dynamics, whereas ICO-hu104 induced increased flexibility in the S2 domain. This may categorize them as class 2 and class 1 neutralizing mAbs, respectively.

7.2 Future Work

7.2.1 The Potential of MS and HDX-MS to Study Complex Systems

The continued improvement in throughput, versatility, and robustness of MS, HDX-MS, and front-end tools leaves so much potential for tackling increasingly larger, more dynamic, and more complex biological systems. MS is one of the few biophysical tools that can report on all levels of protein structure and interactions using mass, mass shift, fragmentation, footprinting techniques, and collision-induced unfolding (to name a few).[254] The future of MS instrumentation is bright; it has and will continue to evolve in various aspects, including user interface, forms of fragmentation, sensitivity, and applications.

Front-end analytical separation enhances sensitivity (on-column concentration and partitioned ionization) while simultaneously desalting analytes to be MS-compatible. It is no longer a requirement for samples to be high purity, buffer exchanged into ammonium acetate, or available in high concentrations. Sample matrices can include, for example, lipid environments for studying membrane-associated proteins, such as nanodiscs, liposomes, or virus-like particles (VLPs).[255] Lipids can be removed by introducing zirconium oxide.[256] In fact, lipid removal modules are presently commercially available as add-ons to robotic tool change systems (e.g., LEAP X-Press).[257] LC pump and valve-based fluid delivery are also modular, enabling introduction of secondary analytical columns, such as SEC, prior to proteolysis which can be used to separate the analyte from free lipids and detergents, even in acidic quench conditions.[258] There is also evidence that higher flowrate pumping systems (e.g., Waters I Class UPLC) can yield comparable or even improved separation for HDX-MS.[259] This is promising for reducing back-exchange and duration of HDX experiments.

Moreover, IMS-MS instruments offer multiple dimensions of sample separation that yield low complexity, interpretable mass spectra. Excitingly, the Select Series cIMS-MS has improved HDX-MS sensitivity and separation compared to its predecessor.[57] This is beneficial to HDX and other fields such as proteomics and metabolomics which can be extremely heterogenous and complex within a narrow mass range.

7.2.2 Unexpected Limitations

A lived scenario that ties many of the preceding ideas together was the (attempt to conduct) epitope mapping of a mAb targeting a membrane protein expressed on the surface of a virus-like

particle (VLP).[260] VLPs harness a viral structural protein (e.g., gag protein of HIV-1) to bud off of a host cell, generating a host membrane-enveloped vehicle.[261] Accordingly, VLPs carry endogenous host cell cargo, viral structural proteins, and a protein of interest (POI) that has been overexpressed in a highly physiologically relevant membrane. A possible means to obtain epitope information would be to conduct HDX-MS with and without the mAb, use SEC prior to proteolysis to remove lipids/small contaminants, and rely on reverse phase peptide separation and cIMS to detect POI peptides. In practice, the POI was not abundant enough relative to all the other matrix components to get sufficiently proteolyzed for reasonable sequence coverage or MS signal intensity. Then, sequence coverage was obtained by limited proteolysis, although none spanning the surface exposed region of the POI (**Fig. G1**), which suggested the sample may have degraded. Perhaps an alternative approach to epitope mapping would have required switching out the VLPs for a simpler lipid model system, such as liposomes.

7.2.3 Where will we see HDX-MS next?

It is safe to say that HDX-MS has plenty to offer in the fast-paced early drug discovery environment. A particular niche where HDX-MS and other solution phase protein footprinting are highly amenable is bifunctional molecule development, such as PROTACs and DUBTACs, where the dynamic system requires dilute protein-ligand concentrations to prevent the hook effect.[223,224] The hook effect occurs when both binary complexes saturate out with bifunctional molecule in place of ternary complex formation. Since HDX-MS thrives amongst dilute, dynamic samples, ternary complex formation could be simultaneously confirmed and structurally characterized using changes in uptake kinetics or changes in footprint relative to the binary complex. SAR studies can be conducted to observe how changes to ligand functional groups or linker length impact HDX magnitudes. These results can be additionally corroborated using native or intact MS, provided the complex is stable in ammonium acetate or water, etc., or in acidic conditions, respectively.

On the topic of massive, dynamic proteins that can be studied by HDX-MS, the Huntingtin (HTT) / HTT-associated protein 40 (HAP40) complex is nearly 400 kDa. Full-length HTT structure was obtained for the first time in 2018 by Cryo-EM, in complex with HAP40, but has never been crystallized.[262] Though the function of HTT is not well understood, mutant HTT is associated with the onset of Huntington's Disease (HD), particularly when its unstructured, dynamic polyglutamate (polyQ) repeat exceeds a threshold of ~35 repeats.[263] The neurodegenerative disorder is debilitating, marked by uncontrolled movements and loss of cognitive function, eventually resulting in death. Although medicines are available to target the *symptoms* of HD, and some RNA/DNA and antibody therapies show promise, no HTT small molecule/drug antagonists are

available.[264] Thus, HDX-MS provides a solution-phase structural characterization platform, orthogonal to Cryo-EM, for early, preclinical drug screening. This was exemplified by the binding characterization of 6 HTT-targeting macrocyclic peptides (an example in **Fig. G2**).

These are only two cases, but, without a doubt, MS and HDX-MS are becoming instrumental in fast-paced environments alongside gold standard structural biology techniques. The throughput, accessibility to HDX reagents, and straight forward data analysis make it particularly appealing. Moreover, the structures of dynamic systems which evade crystallization efforts can be obtained through other means, such as Cryo-EM, NMR, or increasingly common *in silico* prediction, making HDX-MS interpretation by visualization more user friendly. The use of MS and HDX-MS will continue to be more widespread as interdisciplinary collaboration grows and instrumentation becomes less cost prohibitive.

References

(1) Kale, J.; Osterlund, E. J.; Andrews, D. W. BCL-2 Family Proteins: Changing Partners in the Dance towards Death. *Cell Death Differ.* **2018**, *25* (1), 65–80. https://doi.org/10.1038/cdd.2017.186.

(2) Cui, Z.; Li, H.; Liang, F.; Mu, C.; Mu, Y.; Zhang, X.; Liu, J. Effect of High WDR5 Expression on the Hepatocellular Carcinoma Prognosis. *Oncol. Lett.* **2018**, *15* (5), 7864–7870. https://doi.org/10.3892/ol.2018.8298.

(3) Taka, E.; Yilmaz, S. Z.; Golcuk, M.; Kilinc, C.; Aktas, U.; Yildiz, A.; Gur, M. Critical Interactions Between the SARS-CoV-2 Spike Glycoprotein and the Human ACE2 Receptor. *J. Phys. Chem. B* **2021**, *125* (21), 5537–5548. https://doi.org/10.1021/acs.jpcb.1c02048.

(4) Liu, Z.; Hu, M.; Yang, Y.; Du, C.; Zhou, H.; Liu, C.; Chen, Y.; Fan, L.; Ma, H.; Gong, Y.; Xie, Y. An Overview of PROTACs: A Promising Drug Discovery Paradigm. *Mol. Biomed.* **2022**, *3*. https://doi.org/10.1186/s43556-022-00112-0.

(5) Henning, N. J.; Boike, L.; Spradlin, J. N.; Ward, C. C.; Liu, G.; Zhang, E.; Belcher, B. P.; Brittain, S. M.; Hesse, M. J.; Dovala, D.; McGregor, L. M.; Valdez Misiolek, R.; Plasschaert, L. W.; Rowlands, D. J.; Wang, F.; Frank, A. O.; Fuller, D.; Estes, A. R.; Randal, K. L.; Panidapu, A.; McKenna, J. M.; Tallarico, J. A.; Schirle, M.; Nomura, D. K. Deubiquitinase-Targeting Chimeras for Targeted Protein Stabilization. *Nat. Chem. Biol.* **2022**, *18* (4), 412–421. https://doi.org/10.1038/s41589-022-00971-2.

(6) Chen, G.; Tao, L.; Li, Z. Recent Advancements in Mass Spectrometry for Higher Order Structure Characterization of Protein Therapeutics. *Drug Discov. Today* **2022**, *27* (1), 196–206. https://doi.org/10.1016/j.drudis.2021.09.010.

(7) Dueñas, M. E.; Peltier-Heap, R. E.; Leveridge, M.; Annan, R. S.; Büttner, F. H.; Trost, M. Advances in High-throughput Mass Spectrometry in Drug Discovery. *EMBO Mol. Med.* **2022**, *15* (1), e14850. https://doi.org/10.15252/emmm.202114850.

(8) Griffiths, J. A Brief History of Mass Spectrometry. *Anal. Chem.* **2008**, *80* (15), 5678–5683. https://doi.org/10.1021/ac8013065.

(9) McLafferty, F. W. Mass Spectrometric Analysis. Molecular Rearrangements. *Anal. Chem.* **1959**, *31* (1), 82–87. https://doi.org/10.1021/ac60145a015.

(10) Di Girolamo, F.; Lante, I.; Muraca, M.; Putignani, L. The Role of Mass Spectrometry in the “Omics” Era. *Curr. Org. Chem.* **2013**, *17* (23), 2891–2905. https://doi.org/10.2174/13852728178881311181627 25.

(11) Cabral, E. C.; Mirabelli, M. F.; Perez, C. J.; Ifa, D. R. Blotting Assisted by Heating and Solvent Extraction for DESI-MS Imaging. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (6), 956–965. https://doi.org/10.1007/s13361-013-0616-y.

(12) Miller, L. M.; Jarrold, M. F. Charge Detection Mass Spectrometry for the Analysis of Viruses and Virus-like Particles. *Essays Biochem.* **2023**, *67* (2), 315. https://doi.org/10.1042/EBC20220101.

(13) Leskovar, B. Microchannel Plates. *Phys. Today* **1977**, *30* (11), 42–49. https://doi.org/10.1063/1.3037791.

(14) Ladislas Wiza, J. Microchannel Plate Detectors. *Nucl. Instrum. Methods* **1979**, *162* (1), 587–601. https://doi.org/10.1016/0029-554X(79)90734-1.

(15) Tanaka, M.; Iijima, Y.; Masuda, Y.; Sato, T.; Mineta, T. Adjustment Method of MEMS Dual-Cantilever Deflection Using Plastic Deformation of Al and Ni Thin Film by Thermal Annealing. *Electronics* **2023**, *12* (14), 3153. https://doi.org/10.3390/electronics12143153.

(16) Carr, S. A.; Annan, R. S. Overview of Peptide and Protein Analysis by Mass Spectrometry. *Curr. Protoc. Protein Sci.* **1996**, *4* (1), 16.1.1-16.1.27. https://doi.org/10.1002/0471140864.ps1601s04.

(17) Viswanathan, B. Chapter 5 - Nuclear Fission. In *Energy Sources*; Viswanathan, B., Ed.; Elsevier: Amsterdam, 2017; pp 113–126. https://doi.org/10.1016/B978-0-444-56353-8.00005-8.

(18) Winkler, R. ESIprot: A Universal Tool for Charge State Determination and Molecular Weight Calculation of Proteins from Electrospray Ionization Mass Spectrometry Data. *Rapid Commun. Mass Spectrom.* **2010**, *24* (3), 285–294. https://doi.org/10.1002/rcm.4384.

(19) Konermann, L.; Douglas, D. J. Equilibrium Unfolding of Proteins Monitored by Electrospray Ionization Mass Spectrometry: Distinguishing Two-State from Multi-State Transitions. *Rapid Commun. Mass Spectrom.* **1998**, *12* (8), 435–442. https://doi.org/10.1002/(SICI)1097-0231(19980430)12:8<435::AID-RCM181>3.0.CO;2-F.

(20) Dole, M.; Hines, R. L.; Mack, L. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Gas Phase Macroions. *Macromolecules* **1968**, *1* (1), 96–97. https://doi.org/10.1021/ma60001a017.

(21) Kebarle, P.; Verkerk, U. H. Electrospray: From Ions in Solution to Ions in the Gas Phase, What We Know Now. *Mass Spectrom. Rev.* **2009**, *28* (6), 898–917. https://doi.org/10.1002/mas.20247.

(22) Hiraoka, K. Fundamentals of Electrospray. In *Fundamentals of Mass Spectrometry*; Hiraoka, K., Ed.; Springer: New York, NY, 2013; pp 145–171. https://doi.org/10.1007/978-1-4614-7233-9_8.

(23) Taylor, G. I.; McEwan, A. D. The Stability of a Horizontal Fluid Interface in a Vertical Electric Field. *J. Fluid Mech.* **1965**, *22* (1), 1–15. https://doi.org/10.1017/S0022112065000538.

(24) Fernández de la Mora, J. The Fluid Dynamics of Taylor Cones. *Annu. Rev. Fluid Mech.* **2007**, *39* (1), 217–243. https://doi.org/10.1146/annurev.fluid.39.050905.110159.

(25) Hauner, I. M.; Deblais, A.; Beattie, J. K.; Kellay, H.; Bonn, D. The Dynamic Surface Tension of Water. *J. Phys. Chem. Lett.* **2017**, *8* (7), 1599–1603. https://doi.org/10.1021/acs.jpclett.7b00267.

(26) Konermann, L. A Simple Model for the Disintegration of Highly Charged Solvent Droplets during Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (3), 496–506. https://doi.org/10.1016/j.jasms.2008.11.007.

(27) Rayleigh, Lord. On the Equilibrium of Liquid Conducting Masses Charged with Electricity. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1882**, *14* (87), 184–186. https://doi.org/10.1080/14786448208628425.

(28) 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.

(29) Mehmood, S.; Allison, T. M.; Robinson, C. V. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu. Rev. Phys. Chem.* **2015**, *66* (Volume 66, 2015), 453–474. https://doi.org/10.1146/annurev-physchem-040214-121732.

(30) Harvey, S. R.; Gadkari, V. V.; Ruotolo, B. T.; Russell, D. H.; Wysocki, V. H.; Zhou, M. Expanding Native Mass Spectrometry to the Masses. *J. Am. Soc. Mass Spectrom.* **2024**, *35* (3), 646–652. https://doi.org/10.1021/jasms.3c00352.

(31) Ouellette, R. J.; Rawn, J. D. 3 - Alkanes and Cycloalkanes. In *Principles of Organic Chemistry*; Ouellette, R. J., Rawn, J. D., Eds.; Elsevier: Boston, 2015; pp 65–94. https://doi.org/10.1016/B978-0-12-802444-7.00003-3.

(32) David, E.; Niculescu, V.-C. Volatile Organic Compounds (VOCs) as Environmental Pollutants: Occurrence and Mitigation Using Nanomaterials. *Int. J. Environ. Res. Public. Health* **2021**, *18* (24), 13147. https://doi.org/10.3390/ijerph182413147.

(33) The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49* (D1), D480–D489. https://doi.org/10.1093/nar/gkaa1100.

(34) Davis, B. T. V.; Velyvis, A.; Vahidi, S. Fluorinated Ethylamines as Electrospray-Compatible Neutral pH Buffers for Native Mass Spectrometry. *Anal. Chem.* **2023**, *95* (48), 17525–17532. https://doi.org/10.1021/acs.analchem.3c02640.

(35) Chang, H.-H. Computer Simulations of Isolated Conductors in Electrostatic Equilibrium. *Phys. Rev. E* **2008**, *78* (5), 056704. https://doi.org/10.1103/PhysRevE.78.056704.

(36) Ahadi, E.; Konermann, L. Ejection of Solvated Ions from Electrosprayed Methanol/Water Nanodroplets Studied by Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2011**, *133* (24), 9354–9363. https://doi.org/10.1021/ja111492s.

(37) Ahadi, E.; Konermann, L. Modeling the Behavior of Coarse-Grained Polymer Chains in Charged Water Droplets: Implications for the Mechanism of Electrospray Ionization. *J. Phys. Chem. B* **2012**, *116* (1), 104–112. https://doi.org/10.1021/jp209344z.

(38) Konermann, L.; Rodriguez, A. D.; Liu, J. On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. *Anal. Chem.* **2012**, *84* (15), 6798–6804. https://doi.org/10.1021/ac301298g.

(39) Page, J. S.; Kelly, R. T.; Tang, K.; Smith, R. D. Ionization and Transmission Efficiency in an Electrospray Ionization–Mass Spectrometry Interface. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (9), 1582–1590. https://doi.org/10.1016/j.jasms.2007.05.018.

(40) Waters. *Triwave - More Complete Characterization of Mixtures and Molecules.*; White Paper Library Number: LITR134678884. Part Number: 720004176en.; 2012. https://www.waters.com/webassets/cms/library/docs/720004176en.pdf.

(41) Finehout, E. J.; Lee, K. H. An Introduction to Mass Spectrometry Applications in Biological Research. *Biochem. Mol. Biol. Educ.* **2004**, *32* (2), 93–100. https://doi.org/10.1002/bmb.2004.494032020331.

(42) Wiley, W. C.; McLaren, I. H. Time-of-Flight Mass Spectrometer with Improved Resolution. *Rev. Sci. Instrum.* **1955**, *26* (12), 1150–1157. https://doi.org/10.1063/1.1715212.

(43) Honour, J. W. Benchtop Mass Spectrometry in Clinical Biochemistry. *Ann. Clin. Biochem.* **2003**, *40* (6), 628–638. https://doi.org/10.1258/000456303770367216.

(44) Price, D. Time-of-Flight Mass Spectrometry. In *Time-of-Flight Mass Spectrometry*; ACS Symposium Series; American Chemical Society, 1993; Vol. 549, pp 1–15. https://doi.org/10.1021/bk-1994-0549.ch001.

(45) Hoffmann, E. de; Stroobant, V. 2.1 Quadrupole Analyzers. In *Mass Spectrometry: Principles and Applications*; Wiley-Interscience: Chichester, West Sussex, England ; Hoboken, NJ, 2007; pp 88–93.

(46) Clark, A. E.; Kaleta, E. J.; Arora, A.; Wolk, D. M. Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry: A Fundamental Shift in the Routine Practice of Clinical Microbiology. *Clin. Microbiol. Rev.* **2013**, *26* (3), 547–603. https://doi.org/10.1128/cmr.00072-12.

(47) Cunsolo, V.; Muccilli, V.; Saletti, R.; Foti, S. Mass Spectrometry in Food Proteomics: A Tutorial. *J. Mass Spectrom.* **2014**, *49* (9), 768–784. https://doi.org/10.1002/jms.3374.

(48) Han, X.; Yang, K.; Gross, R. W. Multi-Dimensional Mass Spectrometry-Based Shotgun Lipidomics and Novel Strategies for Lipidomic Analyses. *Mass Spectrom. Rev.* **2012**, *31* (1), 134–178. https://doi.org/10.1002/mas.20342.

(49) Michaud, S. A.; Pětrošová, H.; Sinclair, N. J.; Kinnear, A. L.; Jackson, A. M.; McGuire, J. C.; Hardie, D. B.; Bhowmick, P.; Ganguly, M.; Flenniken, A. M.; Nutter, L. M. J.; McKerlie, C.; Smith, D.; Mohammed, Y.; Schibli, D.; Sickmann, A.; Borchers, C. H. Multiple Reaction Monitoring Assays for Large-Scale Quantitation of Proteins from 20 Mouse Organs and Tissues. *Commun. Biol.* **2024**, *7* (1), 1–12. https://doi.org/10.1038/s42003-023-05687-0.

(50) Delvaux, A.; Rathahao-Paris, E.; Alves, S. Different Ion Mobility-Mass Spectrometry Coupling Techniques to Promote Metabolomics. *Mass Spectrom. Rev.* **2022**, *41* (5), 695–721. https://doi.org/10.1002/mas.21685.

(51) Morris, C. B.; May, J. C.; Leaptrot, K. L.; McLean, J. A. Evaluating Separation Selectivity and Collision Cross Section Measurement Reproducibility in Helium, Nitrogen, Argon, and Carbon Dioxide Drift Gases for Drift Tube Ion Mobility–Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 1059–1068. https://doi.org/10.1007/s13361-019-02151-4.

(52) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan Jr, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry

Measurements. *Mass Spectrom. Rev.* **2019**, *38* (3), 291–320. https://doi.org/10.1002/mas.21585.

(53) Revercomb, H. E.; Mason, E. A. Theory of Plasma Chromatography/Gaseous Electrophoresis. Review. *Anal. Chem.* **1975**, *47* (7), 970–983. https://doi.org/10.1021/ac60357a043.

(54) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51–59. https://doi.org/10.1016/j.cbpa.2017.10.022.

(55) May, J. C.; McLean, J. A. The Influence of Drift Gas Composition on the Separation Mechanism in Traveling Wave Ion Mobility Spectrometry: Insight from Electrodynamic Simulations. *Int. J. Ion Mobil. Spectrom.* **2003**, *16* (2), 85–94. https://doi.org/10.1007/s12127-013-0123-7.

(56) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 1028–1037. https://doi.org/10.1007/s13361-019-02168-9.

(57) Griffiths, D.; Anderson, M.; Richardson, K.; Inaba-Inoue, S.; Allen, W. J.; Collinson, I.; Beis, K.; Morris, M.; Giles, K.; Politis, A. Cyclic Ion Mobility for Hydrogen/Deuterium Exchange-Mass Spectrometry Applications. *Anal. Chem.* **2024**. https://doi.org/10.1021/acs.analchem.3c05753.

(58) Morsa, D.; Gabelica, V.; De Pauw, E. Effective Temperature of Ions in Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **2011**, *83* (14), 5775–5782. https://doi.org/10.1021/ac201509p.

(59) Nishimura, T. Unimolecular Dissociations in Mass Spectrometry. In *Fundamentals of Mass Spectrometry*; Hiraoka, K., Ed.; Springer: New York, NY, 2013; pp 29–54. https://doi.org/10.1007/978-1-4614-7233-9_2.

(60) Bayat, P.; Lesage, D.; Cole, R. B. Tutorial: Ion Activation in Tandem Mass Spectrometry Using Ultra-High Resolution Instrumentation. *Mass Spectrom. Rev.* **2020**, *39* (5–6), 680–702. https://doi.org/10.1002/mas.21623.

(61) Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in Travelling Wave Ion Mobility Resolution. *Rapid Commun. Mass Spectrom.* **2011**, *25* (11), 1559–1566. https://doi.org/10.1002/rcm.5013.

(62) Zhang, Q.; Frolov, A.; Tang, N.; Hoffmann, R.; van de Goor, T.; Metz, T. O.; Smith, R. D. Application of Electron Transfer Dissociation Mass Spectrometry in Analyses of Non-Enzymatically Glycated Peptides. *Rapid Commun. Mass Spectrom. RCM* **2007**, *21* (5), 661–666. https://doi.org/10.1002/rcm.2884.

(63) Hjernø, K.; Højrup, P. Interpretation of Tandem Mass Spectrometry (MSMS) Spectra for Peptide Analysis. In *Peptide Antibodies*; Houen, G., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2015; Vol. 1348, pp 83–102. https://doi.org/10.1007/978-1-4939-2999-3_10.

(64) Boyd, R.; Somogyi, Á. The Mobile Proton Hypothesis in Fragmentation of Protonated Peptides: A Perspective. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (8), 1275–1278. https://doi.org/10.1016/j.jasms.2010.04.017.

(65) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406. https://doi.org/10.1002/1096-9888(200012)35:12<1399::AID-JMS86>3.0.CO;2-R.

(66) Hydrogen Exchange: A Sensitive Analytical Window into Protein Conformation and Dynamics. In *Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods, and Applications*; Weis, D. D., Ed.; John Wiley & Sons, Ltd, 2016; pp 1–15.

(67) Perrin, C. L. Proton Exchange in Amides: Surprises from Simple Systems. *Acc. Chem. Res.* **1989**, *22* (8), 268–275. https://doi.org/10.1021/ar00164a002.

(68) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **1993**, *17* (1), 75–86. https://doi.org/10.1002/prot.340170110.

(69) Smith, D. L.; Deng, Y.; Zhang, Z. Probing the Non-Covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry. *J. Mass Spectrom.* **1997**, *32* (2), 135–146. https://doi.org/10.1002/(SICI)1096-9888(199702)32:2<135::AID-JMS486>3.0.CO;2-M.

(70) Covington, A. K.; Robinson, R. A.; Bates, R. G. The Ionization Constant of Deuterium Oxide from 5 to 50 Degrees. *J Phys Chem 70 3820-4Dec 15 1966* **1966**. https://doi.org/10.1021/j100884a011.

(71) Covington, A. K.; Paabo, Maya.; Robinson, R. Anthony.; Bates, R. G. Use of the Glass Electrode in Deuterium Oxide and the Relation between the Standardized pD (paD) Scale and the Operational pH in Heavy Water. *Anal. Chem.* **1968**, *40* (4), 700–706. https://doi.org/10.1021/ac60260a013.

(72) Hamuro, Y. Tutorial: Chemistry of Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (1), 133–151. https://doi.org/10.1021/jasms.0c00260.

(73) Molday, R. S.; Englander, S. W.; Kallen, R. G. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Biochemistry* **1972**, *11* (2), 150–158. https://doi.org/10.1021/bi00752a003.

(74) Klotz, I. M.; Tam, J. W. O. Nuclear Magnetic Resonance Studies of Poly-DL-Alanine and Poly-L-Alanine in Solvents with Strong Acids. *J. Am. Chem. Soc.* **1971**, *93* (5), 1313–1315. https://doi.org/10.1021/ja00734a075.

(75) Bandura, A. V.; Lvov, S. N. The Ionization Constant of Water over Wide Ranges of Temperature and Density. *J. Phys. Chem. Ref. Data* **2005**, *35* (1), 15–30. https://doi.org/10.1063/1.1928231.

(76) Englander, S. W.; Downer, N. W.; Teitelbaum, H. Hydrogen Exchange. *Annu. Rev. Biochem.* **1972**, *41*, 903–924. https://doi.org/10.1146/annurev.bi.41.070172.004351.

(77) Schellman, J. A.; Schellman, C. G. Kaj Ulrik Linderstrøm-Lang (1896–1959). *Protein Sci.* **1997**, *6* (5), 1092–1100. https://doi.org/10.1002/pro.5560060516.

(78) Berger, A.; Linderstrøm-Lang, K. Deuterium Exchange of Poly-Dl-Alanine in Aqueous Solution. *Arch. Biochem. Biophys.* **1957**, *69*, 106–118. https://doi.org/10.1016/0003-9861(57)90478-2.

(79) Pauling, L.; Corey, R. B.; Branson, H. R. The Structure of Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37* (4), 205–211.

(80) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. Protein Analysis by Hydrogen Exchange Mass Spectrometry. *Annu. Rev. Biophys.* **2003**, *32* (Volume 32, 2003), 1–25. https://doi.org/10.1146/annurev.biophys.32.110601.142417.

(81) Englander, S. W.; Kallenbach, N. R. Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic Acids. *Q. Rev. Biophys.* **1983**, *16* (4), 521–655. https://doi.org/10.1017/s0033583500005217.

(82) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. In *Advances in Protein Chemistry*; Anfinsen, C. B., Anson, M. L., Edsall, J. T., Richards, F. M., Eds.; Academic Press, 1966; Vol. 21, pp 287–386. https://doi.org/10.1016/S0065-3233(08)60129-1.

(83) Weis, D. D.; Wales, T. E.; Engen, J. R.; Hotchko, M.; Ten Eyck, L. F. Identification and Characterization of EX1 Kinetics in H/D Exchange Mass Spectrometry by Peak Width Analysis. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (11), 1498–1509. https://doi.org/10.1016/j.jasms.2006.05.014.

(84) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* **2009**, *131* (35), 12801–12808. https://doi.org/10.1021/ja904379w.

(85) Hamuro, Y. Quantitative Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (12), 2711–2727. https://doi.org/10.1021/jasms.1c00216.

(86) Comamala, G.; Wagner, C.; de la Torre, P. S.; Jakobsen, R. U.; Hilger, M.; Brouwer, H.-J.; Rand, K. D. Hydrogen/Deuterium Exchange Mass Spectrometry with Improved Electrochemical Reduction Enables Comprehensive Epitope Mapping of a Therapeutic Antibody to the Cysteine-Knot Containing Vascular Endothelial Growth Factor. *Anal. Chim. Acta* **2020**, *1115*, 41–51. https://doi.org/10.1016/j.aca.2020.04.014.

(87) Paris, A.; Verbitskiy, N.; Nefedov, A.; Wang, Y.; Fedorov, A.; Haberer, D.; Oehzelt, M.; Petaccia, L.; Usachov, D.; Vyalikh, D.; Sachdev, H.; Wöll, C.; Knupfer, M.; Büchner, B.; Calliari, L.; Yashina, L.; Irle, S.; Grüneis, A. Kinetic Isotope Effect in the Hydrogenation and Deuteration of Graphene. *Adv. Funct. Mater.* **2013**, *23* (13), 1628–1635. https://doi.org/10.1002/adfm.201202355.

(88) Shi, C.; Zhang, X.; Yu, C.-H.; Yao, Y.-F.; Zhang, W. Geometric Isotope Effect of Deuteration in a Hydrogen-Bonded Host–Guest Crystal. *Nat. Commun.* **2018**, *9* (1), 481. https://doi.org/10.1038/s41467-018-02931-8.

(89) Pohl, R.; Amaro, F. D.; Antognini, A.; Biraben, F.; Cardoso, J. M. R.; Covita, D. S.; Dax, A.; Dhawan, S.; Fernandes, L. M. P.; Giesen, A.; Graf, T.; Hänsch, T. W.; Indelicato, P.; Julien, L.; Kao, C.-Y.; Knowles, P.; Bigot, E.-O. L.; Liu, Y.-W.; Lopes, J. A. M.; Ludhova, L.; Monteiro, C. M. B.; Mulhauser, F.; Nebel, T.; Nez, F.; Rabinowitz, P.; Santos, J. M. F. dos; Schaller, L. A.;

Schuhmann, K.; Schwob, C.; Taqqu, D.; Veloso, J. F. C. A.; Kottmann, F. The Size of the Proton and the Deuteron. *J. Phys. Conf. Ser.* **2011**, *264* (1), 012008. https://doi.org/10.1088/1742-6596/264/1/012008.

(90) Grocholska, P.; Bąchor, R. Trends in the Hydrogen−Deuterium Exchange at the Carbon Centers. Preparation of Internal Standards for Quantitative Analysis by LC-MS. *Molecules* **2021**, *26* (10), 2989. https://doi.org/10.3390/molecules26102989.

(91) Connelly, G. P.; Bai, Y.; Jeng, M.-F.; Englander, S. W. Isotope effects in peptide group hydrogen exchange. *Proteins Struct. Funct. Bioinforma.* **1993**, *17* (1), 87–92. https://doi.org/10.1002/prot.340170111.

(92) Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R. Binding of a Reduced Peptide Inhibitor to the Aspartic Proteinase from Rhizopus Chinensis: Implications for a Mechanism of Action. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (20), 7009–7013.

(93) KUBOTA, K.; METOKI, Y.; ATHAUDA, S. B. P.; SHIBATA, C.; TAKAHASHI, K. Stability Profiles of Nepenthesin in Urea and Guanidine Hydrochloride: Comparison with Porcine Pepsin A. *Biosci. Biotechnol. Biochem.* **2010**, *74* (11), 2323–2326. https://doi.org/10.1271/bbb.100391.

(94) Yang, M.; Hoeppner, M.; Rey, M.; Kadek, A.; Man, P.; Schriemer, D. C. Recombinant Nepenthesin II for Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2015**, *87* (13), 6681–6687. https://doi.org/10.1021/acs.analchem.5b00831.

(95) Žuvela, P.; Skoczylas, M.; Jay Liu, J.; Bączek, T.; Kaliszan, R.; Wong, M. W.; Buszewski, B. Column Characterization and Selection Systems in Reversed-Phase High-Performance Liquid Chromatography. *Chem. Rev.* **2019**, *119* (6), 3674–3729. https://doi.org/10.1021/acs.chemrev.8b00246.

(96) Xu, A. M.; Tang, L. C.; Jovanovic, M.; Regev, O. Uncovering Distinct Peptide Charging Behaviors in Electrospray Ionization Mass Spectrometry Using a Large-Scale Dataset. *J. Am. Soc. Mass Spectrom.* **2024**, *35* (1), 90–99. https://doi.org/10.1021/jasms.3c00325.

(97) 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.

(98) Narang, D.; James, D. A.; Balmer, M. T.; Wilson, D. J. Protein Footprinting, Conformational Dynamics, and Core Interface-Adjacent Neutralization “Hotspots” in the SARS-CoV-2 Spike Protein Receptor Binding Domain/Human ACE2 Interaction. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (7), 1593–1600. https://doi.org/10.1021/jasms.0c00465.

(99) Analysis of Disordered Proteins by Hydrogen Exchange Mass Spectrometry. In *Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods, and Applications*; Weis, D. D., Ed.; John Wiley & Sons, Ltd, 2016; pp 295–321.

(100) Chow, V.; Wolf, E.; Lento, C.; Wilson, D. J. Developments in Rapid Hydrogen-Deuterium Exchange Methods. *Essays Biochem.* **2023**, *67* (2), 165–174. https://doi.org/10.1042/EBC20220174.

(101) Kish, M.; Smith, V.; Lethbridge, N.; Cole, L.; Bond, Nicholas. J.; Phillips, J. J. Online Fully Automated System for Hydrogen/Deuterium-Exchange Mass Spectrometry with Millisecond Time Resolution. *Anal. Chem.* **2023**, *95* (11), 5000–5008. https://doi.org/10.1021/acs.analchem.2c05310.

(102) Wilson, D. J.; Konermann, L. A Capillary Mixer with Adjustable Reaction Chamber Volume for Millisecond Time-Resolved Studies by Electrospray Mass Spectrometry. *Anal. Chem.* **2003**, *75* (23), 6408–6414. https://doi.org/10.1021/ac0346757.

(103) Resetca, D.; Wilson, D. J. Mapping Ligand Binding Using Microfluidics-Enabled Millisecond Timescale Hydrogen-Deuterium Exchange. *Int. J. Mass Spectrom.* **2017**, *420*, 67–73. https://doi.org/10.1016/j.ijms.2017.06.002.

(104) Lento, C.; Zhu, S.; Brown, K. A.; Knox, R.; Liuni, P.; Wilson, D. J. Time-Resolved ElectroSpray Ionization Hydrogen-Deuterium Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *JoVE J. Vis. Exp.* **2017**, No. 122, e55464. https://doi.org/10.3791/55464.

(105) Rob, T.; Gill, P. K.; Golemi-Kotra, D.; Wilson, D. J. An Electrospray Ms-Coupled Microfluidic Device for Sub-Second Hydrogen/Deuterium Exchange Pulse-Labelling Reveals Allosteric Effects in Enzyme Inhibition. *Lab. Chip* **2013**, *13* (13), 2528–2532. https://doi.org/10.1039/C3LC00007A.

(106) 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.

(107) Liuni, P.; Rob, T.; Wilson, D. J. A Microfluidic Reactor for Rapid, Low-Pressure Proteolysis with on-Chip Electrospray Ionization. *Rapid Commun. Mass Spectrom.* **2010**, *24* (3), 315–320. https://doi.org/10.1002/rcm.4391.

(108) Rob, T. A Versatile Microfluidic Chip for Millisecond Time-Scale Kinetic Studies by Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (1), 124–130.

(109) Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral pH “Buffer” for Native Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1827–1835. https://doi.org/10.1007/s13361-017-1739-3.

(110) 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.

(111) Kale, J.; Osterlund, E. J.; Andrews, D. W. BCL-2 Family Proteins: Changing Partners in the Dance towards Death. *Cell Death Differ.* **2018**, *25* (1), 65–80. https://doi.org/10.1038/cdd.2017.186.

(112) Luo, X.; Budihardjo, I.; Zou, H.; Slaughter, C.; Wang, X. Bid, a Bcl2 Interacting Protein, Mediates Cytochrome c Release from Mitochondria in Response to Activation of Cell Surface Death Receptors. *Cell* **1998**, *94* (4), 481–490. https://doi.org/10.1016/s0092-8674(00)81589-5.

(113) Westphal, D.; Kluck, R. M.; Dewson, G. Building Blocks of the Apoptotic Pore: How Bax and Bak Are Activated and Oligomerize during Apoptosis. *Cell Death Differ.* **2014**, *21* (2), 196–205. https://doi.org/10.1038/cdd.2013.139.

(114) Baliga, B.; Kumar, S. Apaf-1/Cytochrome c Apoptosome: An Essential Initiatorof Caspase Activation or Just a Sideshow? *Cell Death Differ.* **2003**, *10* (1), 16–18. https://doi.org/10.1038/sj.cdd.4401166.

(115) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25* (17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

(116) Altschul, S. F.; Wootton, J. C.; Gertz, E. M.; Agarwala, R.; Morgulis, A.; Schäffer, A. A.; Yu, Y.-K. Protein Database Searches Using Compositionally Adjusted Substitution Matrices. *FEBS J.* **2005**, *272* (20), 5101–5109. https://doi.org/10.1111/j.1742-4658.2005.04945.x.

(117) Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8, 2015.

(118) Petros, A. M.; Medek, A.; Nettesheim, D. G.; Kim, D. H.; Yoon, H. S.; Swift, K.; Matayoshi, E. D.; Oltersdorf, T.; Fesik, S. W. Solution Structure of the Antiapoptotic Protein Bcl-2. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (6), 3012–3017. https://doi.org/10.1073/pnas.041619798.

(119) Liu, G.; Poppe, L.; Aoki, K.; Yamane, H.; Lewis, J.; Szyperski, T. High-Quality NMR Structure of Human Anti-Apoptotic Protein Domain Mcl-1(171-327) for Cancer Drug Design. *PLOS ONE* **2014**, *9* (5), e96521. https://doi.org/10.1371/journal.pone.0096521.

(120) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chem. Rev.* **2020**, *120* (10), 4355–4454. https://doi.org/10.1021/acs.chemrev.9b00815.

(121) 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.

(122) Lento, C.; Wilson, D. J. Subsecond Time-Resolved Mass Spectrometry in Dynamic Structural Biology. *Chem. Rev.* **2022**, *122* (8), 7624–7646. https://doi.org/10.1021/acs.chemrev.1c00222.

(123) Lee, S.; Wales, T. E.; Escudero, S.; Cohen, D. T.; Luccarelli, J.; Gallagher, C. G.; Cohen, N. A.; Huhn, A. J.; Bird, G. H.; Engen, J. R.; Walensky, L. D. Allosteric Inhibition of Antiapoptotic MCL-1. *Nat. Struct. Mol. Biol.* **2016**, *23* (6), 600–607. https://doi.org/10.1038/nsmb.3223.

(124) Day, C. L.; Chen, L.; Richardson, S. J.; Harrison, P. J.; Huang, D. C. S.; Hinds, M. G. Solution Structure of Prosurvival Mcl-1 and Characterization of Its Binding by Proapoptotic BH3-Only Ligands*. *J. Biol. Chem.* **2005**, *280* (6), 4738–4744. https://doi.org/10.1074/jbc.M411434200.

(125) Wilson, D. J.; Konermann, L. A Capillary Mixer with Adjustable Reaction Chamber Volume for Millisecond Time-Resolved Studies by Electrospray Mass Spectrometry. *Anal. Chem.* **2003**, *75* (23), 6408–6414. https://doi.org/10.1021/ac0346757.

(126) Lento, C.; Zhu, S.; Brown, K. A.; Knox, R.; Liuni, P.; Wilson, D. J. Time-Resolved ElectroSpray Ionization Hydrogen-Deuterium Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *J. Vis. Exp. JoVE* **2017**, No. 122. https://doi.org/10.3791/55464.

(127) Rey, M.; Sarpe, V.; Burns, K. M.; Buse, J.; Baker, C. A. H.; van Dijk, M.; Wordeman, L.; Bonvin, A. M. J. J.; Schriemer, D. C. Mass Spec Studio for Integrative Structural Biology. *Struct. Lond. Engl. 1993* **2014**, *22* (10), 1538–1548. https://doi.org/10.1016/j.str.2014.08.013.

(128) Gattiker, A.; Bienvenut, W. V.; Bairoch, A.; Gasteiger, E. FindPept, a Tool to Identify Unmatched Masses in Peptide Mass Fingerprinting Protein Identification. *PROTEOMICS* **2002**, *2* (10), 1435–1444. https://doi.org/10.1002/1615-9861(200210)2:10<1435::AID-PROT1435>3.0.CO;2-9.

(129) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46* (W1), W296–W303. https://doi.org/10.1093/nar/gky427.

(130) Studer, G.; Rempfer, C.; Waterhouse, A. M.; Gumienny, R.; Haas, J.; Schwede, T. QMEANDisCo—Distance Constraints Applied on Model Quality Estimation. *Bioinformatics* **2020**, *36* (6), 1765–1771. https://doi.org/10.1093/bioinformatics/btz828.

(131) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; Thompson, C. B.; Fesik, S. W. Structure of Bcl-xL-Bak Peptide Complex: Recognition Between Regulators of Apoptosis. *Science* **1997**. https://doi.org/10.1126/science.275.5302.983.

(132) Pu, J.; Dewey, J. A.; Hadji, A.; LaBelle, J. L.; Dickinson, B. C. RNA Polymerase Tags To Monitor Multidimensional Protein–Protein Interactions Reveal Pharmacological Engagement of Bcl-2 Proteins. *J. Am. Chem. Soc.* **2017**, *139* (34), 11964–11972. https://doi.org/10.1021/jacs.7b06152.

(133) Carlson, J. C.; Badran, A. H.; Guggiana-Nilo, D. A.; Liu, D. R. Negative Selection and Stringency Modulation in Phage-Assisted Continuous Evolution. *Nat. Chem. Biol.* **2014**, *10* (3), 216–222. https://doi.org/10.1038/nchembio.1453.

(134) Certo, M.; Del Gaizo Moore, V.; Nishino, M.; Wei, G.; Korsmeyer, S.; Armstrong, S. A.; Letai, A. Mitochondria Primed by Death Signals Determine Cellular Addiction to Antiapoptotic BCL-2 Family Members. *Cancer Cell* **2006**, *9* (5), 351–365. https://doi.org/10.1016/j.ccr.2006.03.027.

(135) Letai, A.; Bassik, M. C.; Walensky, L. D.; Sorcinelli, M. D.; Weiler, S.; Korsmeyer, S. J. Distinct BH3 Domains Either Sensitize or Activate Mitochondrial Apoptosis, Serving as Prototype Cancer Therapeutics. *Cancer Cell* **2002**, *2* (3), 183–192. https://doi.org/10.1016/S1535-6108(02)00127-7.

(136) Smith, A. J.; Dai, H.; Correia, C.; Takahashi, R.; Lee, S.-H.; Schmitz, I.; Kaufmann, S. H. Noxa/Bcl-2 Protein Interactions Contribute to Bortezomib Resistance in Human Lymphoid Cells. *J. Biol. Chem.* **2011**, *286* (20), 17682–17692. https://doi.org/10.1074/jbc.M110.189092.

(137) Chen, L.; Willis, S. N.; Wei, A.; Smith, B. J.; Fletcher, J. I.; Hinds, M. G.; Colman, P. M.; Day, C. L.; Adams, J. M.; Huang, D. C. S. Differential Targeting of Prosurvival Bcl-2 Proteins by Their BH3-Only Ligands Allows Complementary Apoptotic Function. *Mol. Cell* **2005**, *17* (3), 393–403. https://doi.org/10.1016/j.molcel.2004.12.030.

(138) Dai, H.; Pang, Y.-P.; Ramirez-Alvarado, M.; Kaufmann, S. H. Evaluation of the BH3-Only Protein Puma as a Direct Bak Activator. *J. Biol. Chem.* **2014**, *289* (1), 89–99. https://doi.org/10.1074/jbc.M113.505701.

(139) Du, Y.; Nikolovska-Coleska, Z.; Qui, M.; Li, L.; Lewis, I.; Dingledine, R.; Stuckey, J. A.; Krajewski, K.; Roller, P. P.; Wang, S.; Fu, H. A Dual-Readout F2 Assay That Combines Fluorescence Resonance Energy Transfer and Fluorescence Polarization for Monitoring Bimolecular Interactions. *Assay Drug Dev. Technol.* **2011**, *9* (4), 382–393. https://doi.org/10.1089/adt.2010.0292.

(140) Liu, Q.; Moldoveanu, T.; Sprules, T.; Matta-Camacho, E.; Mansur-Azzam, N.; Gehring, K. Apoptotic Regulation by MCL-1 through Heterodimerization. *J. Biol. Chem.* **2010**, *285* (25), 19615–19624. https://doi.org/10.1074/jbc.M110.105452.

(141) Bhat, V.; Olenick, M. B.; Schuchardt, B. J.; Mikles, D. C.; McDonald, C. B.; Farooq, A. Biophysical Basis of the Promiscuous Binding of B-Cell Lymphoma Protein 2 Apoptotic Repressor to BH3 Ligands. *J. Mol. Recognit. JMR* **2013**, *26* (10), 501–513. https://doi.org/10.1002/jmr.2295.

(142) Ku, B.; Liang, C.; Jung, J. U.; Oh, B.-H. Evidence That Inhibition of BAX Activation by BCL-2 Involves Its Tight and Preferential Interaction with the BH3 Domain of BAX. *Cell Res.* **2011**, *21* (4), 627–641. https://doi.org/10.1038/cr.2010.149.

(143) Day, C. L.; Smits, C.; Fan, F. C.; Lee, E. F.; Fairlie, W. D.; Hinds, M. G. Structure of the BH3 Domains from the P53-Inducible BH3-Only Proteins Noxa and Puma in Complex with Mcl-1. *J. Mol. Biol.* **2008**, *380* (5), 958–971. https://doi.org/10.1016/j.jmb.2008.05.071.

(144) Czabotar, P. E.; Lee, E. F.; van Delft, M. F.; Day, C. L.; Smith, B. J.; Huang, D. C. S.; Fairlie, W. D.; Hinds, M. G.; Colman, P. M. Structural Insights into the Degradation of Mcl-1 Induced by BH3 Domains. *Proc. Natl. Acad. Sci.* **2007**, *104* (15), 6217–6222. https://doi.org/10.1073/pnas.0701297104.

(145) Anasir, M. I.; Caria, S.; Skinner, M. A.; Kvansakul, M. Structural Basis of Apoptosis Inhibition by the Fowlpox Virus Protein FPV039. *J. Biol. Chem.* **2017**, *292* (22), 9010–9021. https://doi.org/10.1074/jbc.M116.768879.

(146) Dutta, S.; Chen, T. S.; Keating, A. E. Peptide Ligands for Pro-Survival Protein Bfl-1 from Computationally Guided Library Screening. *ACS Chem. Biol.* **2013**, *8* (4), 778–788. https://doi.org/10.1021/cb300679a.

(147) Foight, G. W.; Ryan, J. A.; Gullá, S. V.; Letai, A.; Keating, A. E. Designed BH3 Peptides with High Affinity and Specificity for Targeting Mcl-1 in Cells. *ACS Chem. Biol.* **2014**, *9* (9), 1962–1968. https://doi.org/10.1021/cb500340w.

(148) Winkler, R. ESIprot: A Universal Tool for Charge State Determination and Molecular Weight Calculation of Proteins from Electrospray Ionization Mass Spectrometry Data. *Rapid Commun. Mass Spectrom.* **2010**, *24* (3), 285–294. https://doi.org/10.1002/rcm.4384.

(149) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein–Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 431–441. https://doi.org/10.1007/s13361-011-0311-9.

(150) Macias, L. A.; Santos, I. C.; Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **2020**, *92* (1), 227–251. https://doi.org/10.1021/acs.analchem.9b04859.

(151) 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.

(152) Denis, C.; Sopková-de Oliveira Santos, J.; Bureau, R.; Voisin-Chiret, A. S. Hot-Spots of Mcl-1 Protein. *J. Med. Chem.* **2020**, *63* (3), 928–943. https://doi.org/10.1021/acs.jmedchem.9b00983.

(153) Petros, A. M.; Olejniczak, E. T.; Fesik, S. W. Structural Biology of the Bcl-2 Family of Proteins. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2004**, *1644* (2), 83–94. https://doi.org/10.1016/j.bbamcr.2003.08.012.

(154) Banjara, S.; Suraweera, C. D.; Hinds, M. G.; Kvansakul, M. The Bcl-2 Family: Ancient Origins, Conserved Structures, and Divergent Mechanisms. *Biomolecules* **2020**, *10* (1), 128. https://doi.org/10.3390/biom10010128.

(155) 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.

(156) Huang, X.; Slavkovic, S.; Song, E.; Botta, A.; Mehrazma, B.; Lento, C.; Johnson, P. E.; Sweeney, G.; Wilson, D. J. A Unique Conformational Distortion Mechanism Drives Lipocalin 2 Binding to Bacterial Siderophores. *ACS Chem. Biol.* **2020**, *15* (1), 234–242. https://doi.org/10.1021/acschembio.9b00820.

(157) Campbell, S. T.; Carlson, K. J.; Buchholz, C. J.; Helmers, M. R.; Ghosh, I. Mapping the BH3 Binding Interface of Bcl-xL, Bcl-2, and Mcl-1 Using Split-Luciferase Reassembly. *Biochemistry* **2015**, *54* (16), 2632–2643. https://doi.org/10.1021/bi501505y.

(158) Zhao, R.-N.; Fan, S.; Han, J.-G.; Liu, G. Molecular Dynamics Study of Segment Peptides of Bax, Bim, and Mcl-1 BH3 Domain of the Apoptosis-Regulating Proteins Bound to the Anti-Apoptotic Mcl-1 Protein. *J. Biomol. Struct. Dyn.* **2015**, *33* (5), 1067–1081. https://doi.org/10.1080/07391102.2014.929028.

(159) Fogha, J.; Marekha, B.; De Giorgi, M.; Voisin-Chiret, A. S.; Rault, S.; Bureau, R.; Sopkova-de Oliveira Santos, J. Toward Understanding Mcl-1 Promiscuous and Specific Binding Mode. *J. Chem. Inf. Model.* **2017**, *57* (11), 2885–2895. https://doi.org/10.1021/acs.jcim.7b00396.

(160) Ivanov, S. M.; Huber, R. G.; Warwicker, J.; Bond, P. J. Energetics and Dynamics Across the Bcl-2-Regulated Apoptotic Pathway Reveal Distinct Evolutionary Determinants of Specificity

and Affinity. *Struct. Lond. Engl. 1993* **2016**, *24* (11), 2024–2033. https://doi.org/10.1016/j.str.2016.09.006.

(161) Raghav, P. K.; Verma, Y. K.; Gangenahalli, G. U. Molecular Dynamics Simulations of the Bcl-2 Protein to Predict the Structure of Its Unordered Flexible Loop Domain. *J. Mol. Model.* **2012**, *18* (5), 1885–1906. https://doi.org/10.1007/s00894-011-1201-6.

(162) Zhu, S.; Khatun, R.; Lento, C.; Sheng, Y.; Wilson, D. J. Enhanced Binding Affinity via Destabilization of the Unbound State: A Millisecond Hydrogen–Deuterium Exchange Study of the Interaction between P53 and a Pleckstrin Homology Domain. *Biochemistry* **2017**, *56* (32), 4127–4133. https://doi.org/10.1021/acs.biochem.7b00193.

(163) Xie, V. C.; Pu, J.; Metzger, B. P.; Thornton, J. W.; Dickinson, B. C. Contingency and Chance Erase Necessity in the Experimental Evolution of Ancestral Proteins. *eLife* **2021**, *10*, e67336. https://doi.org/10.7554/eLife.67336.

(164) Hughes, J.; Rees, S.; Kalindjian, S.; Philpott, K. Principles of Early Drug Discovery. *Br. J. Pharmacol.* **2011**, *162* (6), 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x.

(165) *Hydrogen Exchange Mass Spectrometry of Proteins*, 1st ed.; Weis, D. D., Ed.; John Wiley & Sons, Ltd, 2016. https://doi.org/10.1002/9781118703748.

(166) 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.

(167) Englander, S. Walter.; Kallenbach, N. Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic Acids. *Q. Rev. Biophys.* **1983**, *16* (4), 521–655. https://doi.org/10.1017/s0033583500005217.

(168) Deng, B.; Lento, C.; Wilson, D. J. Hydrogen Deuterium Exchange Mass Spectrometry in Biopharmaceutical Discovery and Development – A Review. *Anal. Chim. Acta* **2016**, *940*, 8–20. https://doi.org/10.1016/j.aca.2016.08.006.

(169) Oganesyan, I.; Lento, C.; Wilson, D. J. Contemporary Hydrogen Deuterium Exchange Mass Spectrometry. *Methods* **2018**, *144*, 27–42. https://doi.org/10.1016/j.ymeth.2018.04.023.

(170) Marciano, D. P.; Dharmarajan, V.; Griffin, P. R. HDX-MS Guided Drug Discovery: Small Molecules and Biopharmaceuticals. *Curr. Opin. Struct. Biol.* **2014**, *0*, 105–111. https://doi.org/10.1016/j.sbi.2014.08.007.

(171) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R. Differential Hydrogen/Deuterium Exchange Mass Spectrometry Analysis of Protein–Ligand Interactions. *Expert Rev. Proteomics* **2011**, *8* (1), 43–59. https://doi.org/10.1586/epr.10.109.

(172) Engen, J. R. Analysis of Protein Complexes with Hydrogen Exchange and Mass Spectrometry. *Analyst* **2003**, *128* (6), 623–628. https://doi.org/10.1039/B212800B.

(173) Kochert, B. A.; Iacob, R. E.; Wales, T. E.; Makriyannis, A.; Engen, J. R. Hydrogen-Deuterium Exchange Mass Spectrometry to Study Protein Complexes. In *Protein Complex Assembly*; Marsh, J. A., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2018; Vol. 1764, pp 153–171. https://doi.org/10.1007/978-1-4939-7759-8_10.

(174) Huang, R. Y.-C.; Rempel, D. L.; Gross, M. L. HD Exchange and PLIMSTEX Determine the Affinities and Order of Binding of Ca2+ with Troponin C. *Biochemistry* **2011**, *50* (24), 5426–5435. https://doi.org/10.1021/bi200377c.

(175) Zhu, M. M.; Rempel, D. L.; Du, Z.; Gross, M. L. Quantification of Protein-Ligand Interactions by Mass Spectrometry, Titration, and H/D Exchange: PLIMSTEX. *J. Am. Chem. Soc.* **2003**, *125* (18), 5252–5253. https://doi.org/10.1021/ja029460d.

(176) Jørgensen, T. J. D.; Gårdsvoll, H.; Danø, K.; Roepstorff, P.; Ploug, M. Dynamics of Urokinase Receptor Interaction with Peptide Antagonists Studied by Amide Hydrogen Exchange and Mass Spectrometry. *Biochemistry* **2004**, *43* (47), 15044–15057. https://doi.org/10.1021/bi048706j.

(177) Zhang, Z.; Vachet, R. W. Kinetics of Protein Complex Dissociation Studied by Hydrogen/Deuterium Exchange and Mass Spectrometry. *Anal. Chem.* **2015**, *87* (23), 11777–11783. https://doi.org/10.1021/acs.analchem.5b03123.

(178) Hamuro, Y.; Coales, S. J. Optimization of Feasibility Stage for Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (3), 623–629. https://doi.org/10.1007/s13361-017-1860-3.

(179) Hamuro, Y.; Coales, S. J. Hydrogen/Deuterium Exchange Mass Spectrometry for Weak Binders. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (4), 735–739. https://doi.org/10.1021/jasms.1c00375.

(180) Wysocka, J.; Swigut, T.; Milne, T. A.; Dou, Y.; Zhang, X.; Burlingame, A. L.; Roeder, R. G.; Brivanlou, A. H.; Allis, C. D. WDR5 Associates with Histone H3 Methylated at K4 and Is Essential for H3 K4 Methylation and Vertebrate Development. *Cell* **2005**, *121* (6), 859–872. https://doi.org/10.1016/j.cell.2005.03.036.

(181) Park, S. H.; Ayoub, A.; Lee, Y.-T.; Xu, J.; Kim, H.; Zheng, W.; Zhang, B.; Sha, L.; An, S.; Zhang, Y.; Cianfrocco, M. A.; Su, M.; Dou, Y.; Cho, U.-S. Cryo-EM Structure of the Human MLL1 Core Complex Bound to the Nucleosome. *Nat. Commun.* **2019**, *10* (1), 5540. https://doi.org/10.1038/s41467-019-13550-2.

(182) Grebien, F.; Vedadi, M.; Getlik, M.; Giambruno, R.; Grover, A.; Avellino, R.; Skucha, A.; Vittori, S.; Kuznetsova, E.; Smil, D.; Barsyte-Lovejoy, D.; Li, F.; Poda, G.; Schapira, M.; Wu, H.; Dong, A.; Senisterra, G.; Stukalov, A.; Huber, K. V. M.; Schönegger, A.; Marcellus, R.; Bilban, M.; Bock, C.; Brown, P. J.; Zuber, J.; Bennett, K. L.; Al-awar, R.; Delwel, R.; Nerlov, C.; Arrowsmith, C. H.; Superti-Furga, G. Pharmacological Targeting of the Wdr5-MLL Interaction in C/EBPα N-Terminal Leukemia. *Nat. Chem. Biol.* **2015**, *11* (8), 571–578. https://doi.org/10.1038/nchembio.1859.

(183) Xue, H.; Yao, T.; Cao, M.; Zhu, G.; Li, Y.; Yuan, G.; Chen, Y.; Lei, M.; Huang, J. Structural Basis of Nucleosome Recognition and Modification by MLL Methyltransferases. *Nature* **2019**, *573* (7774), 445–449. https://doi.org/10.1038/s41586-019-1528-1.

(184) Ayoub, A.; Park, S. H.; Lee, Y.-T.; Cho, U.-S.; Dou, Y. Regulation of MLL1 Methyltransferase Activity in Two Distinct Nucleosome Binding Modes. *Biochemistry* **2022**, *61* (1), 1–9. https://doi.org/10.1021/acs.biochem.1c00603.

(185) Guarnaccia, A. D.; Rose, K. L.; Wang, J.; Zhao, B.; Popay, T. M.; Wang, C. E.; Guerrazzi, K.; Hill, S.; Woodley, C. M.; Hansen, T. J.; Lorey, S. L.; Shaw, J. G.; Payne, W. G.; Weissmiller, A. M.; Olejniczak, E. T.; Fesik, S. W.; Liu, Q.; Tansey, W. P. Impact of WIN Site Inhibitor on the WDR5 Interactome. *Cell Rep.* **2021**, *34* (3), 108636. https://doi.org/10.1016/j.celrep.2020.108636.

(186) Schapira, M.; Tyers, M.; Torrent, M.; Arrowsmith, C. H. WD40 Repeat Domain Proteins: A Novel Target Class? *Nat. Rev. Drug Discov.* **2017**, *16* (11), 773–786. https://doi.org/10.1038/nrd.2017.179.

(187) Getlik, M.; Smil, D.; Zepeda-Velázquez, C.; Bolshan, Y.; Poda, G.; Wu, H.; Dong, A.; Kuznetsova, E.; Marcellus, R.; Senisterra, G.; Dombrovski, L.; Hajian, T.; Kiyota, T.; Schapira, M.; Arrowsmith, C. H.; Brown, P. J.; Vedadi, M.; Al-awar, R. Structure-Based Optimization of a Small Molecule Antagonist of the Interaction Between WD Repeat-Containing Protein 5 (WDR5) and Mixed-Lineage Leukemia 1 (MLL1). *J. Med. Chem.* **2016**, *59* (6), 2478–2496. https://doi.org/10.1021/acs.jmedchem.5b01630.

(188) Bolshan, Y.; Getlik, M.; Kuznetsova, E.; Wasney, G. A.; Hajian, T.; Poda, G.; Nguyen, K. T.; Wu, H.; Dombrovski, L.; Dong, A.; Senisterra, G.; Schapira, M.; Arrowsmith, C. H.; Brown, P. J.; Al-awar, R.; Vedadi, M.; Smil, D. Synthesis, Optimization, and Evaluation of Novel Small Molecules as Antagonists of WDR5-MLL Interaction. *ACS Med. Chem. Lett.* **2013**, *4* (3), 353–357. https://doi.org/10.1021/ml300467n.

(189) Senisterra, G.; Wu, H.; Allali-Hassani, A.; Wasney, G. A.; Barsyte-Lovejoy, D.; Dombrovski, L.; Dong, A.; Nguyen, K. T.; Smil, D.; Bolshan, Y.; Hajian, T.; He, H.; Seitova, A.; Chau, I.; Li, F.; Poda, G.; Couture, J.-F.; Brown, P. J.; Al-Awar, R.; Schapira, M.; Arrowsmith, C. H.; Vedadi, M. Small-Molecule Inhibition of MLL Activity by Disruption of Its Interaction with WDR5. *Biochem. J.* **2013**, *449* (Pt 1), 151–159. https://doi.org/10.1042/BJ20121280.

(190) Gurung, R.; Om, D.; Pun, R.; Hyun, S.; Shin, D. Recent Progress in Modulation of WD40-Repeat Domain 5 Protein (WDR5): Inhibitors and Degraders. *Cancers* **2023**, *15* (15), 3910. https://doi.org/10.3390/cancers15153910.

(191) Mandell, J. G.; Falick, A. M.; Komives, E. A. Identification of Protein–Protein Interfaces by Decreased Amide Proton Solvent Accessibility. *Proc. Natl. Acad. Sci.* **1998**, *95* (25), 14705–14710. https://doi.org/10.1073/pnas.95.25.14705.

(192) Hamuro, Y. Tutorial: Chemistry of Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (1), 133–151. https://doi.org/10.1021/jasms.0c00260.

(193) Mann, M. K.; Wolf, E.; Silva, M.; Kwak, H. A.; Wilson, B.; Bolotokova, A.; Wilson, D. J.; Harding, R. J.; Schapira, M. Small Molecule Screen Identifies Non-Catalytic USP3 Chemical Handle. *ACS Omega* **2024**, *9* (1), 917–924. https://doi.org/10.1021/acsomega.3c07070.

(194) Deng, B.; Zhu, S.; Macklin, A. M.; Xu, J.; Lento, C.; Sljoka, A.; Wilson, D. J. Suppressing Allostery in Epitope Mapping Experiments Using Millisecond Hydrogen / Deuterium Exchange Mass Spectrometry. *mAbs* **2017**, *9* (8), 1327–1336. https://doi.org/10.1080/19420862.2017.1379641.

(195) Dong, A.; Dombrovski, L.; Smil, D.; Getlik, M.; Bolshan, Y.; Walker, John. R.; Senisterra, G.; Poda, G.; Al-Awar, R.; Schapira, M.; Vedadi, M.; Bountra, C.; Edwards, A. M.; Arrowsmith, C.

H.; Brown, P. J.; Wu, H.; Structural Genomics Consortium (SGC). Crystal Structure of Human WDR5 in Complex with Compound 9h, 2015. https://doi.org/10.2210/pdb5eal/pdb.

(196) Schuetz, A.; Allali-Hassani, A.; Martín, F.; Loppnau, P.; Vedadi, M.; Bochkarev, A.; Plotnikov, A. N.; Arrowsmith, C. H.; Min, J. Structural Basis for Molecular Recognition and Presentation of Histone H3 By WDR5. *EMBO J.* **2006**, *25* (18), 4245–4252. https://doi.org/10.1038/sj.emboj.7601316.

(197) Ruthenburg, A. J.; Wang, W.; Graybosch, D. M.; Li, H.; Allis, C. D.; Patel, D. J.; Verdine, G. L. Histone H3 Recognition and Presentation by the WDR5 Module of the MLL1 Complex. *Nat. Struct. Mol. Biol.* **2006**, *13* (8), 704–712. https://doi.org/10.1038/nsmb1119.

(198) Song, J.-J.; Kingston, R. E. WDR5 Interacts with Mixed Lineage Leukemia (MLL) Protein via the Histone H3-Binding Pocket. *J. Biol. Chem.* **2008**, *283* (50), 35258–35264. https://doi.org/10.1074/jbc.M806900200.

(199) Han, Z.; Guo, L.; Wang, H.; Shen, Y.; Deng, X. W.; Chai, J. Structural Basis for the Specific Recognition of Methylated Histone H3 Lysine 4 by the WD-40 Protein WDR5. *Mol. Cell* **2006**, *22* (1), 137–144. https://doi.org/10.1016/j.molcel.2006.03.018.

(200) Couture, J.-F.; Collazo, E.; Trievel, R. C. Molecular Recognition of Histone H3 by the WD40 Protein WDR5. *Nat. Struct. Mol. Biol.* **2006**, *13* (8), 698–703. https://doi.org/10.1038/nsmb1116.

(201) Dou, Y.; Milne, T. A.; Tackett, A. J.; Smith, E. R.; Fukuda, A.; Wysocka, J.; Allis, C. D.; Chait, B. T.; Hess, J. L.; Roeder, R. G. Physical Association and Coordinate Function of the H3 K4 Methyltransferase MLL1 and the H4 K16 Acetyltransferase MOF. *Cell* **2005**, *121* (6), 873–885. https://doi.org/10.1016/j.cell.2005.04.031.

(202) Scrosati, P. M.; Yin, V.; Konermann, L. Hydrogen/Deuterium Exchange Measurements May Provide an Incomplete View of Protein Dynamics: A Case Study on Cytochrome c. *Anal. Chem.* **2021**, *93* (42), 14121–14129. https://doi.org/10.1021/acs.analchem.1c02471.

(203) Marrella, S. A.; Brown, K. A.; Mansouri-Noori, F.; Porat, J.; Wilson, D. J.; Bayfield, M. A. An Interdomain Bridge Influences RNA Binding of the Human La Protein. *J. Biol. Chem.* **2019**, *294* (5), 1529–1540. https://doi.org/10.1074/jbc.RA118.003995.

(204) Karatas, H.; Townsend, E. C.; Cao, F.; Chen, Y.; Bernard, D.; Liu, L.; Lei, M.; Dou, Y.; Wang, S. High-Affinity, Small-Molecule Peptidomimetic Inhibitors of MLL1/WDR5 Protein–Protein Interaction. *J. Am. Chem. Soc.* **2013**, *135* (2), 669–682. https://doi.org/10.1021/ja306028q.

(205) Patel, A.; Dharmarajan, V.; Cosgrove, M. S. Structure of WDR5 Bound to Mixed Lineage Leukemia Protein-1 Peptide*. *J. Biol. Chem.* **2008**, *283* (47), 32158–32161. https://doi.org/10.1074/jbc.C800164200.

(206) Avdic, V.; Zhang, P.; Lanouette, S.; Groulx, A.; Tremblay, V.; Brunzelle, J.; Couture, J.-F. Structural and Biochemical Insights into MLL1 Core Complex Assembly. *Structure* **2011**, *19* (1), 101–108. https://doi.org/10.1016/j.str.2010.09.022.

(207) Markowska, A.; Markowski, A. R.; Jarocka-Karpowicz, I. The Importance of 6-Aminohexanoic Acid as a Hydrophobic, Flexible Structural Element. *Int. J. Mol. Sci.* **2021**, *22* (22), 12122. https://doi.org/10.3390/ijms222212122.

(208) Rahman, S.; Hoffmann, N. A.; Worden, E. J.; Smith, M. L.; Namitz, K. E. W.; Knutson, B. A.; Cosgrove, M. S.; Wolberger, C. Multistate Structures of the MLL1-WRAD Complex Bound to H2B-Ubiquitinated Nucleosome. *Proc. Natl. Acad. Sci.* **2022**, *119* (38), e2205691119. https://doi.org/10.1073/pnas.2205691119.

(209) Resetca, D.; Haftchenary, S.; Gunning, P. T.; Wilson, D. J. Changes in Signal Transducer and Activator of Transcription 3 (STAT3) Dynamics Induced by Complexation with Pharmacological Inhibitors of Src Homology 2 (SH2) Domain Dimerization. *J. Biol. Chem.* **2014**, *289* (47), 32538–32547. https://doi.org/10.1074/jbc.M114.595454.

(210) Mann, M. K.; Wolf, E.; Silva, M.; Kwak, H. A.; Wilson, B.; Bolotokova, A.; Wilson, D. J.; Harding, R. J.; Schapira, M. Small Molecule Screen Identifies Non-Catalytic USP3 Chemical Handle. *ACS Omega* **2023**, *9* (1), 917–924. https://doi.org/10.1021/acsomega.3c07070.

(211) Narang, D.; James, D. A.; Balmer, M. T.; Wilson, D. J. Protein Footprinting, Conformational Dynamics, and Core Interface-Adjacent Neutralization “Hotspots” in the SARS-CoV-2 Spike Protein Receptor Binding Domain/Human ACE2 Interaction. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (7), 1593–1600. https://doi.org/10.1021/jasms.0c00465.

(212) Kuo, A. J.; Cheung, P.; Chen, K.; Zee, B. M.; Kioi, M.; Lauring, J.; Xi, Y.; Park, B. H.; Shi, X.; Garcia, B. A.; Li, W.; Gozani, O. NSD2 Links Dimethylation of Histone H3 at Lysine 36 to Oncogenic Programming. *Mol. Cell* **2011**, *44* (4), 609–620. https://doi.org/10.1016/j.molcel.2011.08.042.

(213) Keats, J. J.; Reiman, T.; Maxwell, C. A.; Taylor, B. J.; Larratt, L. M.; Mant, M. J.; Belch, A. R.; Pilarski, L. M. In Multiple Myeloma, t(4;14)(P16;Q32) Is an Adverse Prognostic Factor Irrespective of FGFR3 Expression. *Blood* **2003**, *101* (4), 1520–1529. https://doi.org/10.1182/blood-2002-06-1675.

(214) Qin, S.; Min, J. Structure and Function of the Nucleosome-Binding PWWP Domain. *Trends Biochem. Sci.* **2014**, *39* (11), 536–547. https://doi.org/10.1016/j.tibs.2014.09.001.

(215) Sankaran, S. M.; Wilkinson, A. W.; Elias, J. E.; Gozani, O. A PWWP Domain of Histone-Lysine *N*-Methyltransferase NSD2 Binds to Dimethylated Lys-36 of Histone H3 and Regulates NSD2 Function at Chromatin*. *J. Biol. Chem.* **2016**, *291* (16), 8465–8474. https://doi.org/10.1074/jbc.M116.720748.

(216) Ferreira de Freitas, R.; Liu, Y.; Szewczyk, M. M.; Mehta, N.; Li, F.; McLeod, D.; Zepeda-Velázquez, C.; Dilworth, D.; Hanley, R. P.; Gibson, E.; Brown, P. J.; Al-Awar, R.; James, L. I.; Arrowsmith, C. H.; Barsyte-Lovejoy, D.; Min, J.; Vedadi, M.; Schapira, M.; Allali-Hassani, A. Discovery of Small-Molecule Antagonists of the PWWP Domain of NSD2. *J. Med. Chem.* **2021**, *64* (3), 1584–1592. https://doi.org/10.1021/acs.jmedchem.0c01768.

(217) Dilworth, D.; Hanley, R. P.; Ferreira de Freitas, R.; Allali-Hassani, A.; Zhou, M.; Mehta, N.; Marunde, M. R.; Ackloo, S.; Carvalho Machado, R. A.; Khalili Yazdi, A.; Owens, D. D. G.; Vu, V.; Nie, D. Y.; Alqazzaz, M.; Marcon, E.; Li, F.; Chau, I.; Bolotokova, A.; Qin, S.; Lei, M.; Liu, Y.; Szewczyk, M. M.; Dong, A.; Kazemzadeh, S.; Abramyan, T.; Popova, I. K.; Hall, N. W.; Meiners,

M. J.; Cheek, M. A.; Gibson, E.; Kireev, D.; Greenblatt, J. F.; Keogh, M.-C.; Min, J.; Brown, P. J.; Vedadi, M.; Arrowsmith, C. H.; Barsyte-Lovejoy, D.; James, L. I.; Schapira, M. A Chemical Probe Targeting the PWWP Domain Alters NSD2 Nucleolar Localization. *Nat. Chem. Biol.* **2022**, *18* (1), 56–63. https://doi.org/10.1038/s41589-021-00898-0.

(218) Hanley, R. P.; Nie, D. Y.; Tabor, J. R.; Li, F.; Sobh, A.; Xu, C.; Barker, N. K.; Dilworth, D.; Hajian, T.; Gibson, E.; Szewczyk, M. M.; Brown, P. J.; Barsyte-Lovejoy, D.; Herring, L. E.; Wang, G. G.; Licht, J. D.; Vedadi, M.; Arrowsmith, C. H.; James, L. I. Discovery of a Potent and Selective Targeted NSD2 Degrader for the Reduction of H3K36me2. *J. Am. Chem. Soc.* **2023**, *145* (14), 8176–8188. https://doi.org/10.1021/jacs.3c01421.

(219) Matta-Camacho, E.; Kozlov, G.; Li, F. F.; Gehring, K. Structural Basis of Substrate Recognition and Specificity in the N-End Rule Pathway. *Nat. Struct. Mol. Biol.* **2010**, *17* (10), 1182–1187. https://doi.org/10.1038/nsmb.1894.

(220) Nie, D. Y.; Tabor, J. R.; Li, J.; Kutera, M.; St-Germain, J.; Hanley, R. P.; Wolf, E.; Paulakonis, E.; Kenney, T. M. G.; Duan, S.; Shrestha, S.; Owens, D. D. G.; Pon, A.; Szewczyk, M.; Lamberto, A. J.; Menes, M.; Li, F.; Barsyte-Lovejoy, D.; Brown, N. G.; Barsotti, A. M.; Stamford, A. W.; Collins, J. L.; Wilson, D. J.; Raught, B.; Licht, J. D.; James, L. I.; Arrowsmith, C. H. Recruitment of FBXO22 for Targeted Degradation of NSD2. *BioRxiv Prepr. Serv. Biol.* **2023**, 2023.11.01.564830. https://doi.org/10.1101/2023.11.01.564830.

(221) Borziak, K.; Zhulin, I. B. FIST: A Sensory Domain for Diverse Signal Transduction Pathways in Prokaryotes and Ubiquitin Signaling in Eukaryotes. *Bioinformatics* **2007**, *23* (19), 2518–2521. https://doi.org/10.1093/bioinformatics/btm384.

(222) Barnett, R. E.; Jencks, W. P. Diffusion-Controlled and Concerted Base Catalysis in the Decomposition of Hemithioacetals. *J. Am. Chem. Soc.* **1969**, *91* (24), 6758–6765. https://doi.org/10.1021/ja01052a038.

(223) Cecchini, C.; Pannilunghi, S.; Tardy, S.; Scapozza, L. From Conception to Development: Investigating PROTACs Features for Improved Cell Permeability and Successful Protein Degradation. *Front. Chem.* **2021**, *9*. https://doi.org/10.3389/fchem.2021.672267.

(224) Eron, S. J.; Huang, H.; Agafonov, R. V.; Fitzgerald, M. E.; Patel, J.; Michael, R. E.; Lee, T. D.; Hart, A. A.; Shaulsky, J.; Nasveschuk, C. G.; Phillips, A. J.; Fisher, S. L.; Good, A. Structural Characterization of Degrader-Induced Ternary Complexes Using Hydrogen–Deuterium Exchange Mass Spectrometry and Computational Modeling: Implications for Structure-Based Design. *ACS Chem. Biol.* **2021**, *16* (11), 2228–2243. https://doi.org/10.1021/acschembio.1c00376.

(225) Mann, M. K.; Wolf, E.; Silva, M.; Kwak, H. A.; Wilson, B.; Bolotokova, A.; Wilson, D. J.; Harding, R. J.; Schapira, M. Small Molecule Screen Identifies Non-Catalytic USP3 Chemical Handle. *ACS Omega* **2024**, *9* (1), 917–924. https://doi.org/10.1021/acsomega.3c07070.

(226) Oganesyan, I.; Lento, C.; Wilson, D. J. Contemporary Hydrogen Deuterium Exchange Mass Spectrometry. *Methods* **2018**, *144*, 27–42. https://doi.org/10.1016/j.ymeth.2018.04.023.

(227) James, E. I.; Murphree, T. A.; Vorauer, C.; Engen, J. R.; Guttman, M. Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chem. Rev.* **2022**, *122* (8), 7562–7623. https://doi.org/10.1021/acs.chemrev.1c00279.

(228) Mann, M. K.; Zepeda-Velázquez, C. A.; González-Álvarez, H.; Dong, A.; Kiyota, T.; Aman, A. M.; Loppnau, P.; Li, Y.; Wilson, B.; Arrowsmith, C. H.; Al-Awar, R.; Harding, R. J.; Schapira, M. Structure–Activity Relationship of USP5 Inhibitors. *J. Med. Chem.* **2021**, *64* (20), 15017–15036. https://doi.org/10.1021/acs.jmedchem.1c00889.

(229) Reyes-Turcu, F. E.; Horton, J. R.; Mullally, J. E.; Heroux, A.; Cheng, X.; Wilkinson, K. D. The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell* **2006**, *124* (6), 1197–1208. https://doi.org/10.1016/j.cell.2006.02.038.

(230) 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.

(231) Jørgensen, T. J. D.; Gårdsvoll, H.; Danø, K.; Roepstorff, P.; Ploug, M. Dynamics of Urokinase Receptor Interaction with Peptide Antagonists Studied by Amide Hydrogen Exchange and Mass Spectrometry. *Biochemistry* **2004**, *43* (47), 15044–15057. https://doi.org/10.1021/bi048706j.

(232) Hermet, P.; Delache, B.; Herate, C.; Wolf, E.; Kivi, G.; Juronen, E.; Mumm, K.; Žusinaite, E.; Kainov, D.; Sankovski, E.; Virumäe, K.; Planken, A.; Merits, A.; Besaw, J. E.; Yee, A. W.; Morizumi, T.; Kim, K.; Kuo, A.; Berriche, A.; Dereuddre-Bosquet, N.; Sconosciuti, Q.; Naninck, T.; Relouzat, F.; Cavarelli, M.; Ustav, M.; Wilson, D.; Ernst, O. P.; Männik, A.; LeGrand, R.; Jr, M. U. Broadly Neutralizing Humanized SARS-CoV-2 Antibody Binds to a Conserved Epitope on Spike and Provides Antiviral Protection through Inhalation-Based Delivery in Non-Human Primates. *PLOS Pathog.* **2023**, *19* (8), e1011532. https://doi.org/10.1371/journal.ppat.1011532.

(233) Justo Arevalo, S.; Castillo-Chávez, A.; Uribe Calampa, C. S.; Zapata Sifuentes, D.; Huallpa, C. J.; Landa Bianchi, G.; Garavito-Salini Casas, R.; Quiñones Aguilar, M.; Pineda Chavarría, R. What Do We Know about the Function of SARS-CoV-2 Proteins? *Front. Immunol.* **2023**, *14*, 1249607. https://doi.org/10.3389/fimmu.2023.1249607.

(234) Pizzato, M.; Baraldi, C.; Boscato Sopetto, G.; Finozzi, D.; Gentile, C.; Gentile, M. D.; Marconi, R.; Paladino, D.; Raoss, A.; Riedmiller, I.; Ur Rehman, H.; Santini, A.; Succetti, V.; Volpini, L. SARS-CoV-2 and the Host Cell: A Tale of Interactions. *Front. Virol.* **2022**, *1*. https://doi.org/10.3389/fviro.2021.815388.

(235) Jackson, C. B.; Farzan, M.; Chen, B.; Choe, H. Mechanisms of SARS-CoV-2 Entry into Cells. *Nat. Rev. Mol. Cell Biol.* **2022**, *23* (1), 3–20. https://doi.org/10.1038/s41580-021-00418-x.

(236) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; Wang, X. Structure of the SARS-CoV-2 Spike Receptor-Binding Domain Bound to the ACE2 Receptor. *Nature* **2020**, *581* (7807), 215–220. https://doi.org/10.1038/s41586-020-2180-5.

(237) Baerga-Ortiz, A.; Hughes, C. A.; Mandell, J. G.; Komives, E. A. Epitope Mapping of a Monoclonal Antibody against Human Thrombin by H/D-Exchange Mass Spectrometry Reveals Selection of a Diverse Sequence in a Highly Conserved Protein. *Protein Sci.* **2002**, *11* (6), 1300–1308. https://doi.org/10.1110/ps.4670102.

(238) Yamada, N.; Suzuki, E.; Hirayama, K. Identification of the Interface of a Large Protein–Protein Complex Using H/D Exchange and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2002**, *16* (4), 293–299. https://doi.org/10.1002/rcm.579.

(239) Jethva, P. N.; Gross, M. L. Hydrogen Deuterium Exchange and Other Mass Spectrometry- Based Approaches for Epitope Mapping. *Front. Anal. Sci.* **2023**, *3*. https://doi.org/10.3389/frans.2023.1118749.

(240) Zhu, S.; Liuni, P.; Chen, T.; Houy, C.; Wilson, D. J.; James, D. A. Epitope Screening Using Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS): An Accelerated Workflow for Evaluation of Lead Monoclonal Antibodies. *Biotechnol. J.* **2022**, *17* (2), 2100358. https://doi.org/10.1002/biot.202100358.

(241) Calvaresi, V.; Wrobel, A. G.; Toporowska, J.; Hammerschmid, D.; Doores, K. J.; Bradshaw, R. T.; Parsons, R. B.; Benton, D. J.; Roustan, C.; Reading, E.; Malim, M. H.; Gamblin, S. J.; Politis, A. Structural Dynamics in the Evolution of SARS-CoV-2 Spike Glycoprotein. *Nat. Commun.* **2023**, *14* (1), 1421. https://doi.org/10.1038/s41467-023-36745-0.

(242) Seow, J.; Khan, H.; Rosa, A.; Calvaresi, V.; Graham, C.; Pickering, S.; Pye, V. E.; Cronin, N. B.; Huettner, I.; Malim, M. H.; Politis, A.; Cherepanov, P.; Doores, K. J. A Neutralizing Epitope on the SD1 Domain of SARS-CoV-2 Spike Targeted Following Infection and Vaccination. *Cell Rep.* **2022**, *40* (8), 111276. https://doi.org/10.1016/j.celrep.2022.111276.

(243) Costello, S. M.; Shoemaker, S. R.; Hobbs, H. T.; Nguyen, A. W.; Hsieh, C.-L.; Maynard, J. A.; McLellan, J. S.; Pak, J. E.; Marqusee, S. The SARS-CoV-2 Spike Reversibly Samples an Open-Trimer Conformation Exposing Novel Epitopes. *Nat. Struct. Mol. Biol.* **2022**, *29* (3), 229–238. https://doi.org/10.1038/s41594-022-00735-5.

(244) Chen, C.; Zhu, R.; Hodge, E. A.; Díaz-Salinas, M. A.; Nguyen, A.; Munro, J. B.; Lee, K. K. hACE2-Induced Allosteric Activation in SARS-CoV versus SARS-CoV-2 Spike Assemblies Revealed by Structural Dynamics. *ACS Infect. Dis.* **2023**, *9* (6), 1180–1189. https://doi.org/10.1021/acsinfecdis.3c00010.

(245) Raghuvamsi, P. V.; Tulsian, N. K.; Samsudin, F.; Qian, X.; Purushotorman, K.; Yue, G.; Kozma, M. M.; Hwa, W. Y.; Lescar, J.; Bond, P. J.; MacAry, P. A.; Anand, G. S. SARS-CoV-2 S Protein:ACE2 Interaction Reveals Novel Allosteric Targets. *eLife* **2021**, *10*, e63646. https://doi.org/10.7554/eLife.63646.

(246) Silva, R. P.; Huang, Y.; Nguyen, A. W.; Hsieh, C.-L.; Olaluwoye, O. S.; Kaoud, T. S.; Wilen, R. E.; Qerqez, A. N.; Park, J.-G.; Khalil, A. M.; Azouz, L. R.; Le, K. C.; Bohanon, A. L.; DiVenere, A. M.; Liu, Y.; Lee, A. G.; Amengor, D. A.; Shoemaker, S. R.; Costello, S. M.; Padlan, E. A.; Marqusee, S.; Martinez-Sobrido, L.; Dalby, K. N.; D'Arcy, S.; McLellan, J. S.; Maynard, J. A. Identification of a Conserved S2 Epitope Present on Spike Proteins from All Highly Pathogenic Coronaviruses. *eLife* **2023**, *12*, e83710. https://doi.org/10.7554/eLife.83710.

(247) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058.

(248) Barnes, C. O.; Jette, C. A.; Abernathy, M. E.; Dam, K.-M. A.; Esswein, S. R.; Gristick, H. B.; Malyutin, A. G.; Sharaf, N. G.; Huey-Tubman, K. E.; Lee, Y. E.; Robbiani, D. F.; Nussenzweig, M. C.; West, A. P.; Bjorkman, P. J. SARS-CoV-2 Neutralizing Antibody Structures Inform Therapeutic Strategies. *Nature* **2020**, *588* (7839), 682–687. https://doi.org/10.1038/s41586-020-2852-1.

(249) Barnes, C. O.; Jette, C. A.; Abernathy, M. E.; Dam, K.-M. A.; Esswein, S. R.; Gristick, H. B.; Malyutin, A. G.; Sharaf, N. G.; Huey-Tubman, K. E.; Lee, Y. E.; Robbiani, D. F.; Nussenzweig, M. C.; West, A. P.; Bjorkman, P. J. SARS-CoV-2 Neutralizing Antibody Structures Inform Therapeutic Strategies. *Nature* **2020**, *588* (7839), 682–687. https://doi.org/10.1038/s41586-020-2852-1.

(250) Sowole, M. A.; Konermann, L. Effects of Protein–Ligand Interactions on Hydrogen/Deuterium Exchange Kinetics: Canonical and Noncanonical Scenarios. *Anal. Chem.* **2014**, *86* (13), 6715–6722. https://doi.org/10.1021/ac501849n.

(251) Wang, Q.; Iketani, S.; Li, Z.; Liu, L.; Guo, Y.; Huang, Y.; Bowen, A. D.; Liu, M.; Wang, M.; Yu, J.; Valdez, R.; Lauring, A. S.; Sheng, Z.; Wang, H. H.; Gordon, A.; Liu, L.; Ho, D. D. Alarming Antibody Evasion Properties of Rising SARS-CoV-2 BQ and XBB Subvariants. *Cell* **2023**, *186* (2), 279-286.e8. https://doi.org/10.1016/j.cell.2022.12.018.

(252) Hoffmann, M.; Krüger, N.; Schulz, S.; Cossmann, A.; Rocha, C.; Kempf, A.; Nehlmeier, I.; Graichen, L.; Moldenhauer, A.-S.; Winkler, M. S.; Lier, M.; Dopfer-Jablonka, A.; Jäck, H.-M.; Behrens, G. M. N.; Pöhlmann, S. The Omicron Variant Is Highly Resistant against Antibody-Mediated Neutralization: Implications for Control of the COVID-19 Pandemic. *Cell* **2022**, *185* (3), 447-456.e11. https://doi.org/10.1016/j.cell.2021.12.032.

(253) Imai Masaki; Ito Mutsumi; Kiso Maki; Yamayoshi Seiya; Uraki Ryuta; Fukushi Shuetsu; Watanabe Shinji; Suzuki Tadaki; Maeda Ken; Sakai-Tagawa Yuko; Iwatsuki-Horimoto Kiyoko; Halfmann Peter J.; Kawaoka Yoshihiro. Efficacy of Antiviral Agents against Omicron Subvariants BQ.1.1 and XBB. *N. Engl. J. Med.* **2023**, *388* (1), 89–91. https://doi.org/10.1056/NEJMc2214302.

(254) Liu, X. R.; Zhang, M. M.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications. *Chem. Rev.* **2020**, *120* (10), 4355–4454. https://doi.org/10.1021/acs.chemrev.9b00815.

(255) Keener, J. E.; Zhang, G.; Marty, M. T. Native Mass Spectrometry of Membrane Proteins. *Anal. Chem.* **2021**, *93* (1), 583–597. https://doi.org/10.1021/acs.analchem.0c04342.

(256) Hammerschmid, D.; Calvaresi, V.; Bailey, C.; Russell Lewis, B.; Politis, A.; Morris, M.; Denbigh, L.; Anderson, M.; Reading, E. Chromatographic Phospholipid Trapping for Automated H/D Exchange Mass Spectrometry of Membrane Protein–Lipid Assemblies. *Anal. Chem.* **2023**, *95* (5), 3002–3011. https://doi.org/10.1021/acs.analchem.2c04876.

(257) Anderson, K. W.; Gallagher, E. S.; Hudgens, J. W. Automated Removal of Phospholipids from Membrane Proteins for H/D Exchange Mass Spectrometry Workflows. *Anal. Chem.* **2018**, *90* (11), 6409–6412. https://doi.org/10.1021/acs.analchem.8b00429.

(258) Calvaresi, V.; Redsted, A.; Norais, N.; Rand, K. D. Hydrogen–Deuterium Exchange Mass Spectrometry with Integrated Size-Exclusion Chromatography for Analysis of Complex Protein Samples. *Anal. Chem.* **2021**, *93* (33), 11406–11414. https://doi.org/10.1021/acs.analchem.1c01171.

(259) Peterle, D.; DePice, D.; Wales, T. E.; Engen, J. R. Increase the Flow Rate and Improve Hydrogen Deuterium Exchange Mass Spectrometry. *J. Chromatogr. A* **2023**, *1689*, 463742. https://doi.org/10.1016/j.chroma.2022.463742.

(260) Ludwig, C.; Wagner, R. Virus-like Particles—Universal Molecular Toolboxes. *Curr. Opin. Biotechnol.* **2007**, *18* (6), 537–545. https://doi.org/10.1016/j.copbio.2007.10.013.

(261) Martins, S. A.; Santos, J.; Silva, R. D. M.; Rosa, C.; Cabo Verde, S.; Correia, J. D. G.; Melo, R. How Promising Are HIV-1-Based Virus-like Particles for Medical Applications. *Front. Cell. Infect. Microbiol.* **2022**, *12*, 997875. https://doi.org/10.3389/fcimb.2022.997875.

(262) Guo, Q.; Bin Huang; Cheng, J.; Seefelder, M.; Engler, T.; Pfeifer, G.; Oeckl, P.; Otto, M.; Moser, F.; Maurer, M.; Pautsch, A.; Baumeister, W.; Fernández-Busnadiego, R.; Kochanek, S. The Cryo-Electron Microscopy Structure of Huntingtin. *Nature* **2018**, *555* (7694), 117–120. https://doi.org/10.1038/nature25502.

(263) Donaldson, J.; Powell, S.; Rickards, N.; Holmans, P.; Jones, L. What Is the Pathogenic CAG Expansion Length in Huntington's Disease? *J. Huntingt. Dis.* **2021**, *10* (1), 175–202. https://doi.org/10.3233/JHD-200445.

(264) Ferguson, M. W.; Kennedy, C. J.; Palpagama, T. H.; Waldvogel, H. J.; Faull, R. L. M.; Kwakowsky, A. Current and Possible Future Therapeutic Options for Huntington's Disease. *J. Cent. Nerv. Syst. Dis.* **2022**, *14*, 11795735221092517. https://doi.org/10.1177/11795735221092517.

(265) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, *50* (D1), D439–D444. https://doi.org/10.1093/nar/gkab1061.

(266) 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.

Appendices

Appendix A: Supplementary information for Chapter 1.

Table A1. *L* and *R* Factors for k_{NH} Determination reproduced from Bai et al. (1993) with permission.[68]

	Logkex(X) − Logkex(Ala)			
	Acid catalysis		Base catalysis	
	L	R	L	R
Ala	0	0	0	0
Arg	−0.59	−0.32	0.08	0.22
Asn	−0.58	−0.13	0.49	0.32
Asp(COO−)	-0.9	0.58	−0.30	−0.18
Asp(COOH)	(−0.9)	−0.12	0.69	-0.6
Cys	−0.54	−0.46	0.62	0.55
Cys2	−0.74	−0.58	0.55	0.46
Gly	−0.22	0.22	0.27	0.17
Gln	−0.47	−0.27	0.06	0.2
Glu(COO−)	(−0.9)	0.31	−0.51	−0.15
Glu(COOH)	(−0.6)	−0.27	0.24	0.39
His			−0.10	0.14
His+	(−0.8)	−0.51	-0.8	0.83
Ile	−0.91	−0.59	−0.73	−0.23
Leu	−0.57	−0.13	−0.58	−0.21
Lys	−0.56	−0.29	−0.04	0.12
Met	−0.64	−0.28	−0.01	0.11
Phe	−0.52	−0.43	−0.24	0.06
Pro*(trans)*		−0.19		−0.24
Pro(cis)		−0.85		0.6
Ser	−0.44	−0.39	0.37	0.3
Thr	−0.79	−0.47	−0.07	0.2
Trp	−0.40	−0.44	−0.41	−0.11
Tyr	−0.41	−0.37	−0.27	0.05
Val	−0.74	−0.30	−0.70	−0.14
N-term (NH3+)		−1.32		1.62
C-term (COO−)	0.96		(−1.8)	
C-term (COOH)	-0.05			

Appendix B: Supplementary information for Chapter 2.

Table B1. Sequences of Bcl-2 and Mcl-1.

[a]The starting residue (1) corresponds to the start of WT Bcl2 sequence and continues in Bcl-2 chimera until the loop replacement (34), [b]The sequence corresponding to WT Bcl-X_L is denoted in red (35-50), [c]Note that the sequence corresponding to the loop deletion is omitted which is why the sequence here skips from 51 to 90. [d]GSGS is an artifact after GST tag removal.

Table B2. % Sequence Coverage and Redundancy

State	% Coverage	Redundancy
Bcl2 Only	75	2.28
Bcl2/Bid	69	2.22
Bcl2/Bim	61	1.67
Bcl2/Bad	62	1.76
Bcl2/Noxa	55	1.93
Mcl1 Only	92	2.11
Mcl1/Bid	85	1.54
Mcl1/Bim	84	2.05
Mcl1/Bad	77	1.71
Mcl1/Noxa	92	2.11

Table B3. Mcl1 Peptide List

Mcl-1						
m/z	z	m	Drift Time	Sequence	Start	Stop
676.8288	2	1351.642	3.255	SDELYRQSLEI	^172	181
589.8311	2	1178.655	2.604	EIISRYLRE	180	188
904.4791	1	903.4713	7.161	KPMGRSGAT	197	205
821.4367	3	2461.287	3.038	QATGAKDTKPMGRSGATSRKALET	189	212
644.5959	4	2574.352	2.713	QATGAKDTKPMGRSGATSRKALETL	189	213
581.2994	3	1741.884	2.007	RRVGDGVQRNHETAF	214	228
686.676	3	2057.004	2.333	RRVGDGVQRNHETAFQGM	214	231
447.2531	4	1785.992	2.224	LRKLDIKNEDDVKSL	232	246
710.071	3	2128.201	3.201	LRKLDIKNEDDVKSLSRV	232	249
672.377	3	2015.12	2.93	RKLDIKNEDDVKSLSRV	233	249
703.3295	2	1405.652	3.309	MIHVFSDGVTNW	250	261
637.8085	2	1274.609	2.875	IHVFSDGVTNW	251	261
682.3511	3	2045.036	2.604	MIHVFSDGVTNWGRIVTL	250	267
638.6748	3	1914.01	2.333	IHVFSDGVTNWGRIVTL	251	267
709.9061	2	1417.797	3.689	ISFGAFVAKHLKT	268	280
735.89	2	1470.773	3.58	VAKHLKTINQESC	274	286
861.428	1	861.428	6.835	SCIEPLAE	285	292
667.3611	3	2000.071	2.441	VRTKRDWLVKQRGWDG	299	314
537.5432	4	2147.149	2.17	VRTKRDWLVKQRGWDGF	299	315
1002.437	1	1002.439	8.354	FHVEDLEGG	319	327

Table B4. Bcl2 Peptide List

Bcl-2						
m/z	z	m	Drift Time	Sequence	Start	Stop
449.5257	3	1345.554	2.062	AHAGRTGYDNRE	2	13
507.7625	4	2027.019	2.007	IVMKYIHYKLSQRGYE	14	29
812.7342	3	2435.179	2.984	VEENRTEAPEGTESEPVVHLTL	**91	97
907.7771	3	2720.308	3.309	NRTEAPEGTESEPVVHLTLRQAGDD	*91	103
389.7483	2	777.4809	1.682	PVVHLTL	91	97
511.7341	2	1021.453	2.170	TLRQAGDDF	96	104
434.9001	3	1301.677	1.573	FSRRYRRDF	104	112
573.3018	3	1716.882	2.496	SSQLHLTPFTARGRF	116	130
434.9091	3	1301.704	1.953	HLTPFTARGRF	120	130
583.3159	2	1164.616	2.984	LTPFTARGRF	121	130
503.7474	2	1005.479	2.224	LFRDGVNW	137	144
825.4393	2	1648.863	4.325	LFRDGVNWGRIVAF	137	150
447.2057	2	892.3957	1.736	FRDGVNW	138	144
695.3737	2	1388.732	3.201	FRDGVNWGRIVA	138	149
512.9453	3	1535.812	2.116	FRDGVNWGRIVAF	138	150
789.3725	2	1576.729	3.580	CVESVNREMSPLVD	158	171
829.4379	2	1656.86	3.852	SVNREMSPLVDNIAL	161	175
534.2533	3	1599.736	2.007	WMTEYLNRHLHT	176	187
800.887	2	1599.758	4.120	MTEYLNRHLHTW	177	188
654.7814	2	1307.547	2.930	WIQDNGGWDAF	188	198
469.2087	4	1872.803	2.007	VELYGPSMRHHHHHH	199	207-6His

Table B5. Locations of Hydrophobic and Conserved Residues in BH3 Peptides[a]

Position									H1 (i)				H2 (i+4)			H3 (i+7)		i+9		H4							
Bim		D	M	R	P	E	I	W	I	A	Q	E	L	R	R	I	G	**D**	E	F	N	A	Y	Y	A	R	R
Bid	S	Q	E	D	I	I	R	N	I	A	R	H	L	A	Q	V	G	**D**	S	M	D	R	S	I	P	P	G
Bad		N	L	W	A	A	Q	R	Y	G	R	E	L	R	R	M	S	**D**	E	F	V	D	S	F	K	K	
Noxa			A	E	L	E	V	E	C	A	T	Q	L	R	R	F	G	**D**	K	L	N	F	R	Q	K	L	L

[a]Hydrophobic residues (H1-H4) of BH3 peptides embed in hydrophobic pockets (P1-P4) of the BH3 binding groove in Bcl-2 and Mcl-1, D (i+9) is bolded to indicate conservation across these BH3 peptides.

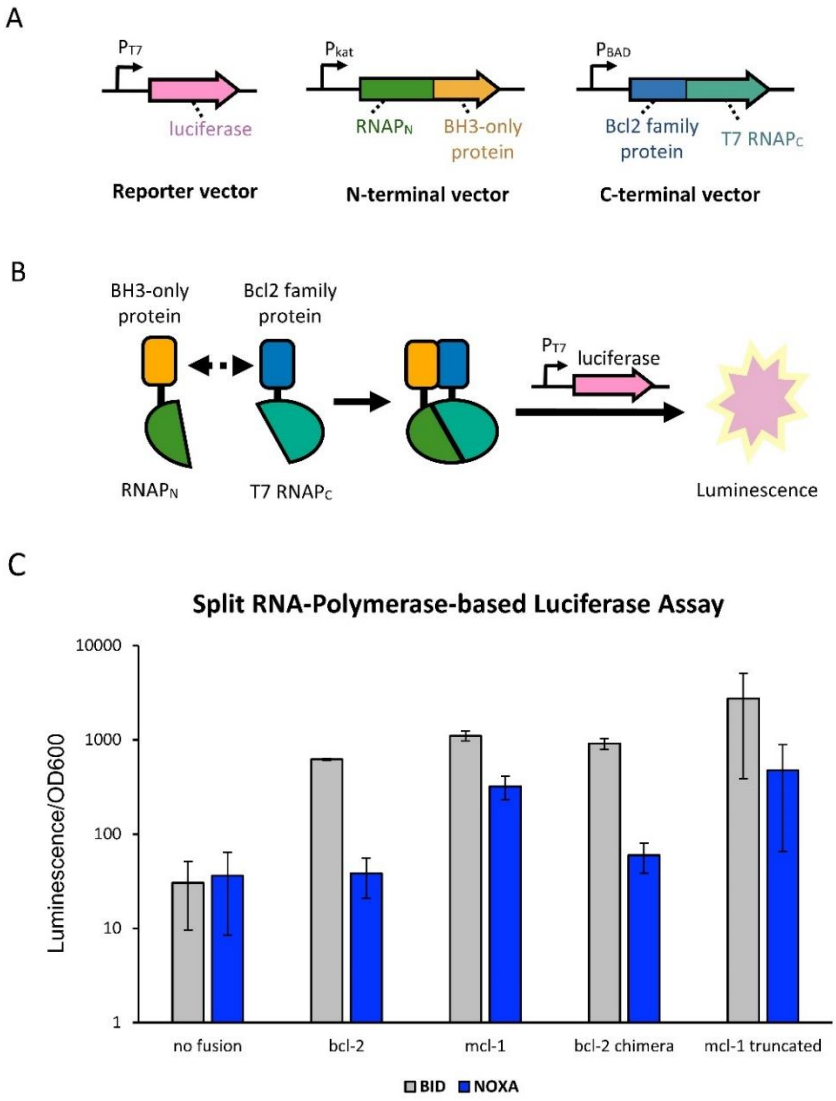


Figure B1. *(Work completed by Jin Pu)* Split RNA polymerase-based luciferase assay comparing WT (ΔTM) Bcl-2 and Mcl-1 to Bcl-2 Chimera and Truncated Mcl-1 in *E. coli*. (A) E. coli vector system to detect the interactions between Bcl-2 family proteins (e.g., Bcl-2, Mcl-1, Bcl-2 Chimera, Truncated Mcl-1) and BH3-only proteins (e.g., BID, NOXA). (B) Schematic of split T7 RNA polymerase-based biosensors to monitor Bcl-2 family PPIs using the vectors shown in (A). A Bcl2-family protein is fused to C-terminal split T7 RNA polymerase (RNAPC), and a BH3-only protein is fused to an evolved proximity-dependent N-terminal T7 RNA polymerase variant (RNAPN). Interaction between the fused proteins lead to the assembly of the RNAP and transcriptional signal of the luminescence reporter in *E. coli*. (C) Luminescence signal output of split RNAPs with different Bcl2-family proteins interaction with either NOXA or BID assayed in E. coli using the system described in (B). *E. coli* cells were induced for 3h with arabinose and then analyzed for luminescence. Error bars are ± sem, n = 3. Detailed methods for this experiment are provided in *J. Am. Chem. Soc*. 2017, 139 (34), 11964.

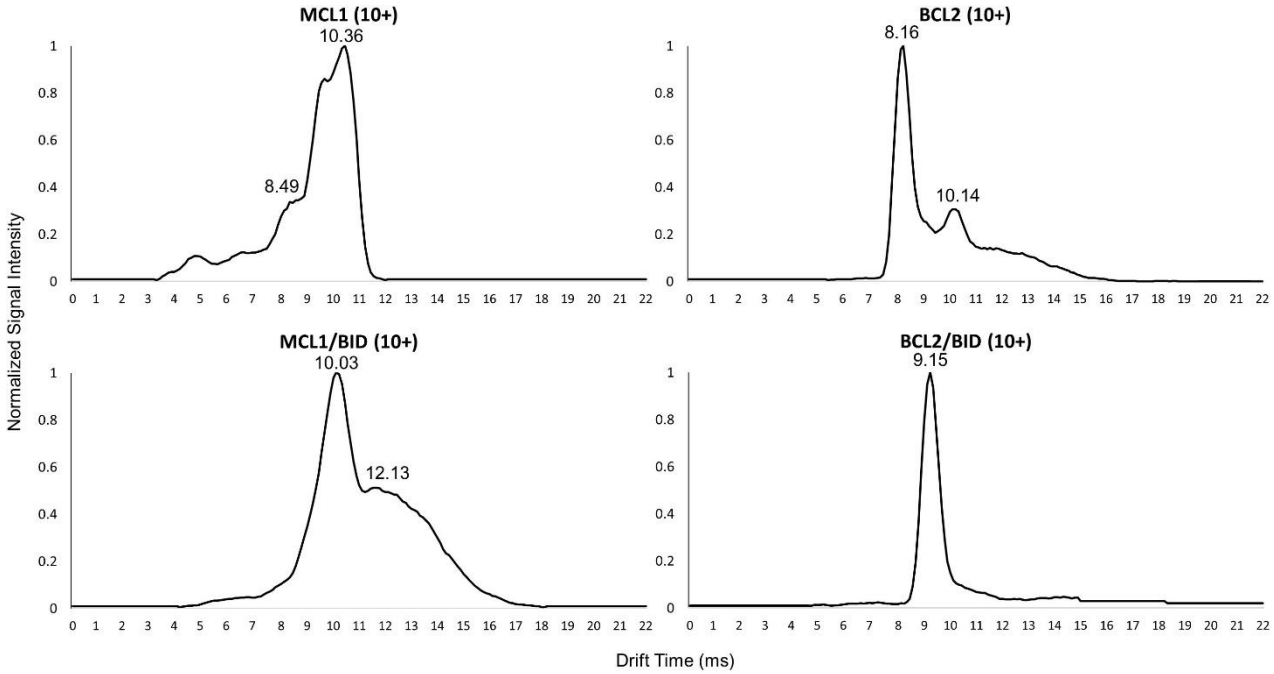


Figure B2. Ion Mobility Chromatograms of Bcl-2 and Mcl-1. The relative signal intensity as a function of drift time (milliseconds) for Mcl-1 (top left), Bcl-2 (top right), Mcl-1 + Bid (bottom left), and Bcl-2 + Bid (bottom right) is shown for their 10+ charge state. These 10+ charge states are less abundant and lower intensity (ion count).

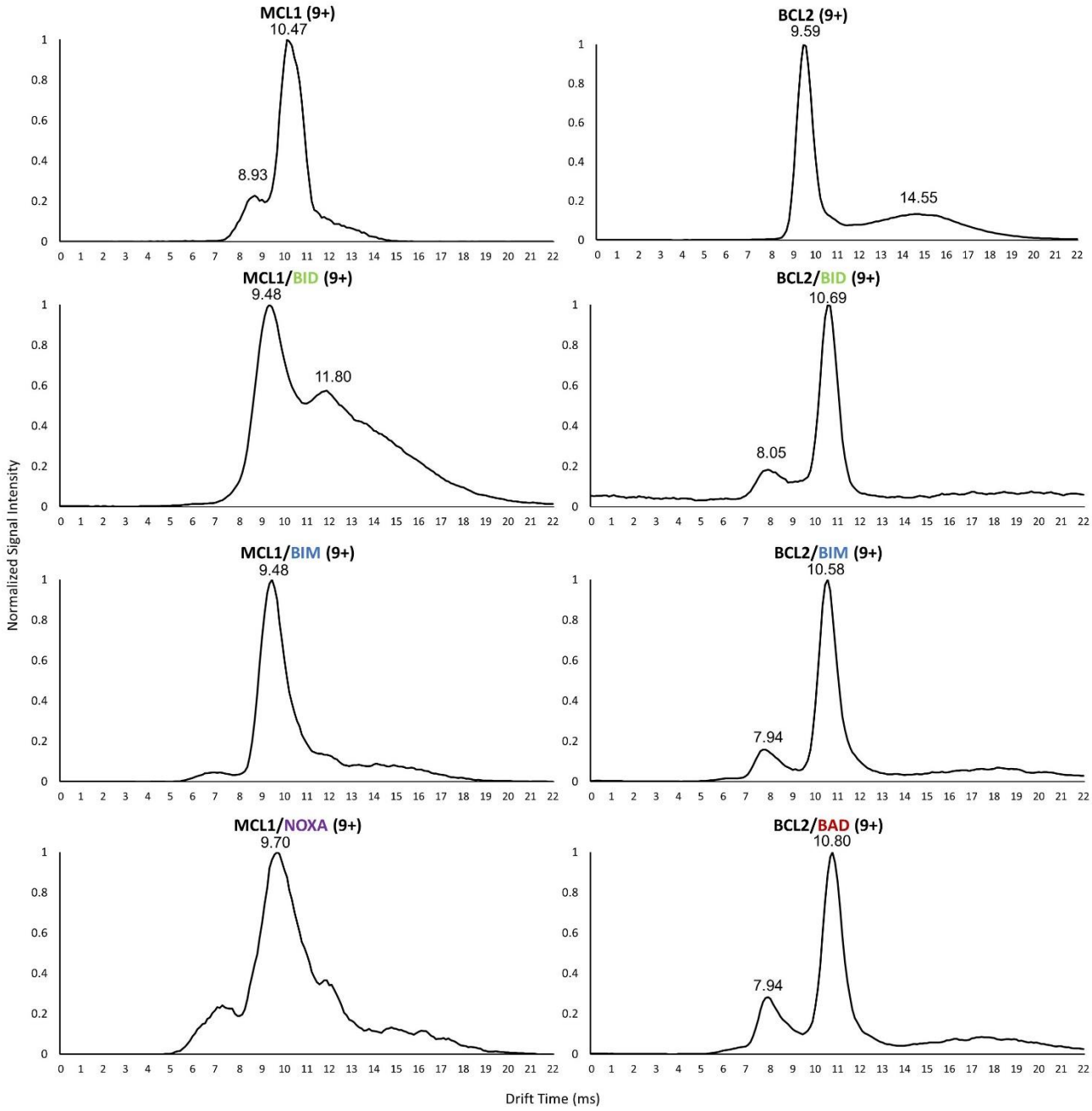


Figure B3. Ion Mobility Chromatograms of Bcl-2 and Mcl-1 with all BH3 peptides. The relative signal intensity as a function of drift time (milliseconds) for the 9+ charge states of Mcl-1 (top left), Bcl-2 (top right), and remainder of bound complexes: Bcl2/Bim, Bcl2/Bad, Mcl1/Bim, and Mcl1/Noxa. Mcl-1 + Bid (bottom left), and Bcl-2 + Bid (bottom right) is shown for their 10+ charge state.

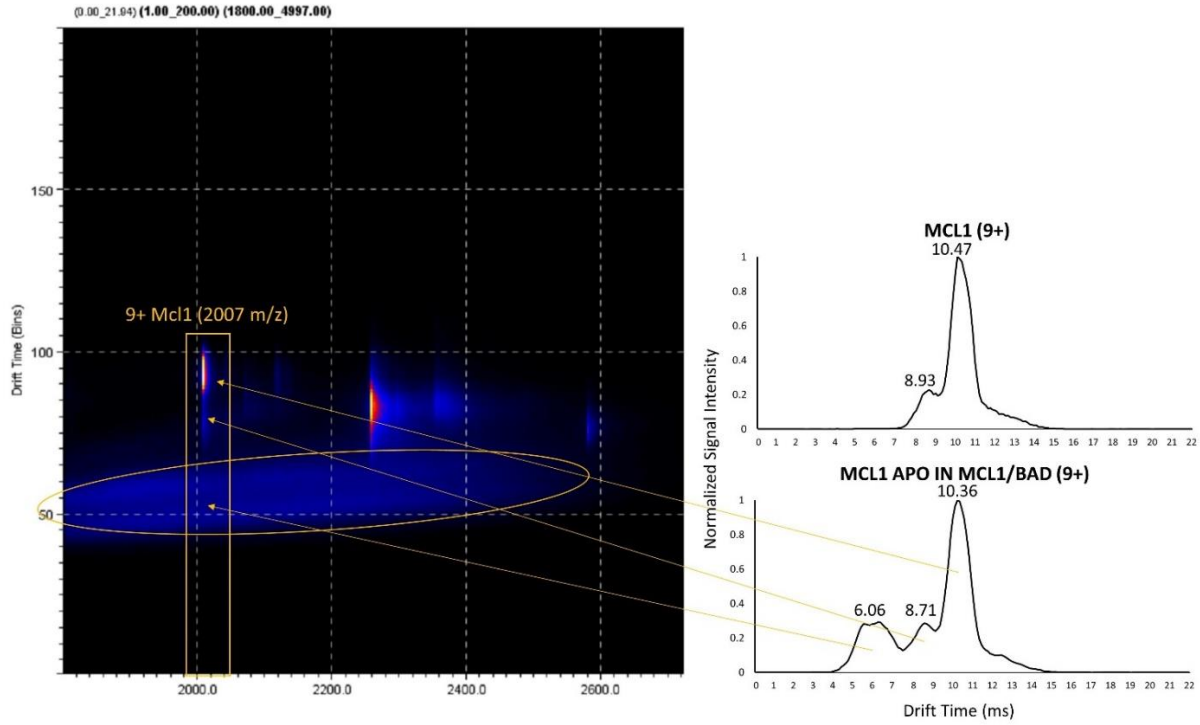


Figure B4. Ion Mobility Chromatograms of unbound Mcl-1 in the "Mcl-1 only" and "Mcl-1/BAD" spectra. The relative signal intensity is plotted as a function of drift time (milliseconds) for the 9+ charge state. The Mcl-1 unbound 9+ peak in Mcl-1/BAD has an artifactual drift time peak at 6.06 which is a result of noise in the sample acquisition (see heat map on the left) to observe noise across a 500 m/z window which does not relate to the species of interest.

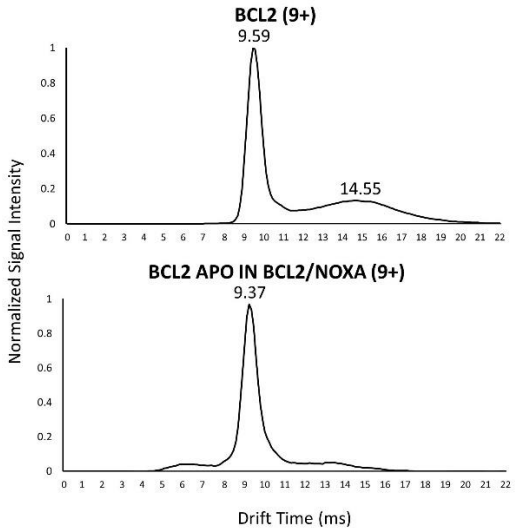


Figure B5. Ion Mobility Chromatograms of unbound Bcl-2 in the "Bcl-2 only" and "Bcl-2/NOXA" spectra. The relative signal intensity is plotted as a function of drift time (milliseconds) for the 9+ charge state. The majority peak of the Bcl2-only 9+ charge state was the same in both states (approximately 9.5 ms).

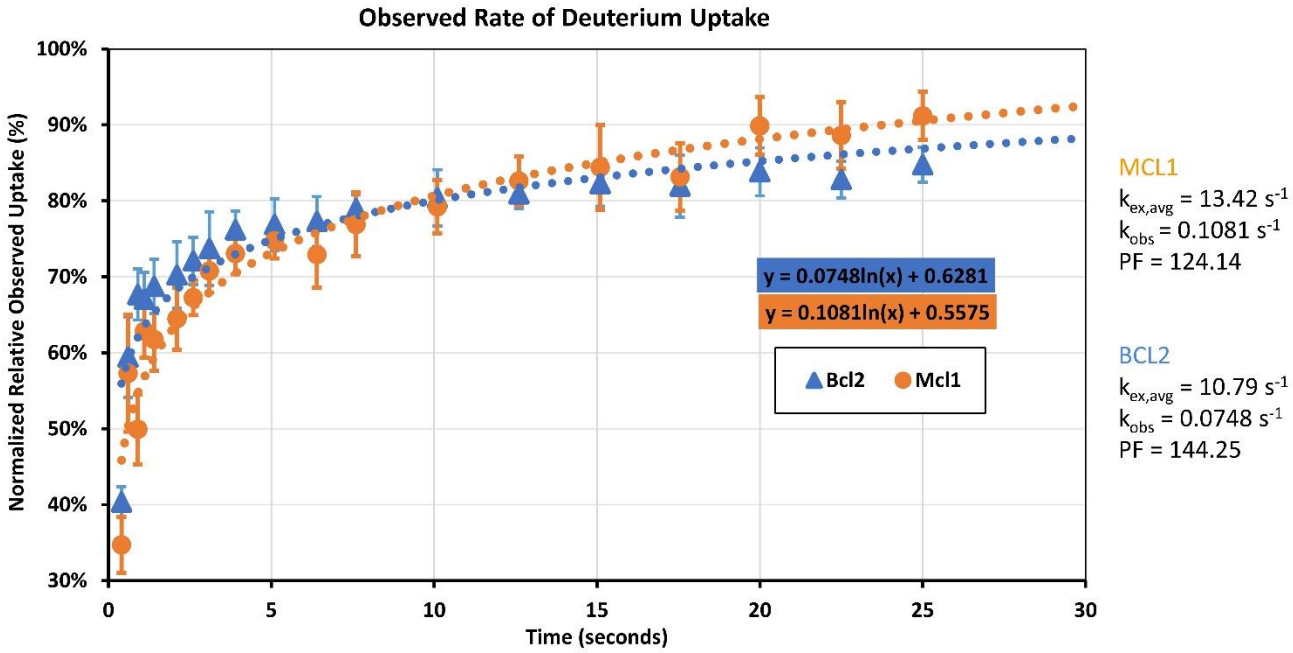


Figure B6. Global TRESI-HDX of intact Bcl-2 and Mcl-1. The normalized relative observed uptake (%) was plotted as a function of time in seconds. Normalized relative uptake was calculated by dividing the deuterium uptake (Da) by the total number of exchangeable amide hydrogens for Bcl-2 and Mcl-1 (165 and 155, respectively). The blue triangles represent Bcl-2 and orange circles represent Mcl-1. The error bars represent 2 sigma which was obtained from a minimum of n=11 charge states. Protection factors (PF) were determined by taking the quotient of the whole protein average intrinsic exchange rate ($k_{ex,avg}$) obtained by averaging the intrinsic rates calculated by Sphere (https://repository.upenn.edu/ dissertations/AAI9532311/) at pH 7.0, 20 ºC and the observed rate of HDX (k_{obs}) which were extracted from the exponential trendline. The PF of Mcl-1 is less than Bcl-2, indicating that Bcl-2 is more structured and protected from HDX than Mcl-1.

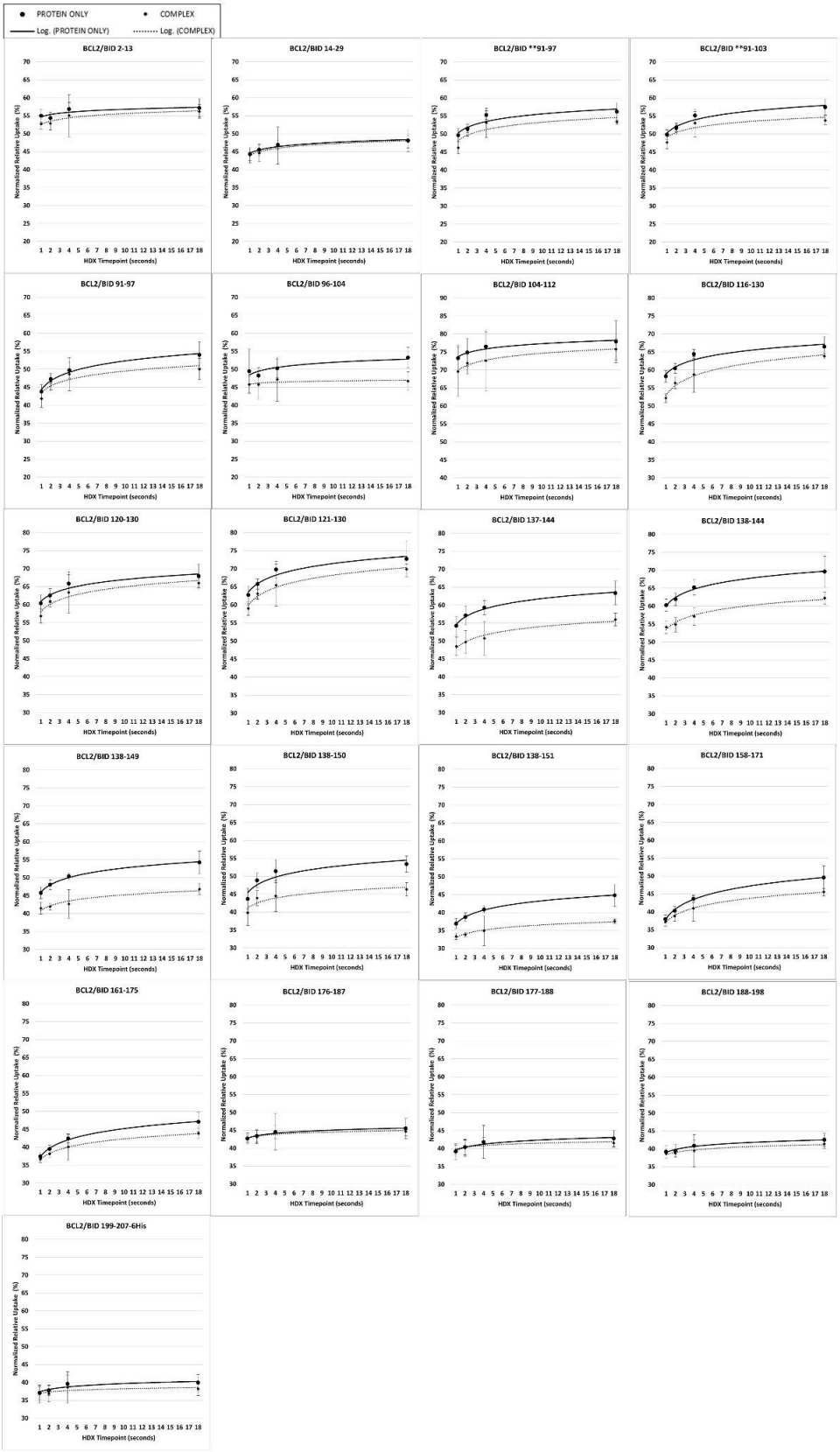


Figure B7. Bcl2 vs. Bcl2/Bid Deuterium Uptake Kinetics Plots.

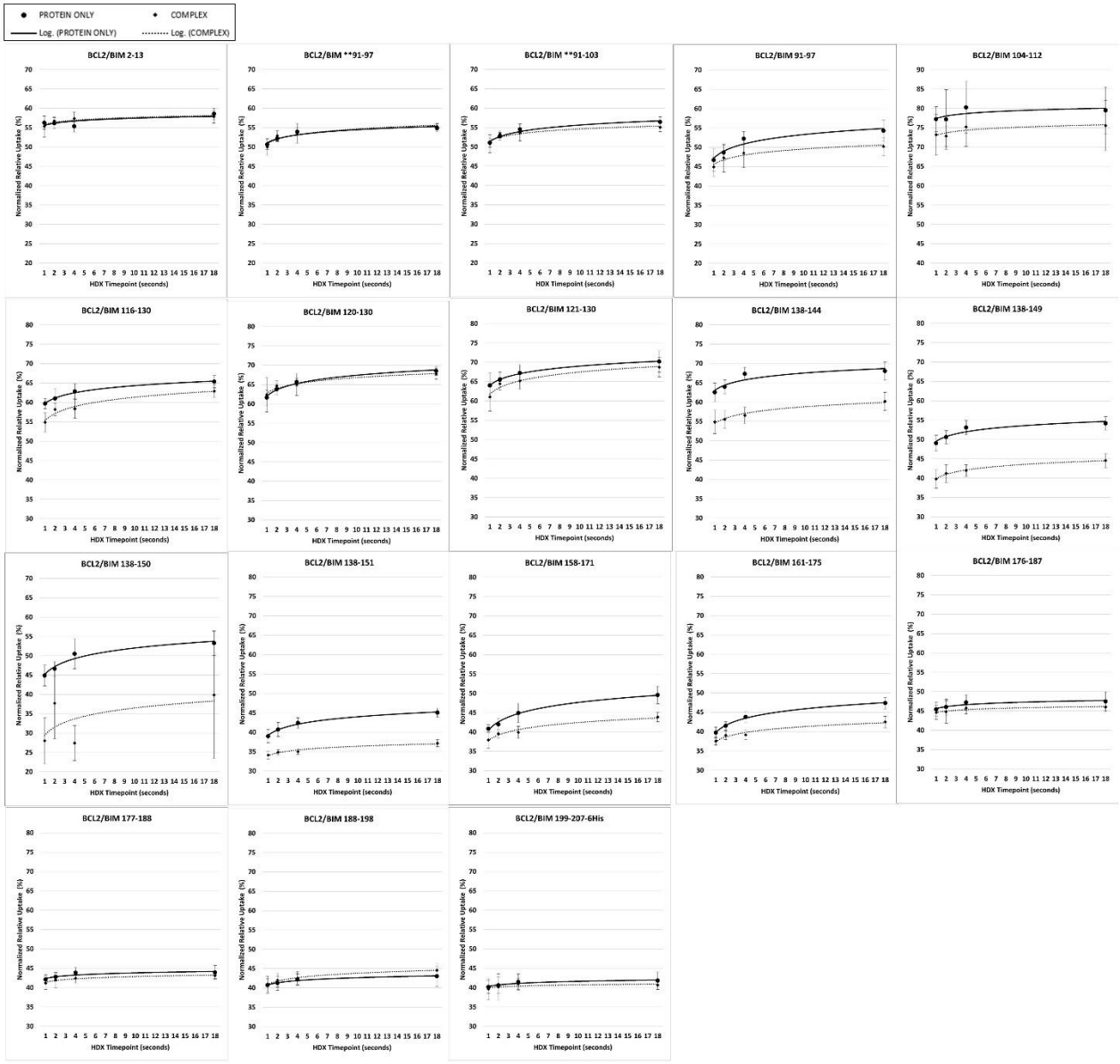


Figure B8. Bcl2 vs. Bcl2/Bim Deuterium Uptake Kinetics Plots.

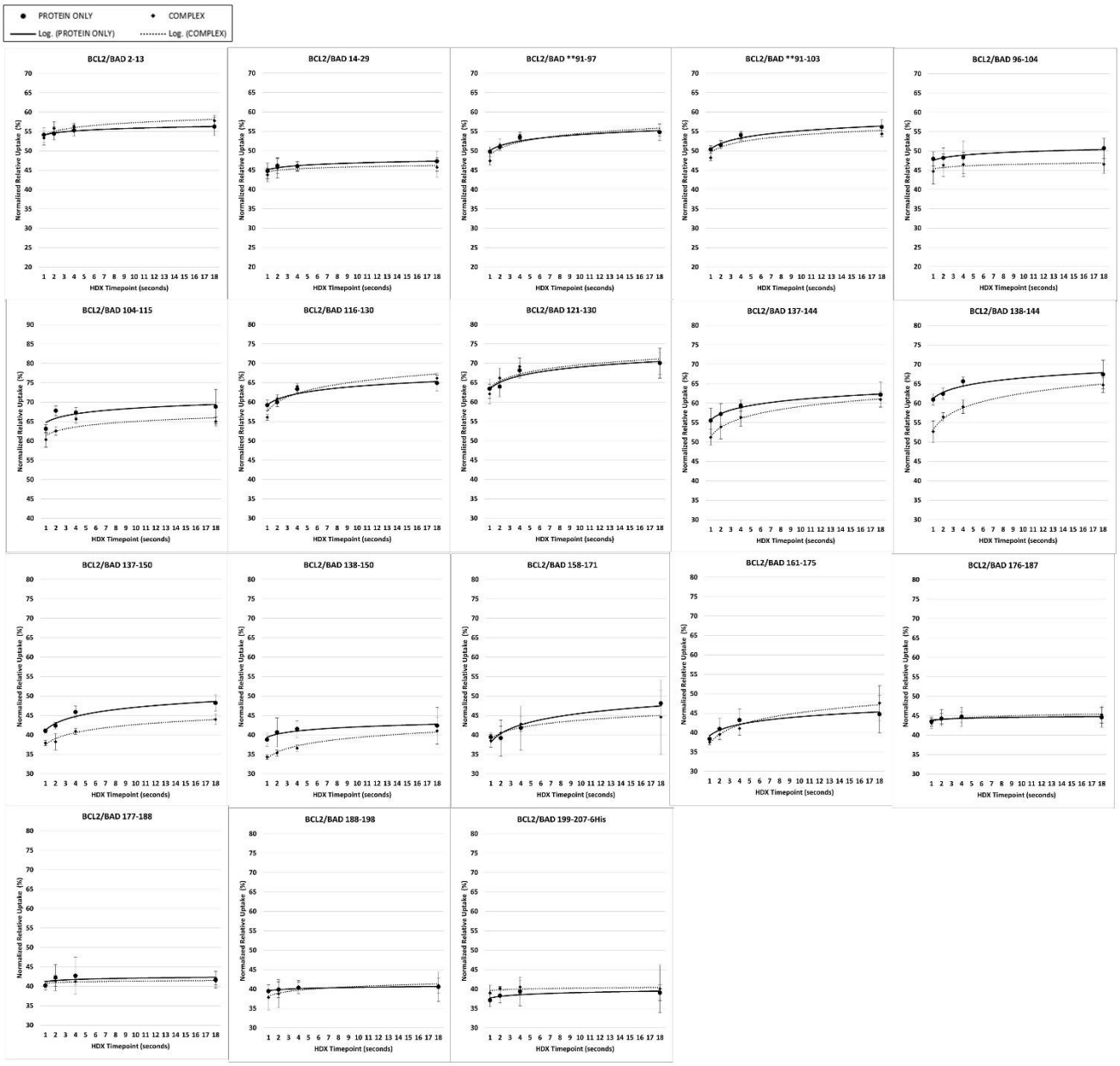


Figure B9. Bcl2 vs. Bcl2/Bad Deuterium Uptake Kinetics Plots.

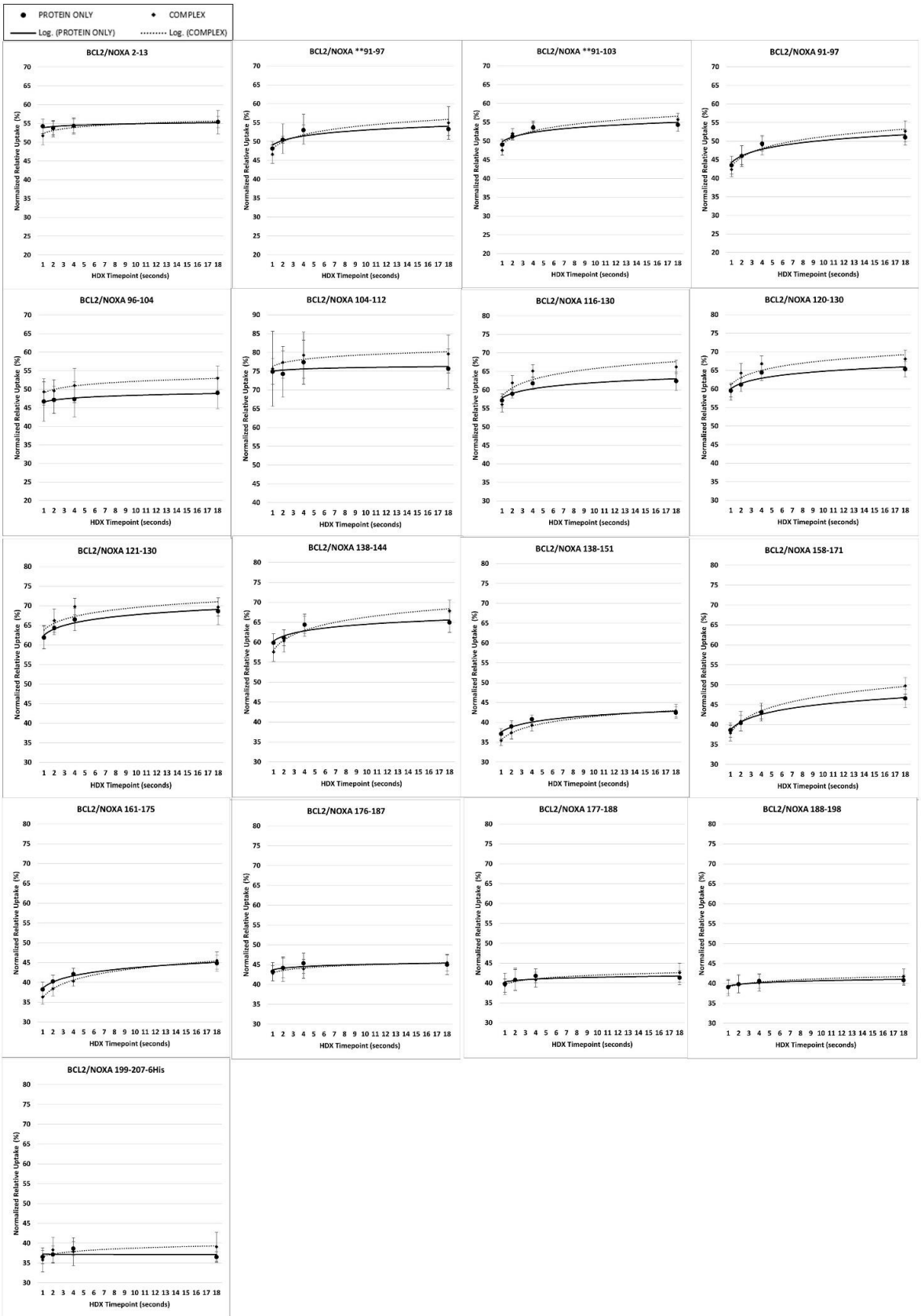


Figure B10. Bcl2 vs. Bcl2/Noxa Deuterium Uptake Kinetics Plots.

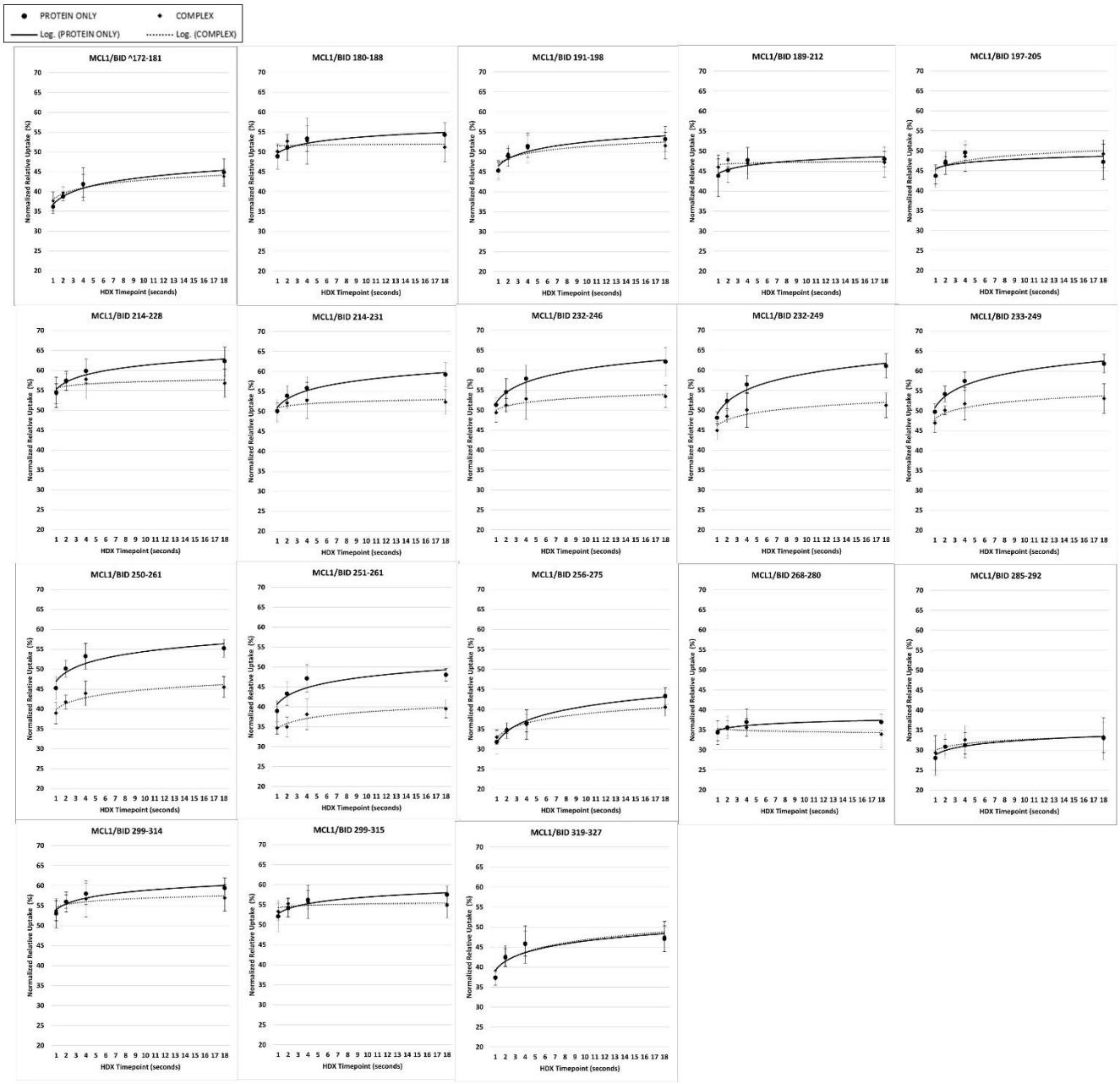


Figure B11. Mcl1 vs. Mcl1/Bid Deuterium Uptake Kinetics Plots.

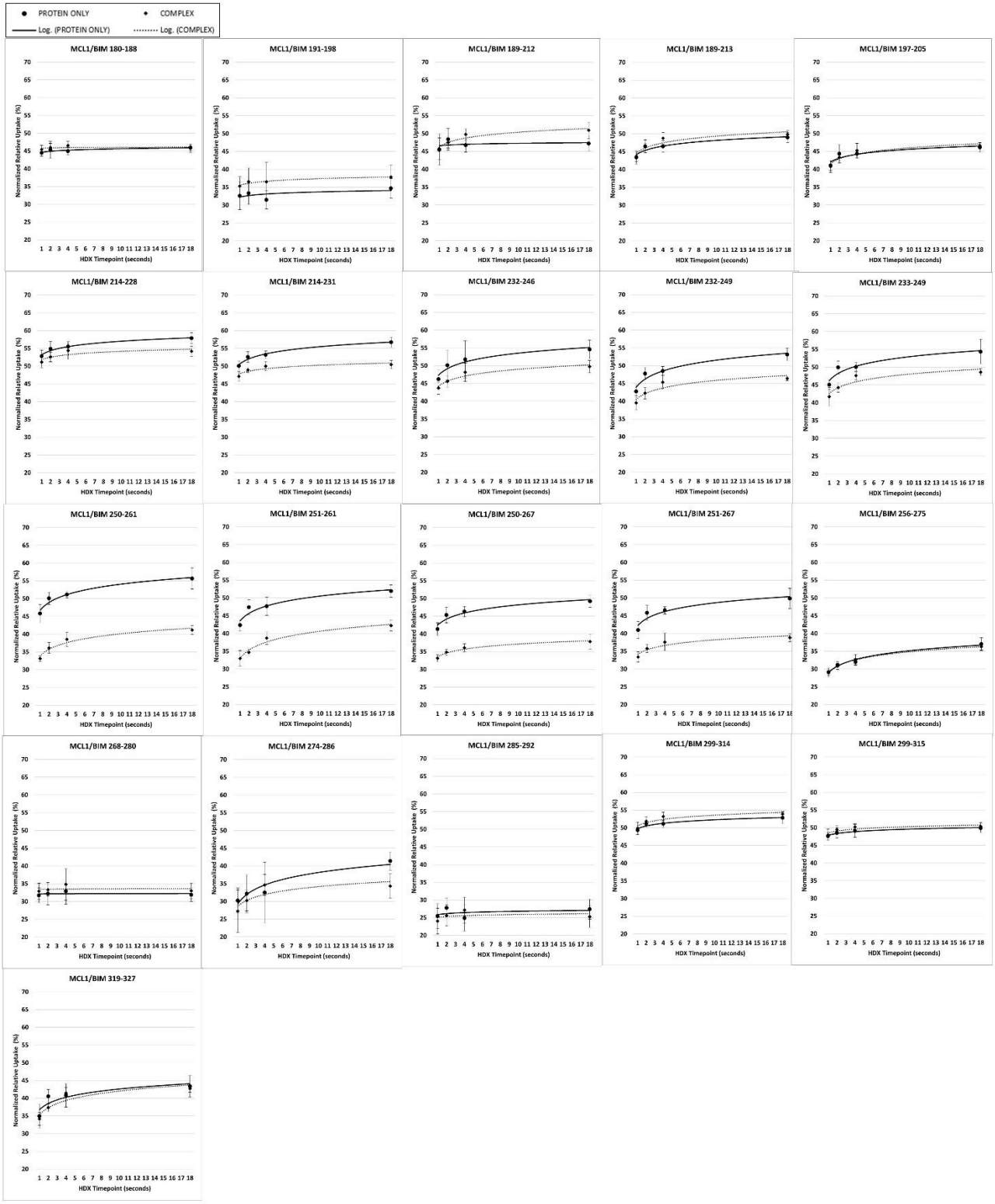


Figure B12. Mcl1 vs. Mcl1/Bid Deuterium Uptake Kinetics Plots.

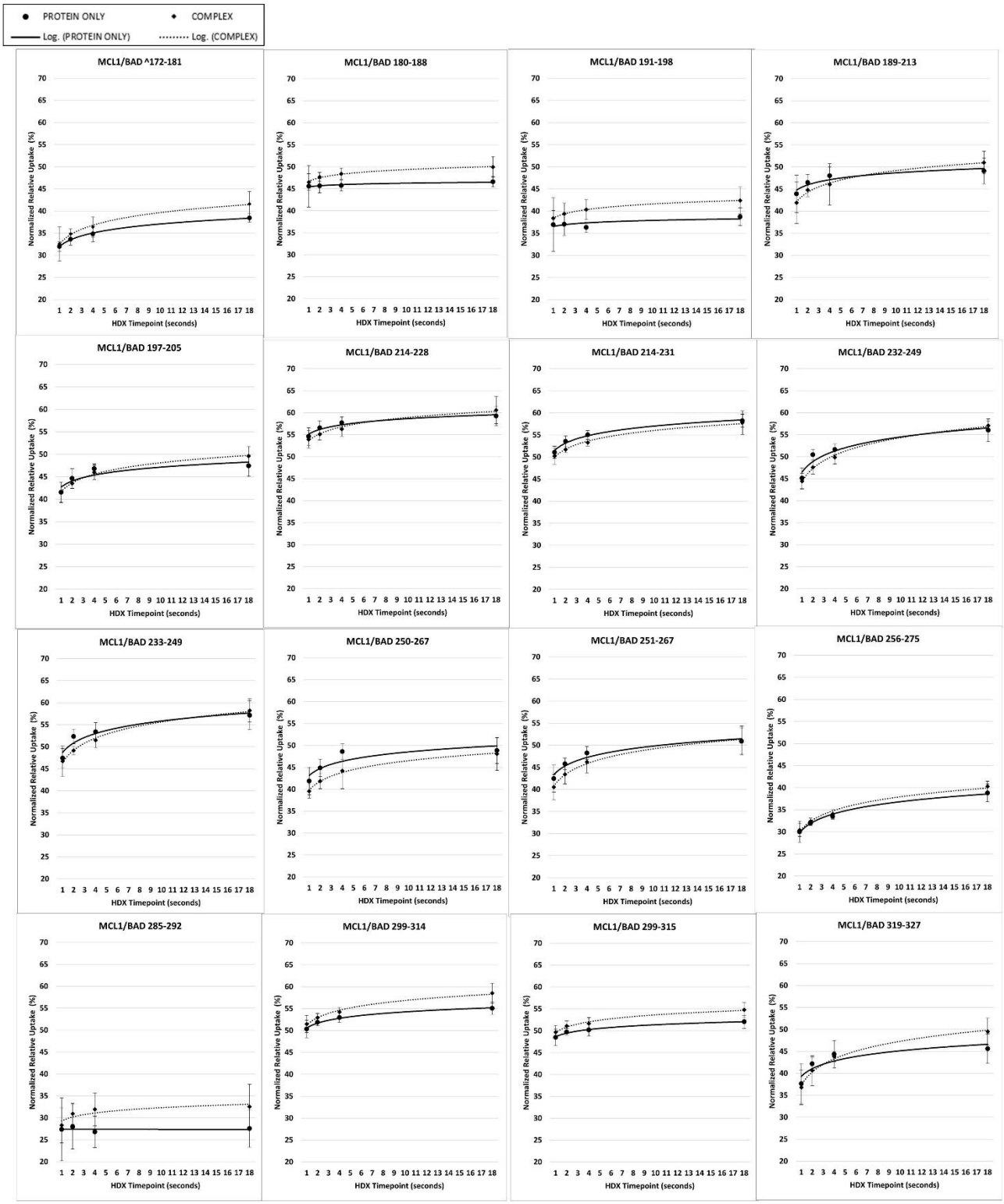


Figure B13. Mcl1 vs. Mcl1/Bad Deuterium Uptake Kinetics Plots.

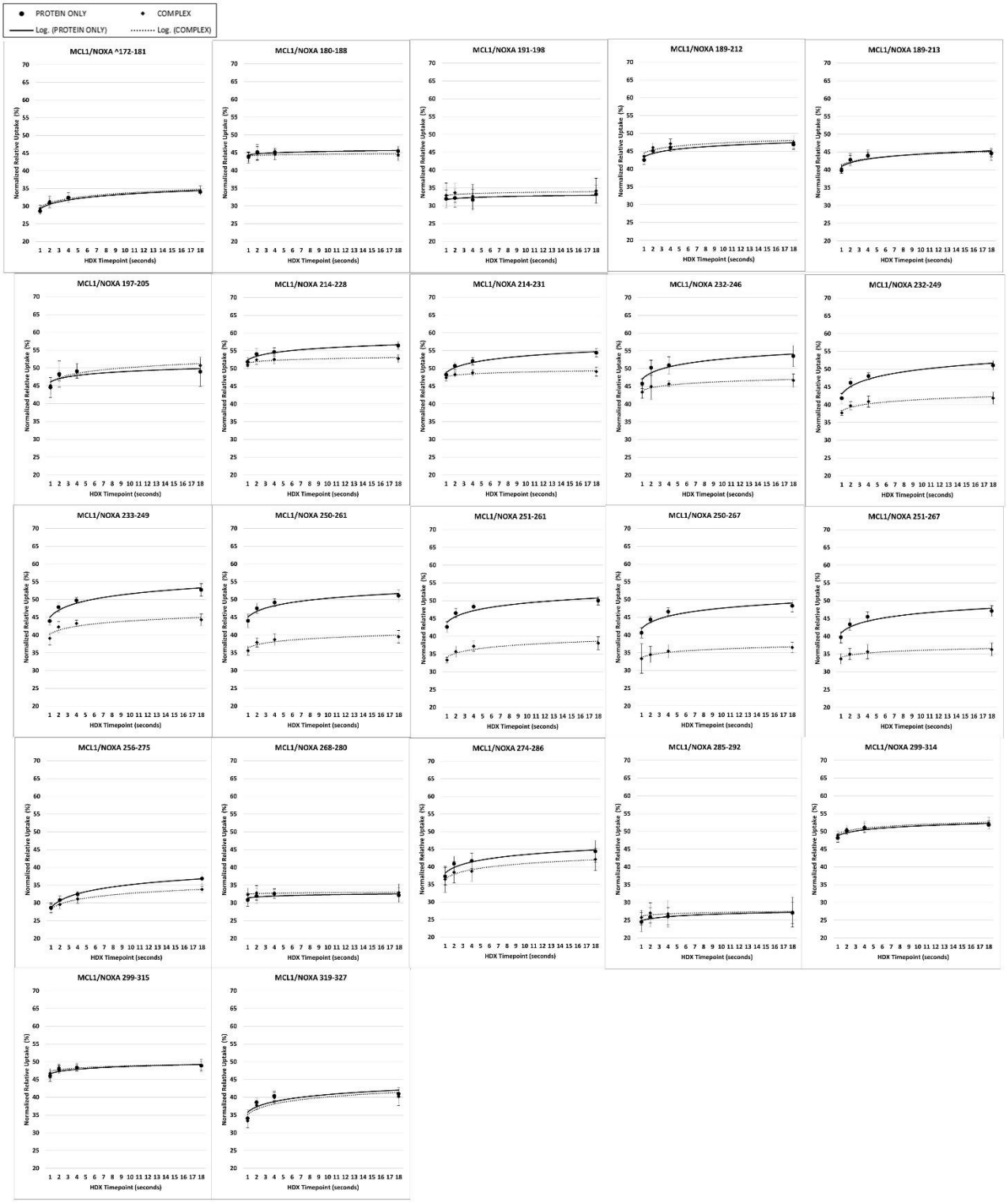


Figure B14. Mcl1 vs. Mcl1/Noxa Deuterium Uptake Kinetics Plots.

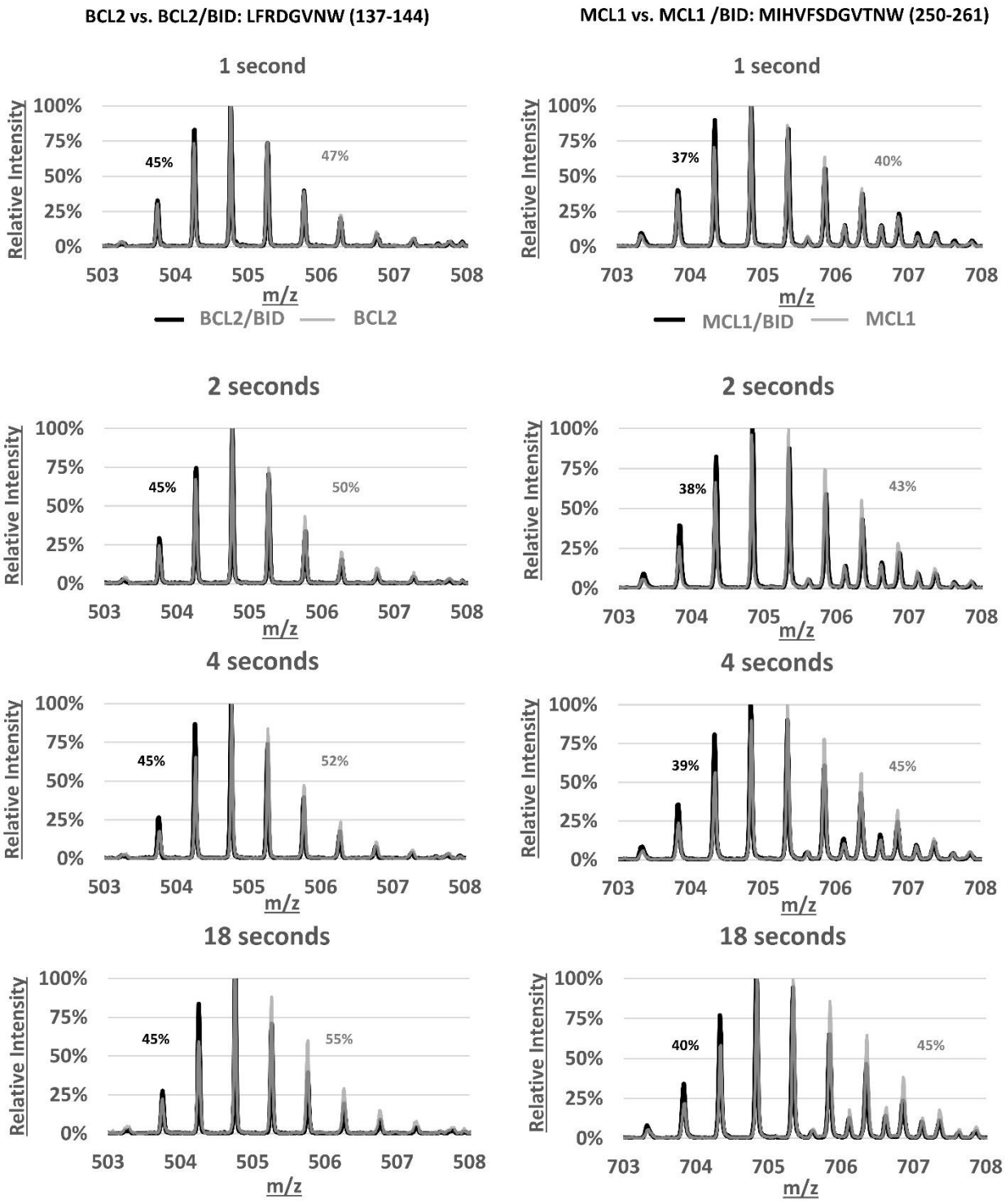


Figure B15. Raw Mass Spectra of Bcl-2 (left) and Mcl-1 (right) experiencing protection from HDX in the presence of BID BH3 peptide. In each case, the peptide corresponding to the protein alone is grey, and the bound protein is black. From top to bottom, the timepoints shown are 1, 2, 4, and 18 seconds. Note that during analysis in MS Studio, spectra were fit using peaks not experiencing spectral overlap.

```
BCL2_Chimera       1 MAHAGRTGYD-----NREIVMKYIHYKLSQRGYEWDAGDDVEENRTEAPE    45
                           |     :.||:.:|:.          :.....::.:.....
MCL1_Trunc         1 ---------DELYRQSLEIISRYLR----------EQATGAKDTKPMGRS    31

                             BH3
BCL2_Chimera      46 GTESEPVVHLTLRQAGDDFSRRYRRDFAEMSSQLHLTPFTARGRFATVVE    95
                     |..|...:. |||:.||...|.:...|..|..:|.:.........:.|:.
MCL1_Trunc        32 GATSRKALE-TLRRVGDGVQRNHETAFQGMLRKLDIKNEDDVKSLSRVMI    80

                          BH1
BCL2_Chimera      96 ELFRDGV-NWGRIVAFFEFGGVMC--VESVNRE--MSPLVDNIALWMTEY   140
                     .:|.||| ||||||....||..:.  ::::|:|  :.||.::|    |:.
MCL1_Trunc        81 HVFSDGVTNWGRIVTLISFGAFVAKHLKTINQESCIEPLAESI----TDV   126

                             BH2
BCL2_Chimera     141 LNRHLHTWIQDNGGWDAFVELY-------GPSMR    167
                     |.|....|:....|||.|||.:       |
MCL1_Trunc       127 LVRTKRDWLVKQRGWDGFVEFFHVEDLEGG----    156

Aligned_sequences: 2
Matrix: EBLOSUM62
Gap_penalty: 10.0
Extend_penalty: 0.5

Length: 184
Identity:      44/184 (23.9%)
Similarity:    72/184 (39.1%)
Gaps:          45/184 (24.5%)
Score: 139.0
```

Figure B16. Sequence Alignment of Bcl-2 Chimera and Truncated Mcl-1. Sequences shown here are those used in this work with cleavage artifacts and purification tags omitted. Sequence alignment completed using EMBOSS Needle. Lines indicate sequence identity whereas dots indicate high (:) or moderate sequence (.) similarity. Bcl-2 Homology domains colour-coded as follows: BH1 (cyan), BH2 (magenta), BH3 (yellow).

Appendix C: Supplementary information for Chapter 3.

Table C1. SMILES and Sequences for 19 WDR5-targeting molecules

	ChemiReg ID / Common Name	SMILES or Amino Acid Sequence
1	XS060584a**/DS0227**	C1=CC(=C(C=C1C1=C(C(=CC=C1)N1C)C=C1)NC(C1=CC=CC(=C1)C)=O)N1CCN(CC1)C
2	XS060479a**/DS0233**	C1=CC(=C(C=C1C1=CN=C(C2=C1)C=CC=C2)NC(C1=CC=CC(=C1)C)=O)N1CCN(CC1)C
3	XS060480a**/DS0234**	C1=CC(=C(C=C1C(=CN=C1)C(=C1C1)C=CC=1)NC(C1=CC=CC(=C1)C)=O)N1CCN(CC1)C
4	XS068123a**/DS0271**	C1=CC(=C(C=C1C1=C(C=CC=C1)CN(CCOC1)C1)NC(C1=CC=CC(=C1)C)=O)N1CCN(CC1)C
5	XS068124a**/DS0273**	C1=CC(=C(C=C1C1C(=C2SC=1)C=CC=C2)NC(C1=CC=CC(=C1)C)=O)N1CCN(CC1)C
6	XS068128a**/DS0278**	C1=CC(=C(C=C1[N+](=O)[O-])NC(C(=N1)C2=C(N1C)C=CC=C2)=O)N1CCN(CC1)C
7	XS083358a**/DS0324**	C1=CC(=C(C=C1[N+](=O)[O-])NC(C(=CC=C1)N2C1=CC=N2)=O)N1CCN(CC1)C
8	XS083452a**/DS0334**	C1=CC(=C(C=C1[N+](=O)[O-])NC(=O)C(=CC(=N1)C)C2=C1C=CC(=C2)O)N1CCN(CC1)C
9	XS083453a**/DS0335**	C1=CC(=C(C=C1[N+](=O)[O-])NC(=O)C1=CC(=C2C=C1)NN=C2)N1CCN(CC1)C
10	MT000951b**/DS0413**	C1=CC(=C(C=C1C1=CC=C(N=C1)NC(C)=O)NC(C(=CNC(C1)=O)C=1C(F)(F)F)=O)N1CCN(CC1)C
11	MT000953b**/DS0415**	C1=CC(=C(C=C1C1=CC(=C2C=C1)C=CN2)NC(C(=CNC(C1)=O)C=1C(F)(F)F)=O)N1CCN(CC1)C
12	MT005481d**/OICR-9429**	C1=CC(=C(C=C1C1=CC(=CC=C1)CN1CCOCC1)NC(C(=CNC(C1)=O)C=1C(F)(F)F)=O)N1CCN(CC1)C
13	MT000954a**/OICR-0547**	C1=CC(=CC(=C1)CN1CCOCC1)C1=CC=C(C(=C1)NC(=O)C(=CNC(C1)=O)C=1C(F)(F)F)N1CCOCC1
14	**MM-102**	CCC(CC)(C(=O)NC(CCCN=C(N)N)C(=O)NC1(CCCC1)C(=O)NC(C2=CC=C(C=C2)F)C3=CC=C(C=C3)F)NC(=O)C(C)C
15	**WDR5-0103**	O=C(OC)C1=CC(NC(C2=CC=CC(OC)=C2)=O)=C(N3CCN(C)CC3)C=C1
16	**WIN-10mer-Ala** (MLL1 Peptide, Q03164)	Ac-^{3764}ARAEVHLRKS^{3773}-Ahx-Ahx-K-NH2 Ac = acetylation, Ahx= 6-amino hexanoic acid
17	**MLL-Win** (MLL1 Peptide, Q03164)	^{3762}GSARAEVHLRKS^{3773}
18	**H3K4Me3-B** (Histone H3.1 Peptide, P68431)	^{2}ARTK(Me$_3$)QTARKSTGGKAPRKQLATKAA^{26}K-B Me$_3$ = Trimethylation, B = Biotinylation
19	**RbBP5-DV** (RbBP5 Peptide, Q15291)	Y^{369}AAEDEEVDVTSVD^{381}

Table C2. Signal-to-noise values for 19 WDR5-targeting molecules.

	ChemiReg ID / Common Name	K_D SPR (μM)	Cumulative	Characteristic	Maximum
1	XS060584a**/DS0227**	10±0	0.18	1.63	0.14
2	XS060479a**/DS0233**	1±0	0.34	4.11	0.38
3	XS060480a**/DS0234**	~116±7	0.1	0.26	-
4	XS068123a**/DS0271**	77±2	0.09	0.14	-
5	XS068124a**/DS0273**	5±0	0.17	1.12	0.08
6	XS068128a**/DS0278**	2±0	0.17	0.61	-
7	XS083358a**/DS0324**	22±2	0.11	0.53	-
8	XS083452a**/DS0334**	60±0	0.16	1.01	0.07
9	XS083453a**/DS0335**	37±0.7	0.2	2.29	0.17
10	MT000951b**/DS0413**	0.02±0.001	0.6	4.49	0.33
11	MT000953b**/DS0415**	0.03±0.002	0.51	6.24	0.46
12	MT005481d**/OICR-9429***	0.02±0.001	0.5	7.22	0.53
13	MT000954a**/OICR-0547****	46±1	0.11	0.3	-
14	**MM-102**	-	0.64	9.57	0.44
15	**WDR5-0103**	-	0.67	7.71	0.35
16	**WIN-10mer-Ala (Peptide)**	-	0.44	8.13	0.78
17	**MLL-Win (Peptide)**	-	0.22	3.95	0.38
18	**H3K4Me3-B (Peptide)**	-	0.32	1.99	0.28
19	**RbBP5-DV (Peptide)**	-	0.2	0.49	0.21

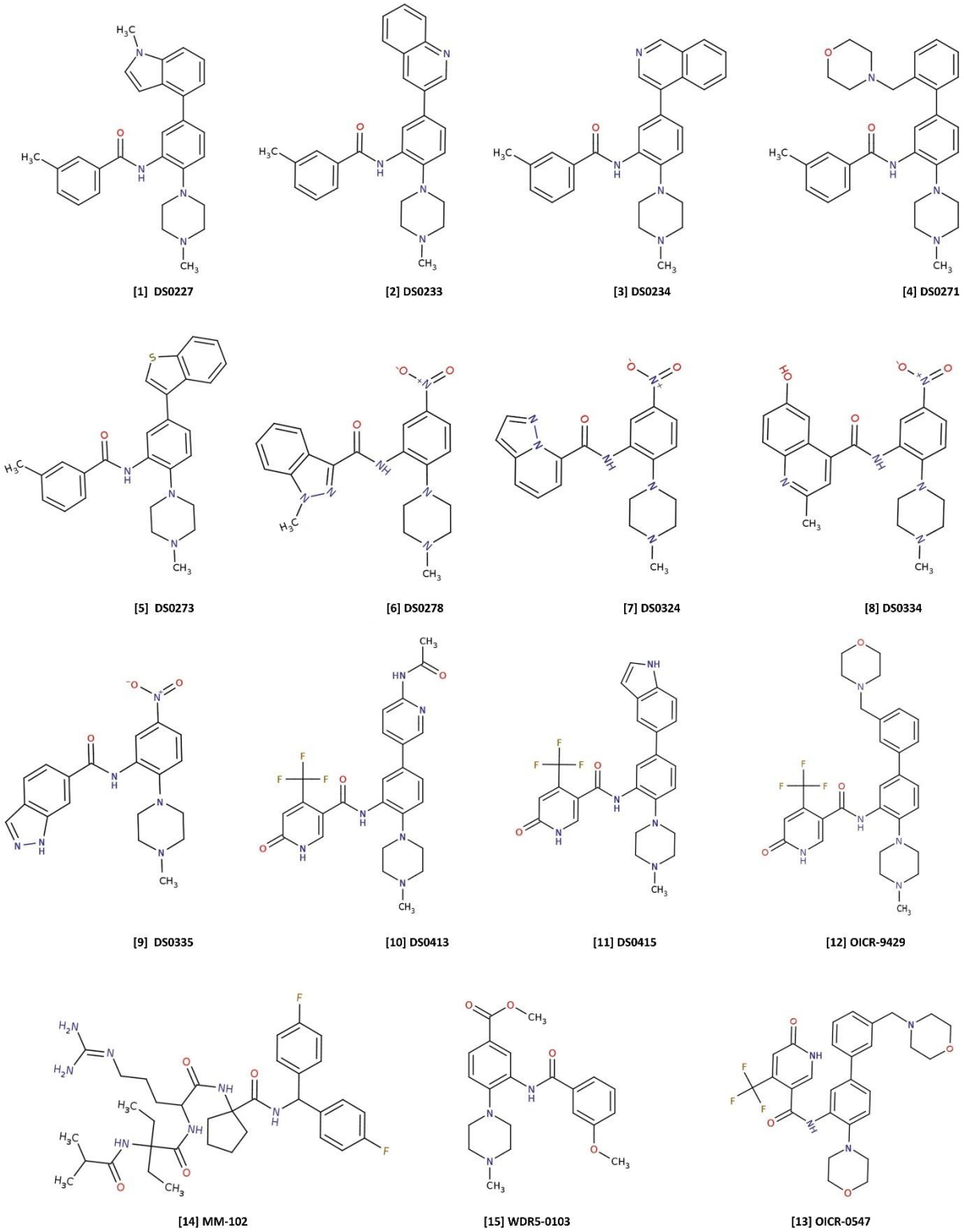


Figure C1. WDR5-Targeting Small Molecule Structures (MolView.org)

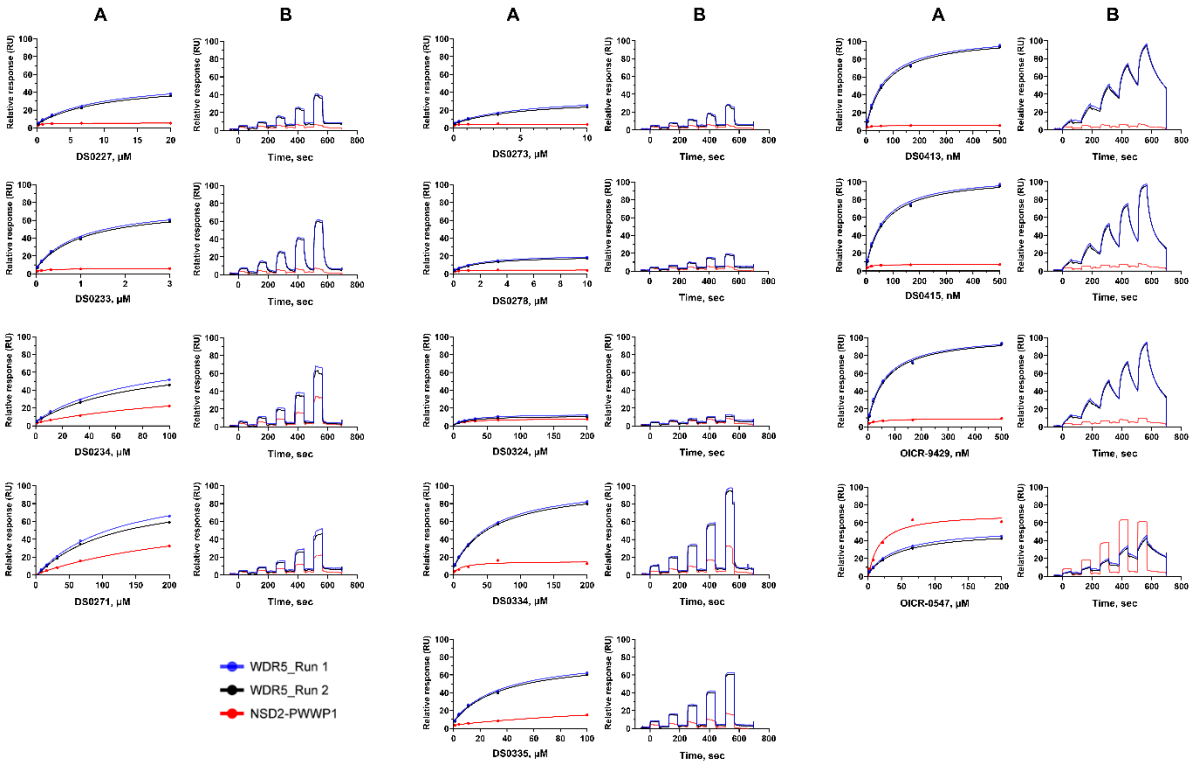


Figure C2. Binding response data obtained by SPR (Performed by O.H.) (A) Response vs concentration plots and (B) Raw sensorgram data profiles for concentrations ranging from 2.5 to 200 µM (compound DS0271, DS0324, DS0334, OICR-0547), from 1.2 to 100 µM (compound DS0234, DS0335), from 0.25 to 20 µM (compound DS0227), from 0.12 to 10 µM (compound DS0273, DS0278), from 0.012 to 3 µM (compound DS0233) and from 6.2 to 500 nM (compound DS0413, DS0415, OICR-9429). The 5-point titration was performed with 3-fold serial dilution in duplicates. Binding response against WDR5 protein (blue and black) and against negative control NSD2-PWWP1 protein (red).

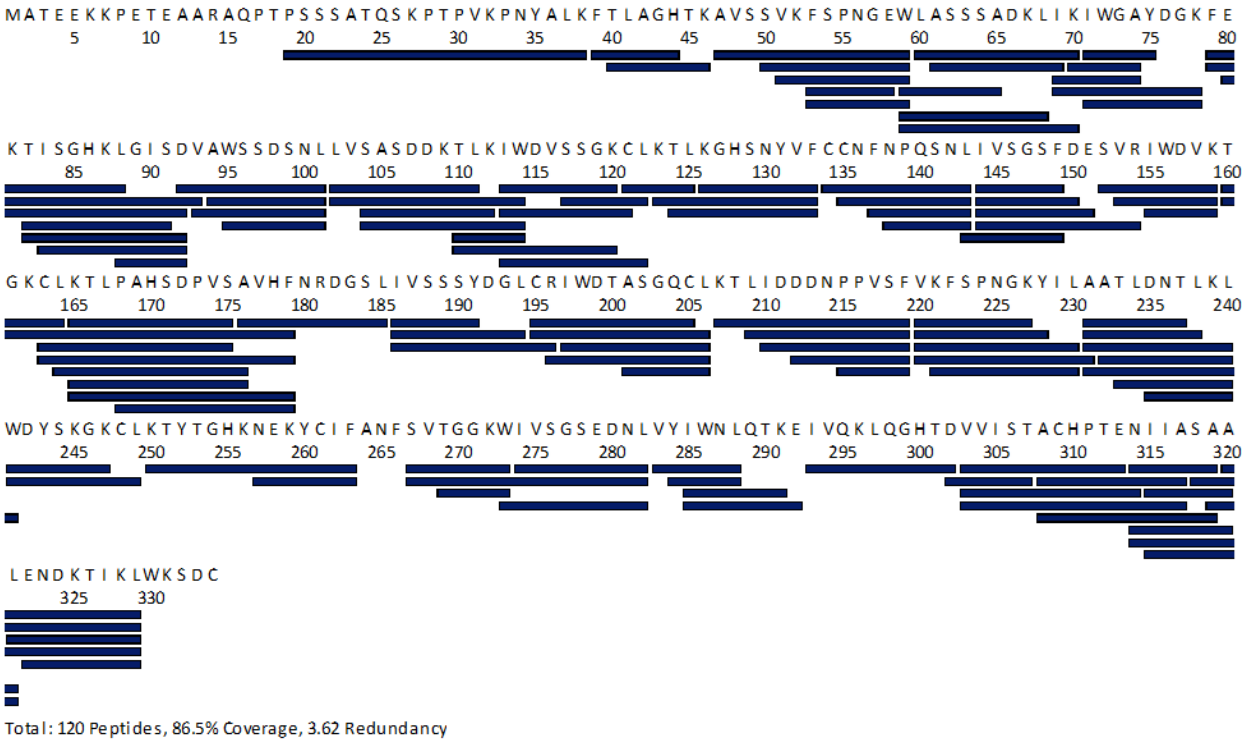


Figure C3. Peptide coverage map of WDR5. A total of 120 peptides were analyzed, yielding a sequence coverage of 86.5% and peptide redundancy of 3.62. Peptides were generated using a 1:1 Nepenthesin-2 Pepsin column (Affipro).

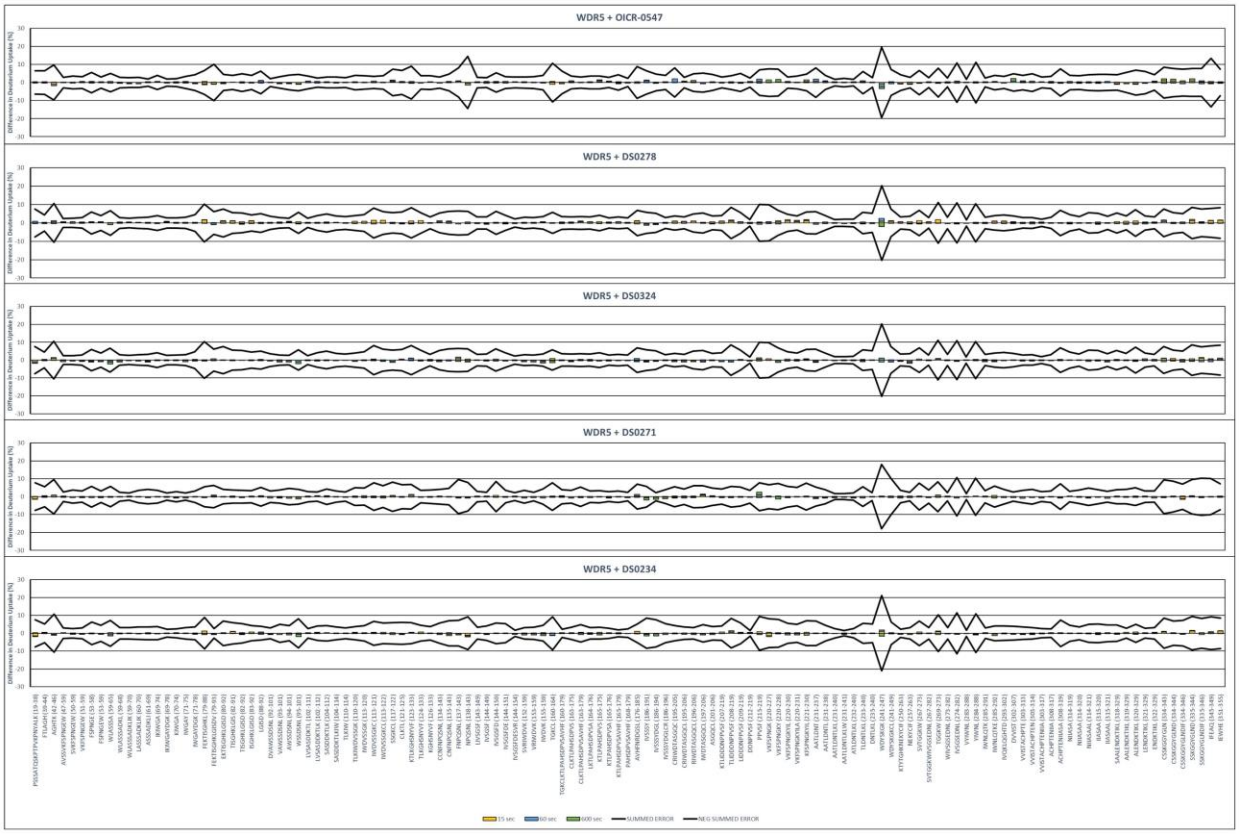


Figure C4. Summed Differences in Deuterium Uptake of Small Molecules with Undetected Binding to WDR5. WDR5 peptide sequences plotted against the extent of D uptake perturbation (%) for each small molecule. To be statistically significant, the summed signal must exceed the summed propagated error (black line). OICR-0547, DS0278, DS0324, DS0271, and DS0234 did not have a statistically significant (HDX-detectable) effect on the WDR5 conformational ensemble.

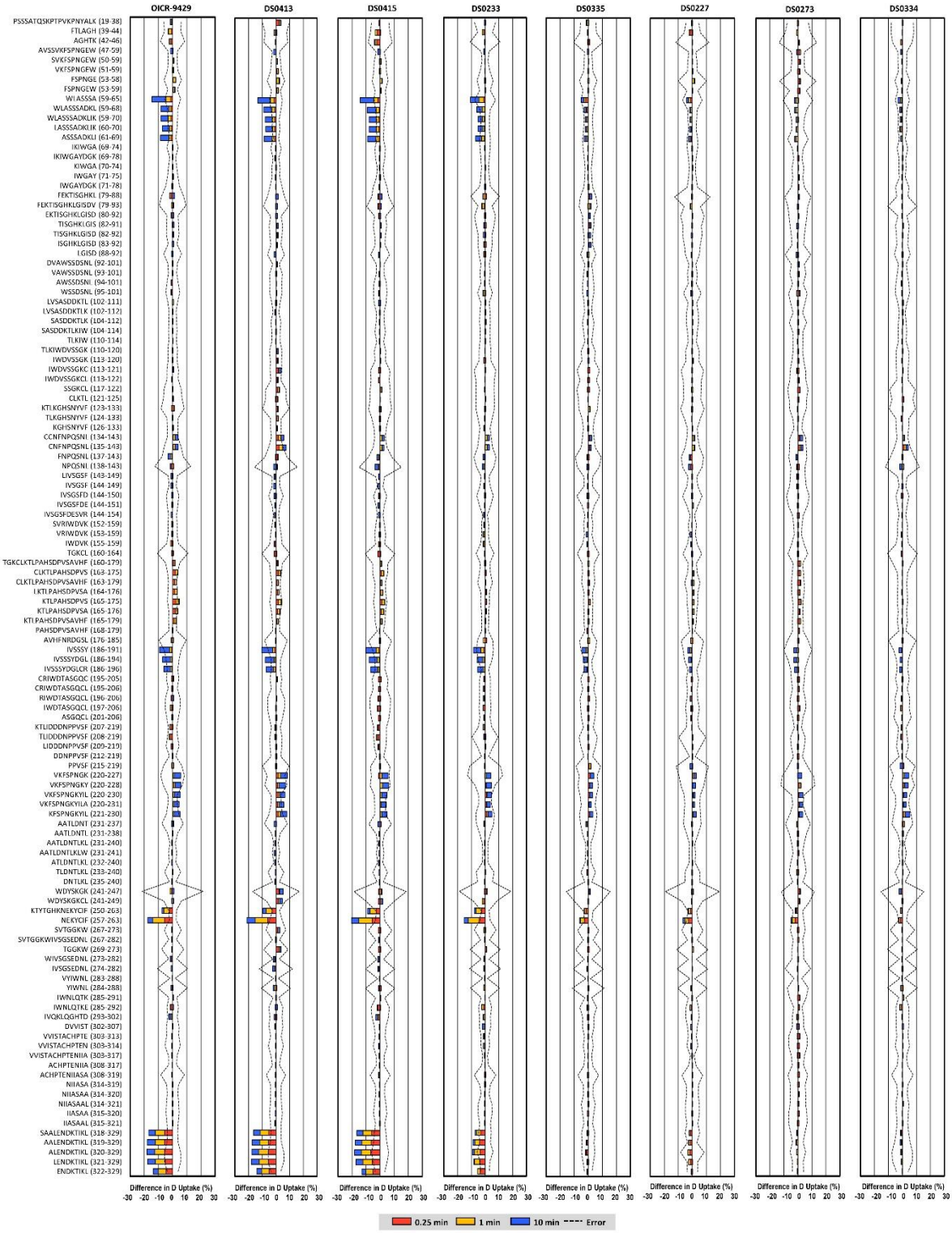


Figure C5. Summed Differences in Deuterium Uptake Aligned to Full Peptide List. To be statistically significant, the summed signal of timepoints 0.25 (red), 1 (yellow), and 10 min (blue) must exceed the summed propagated error (dashed line).

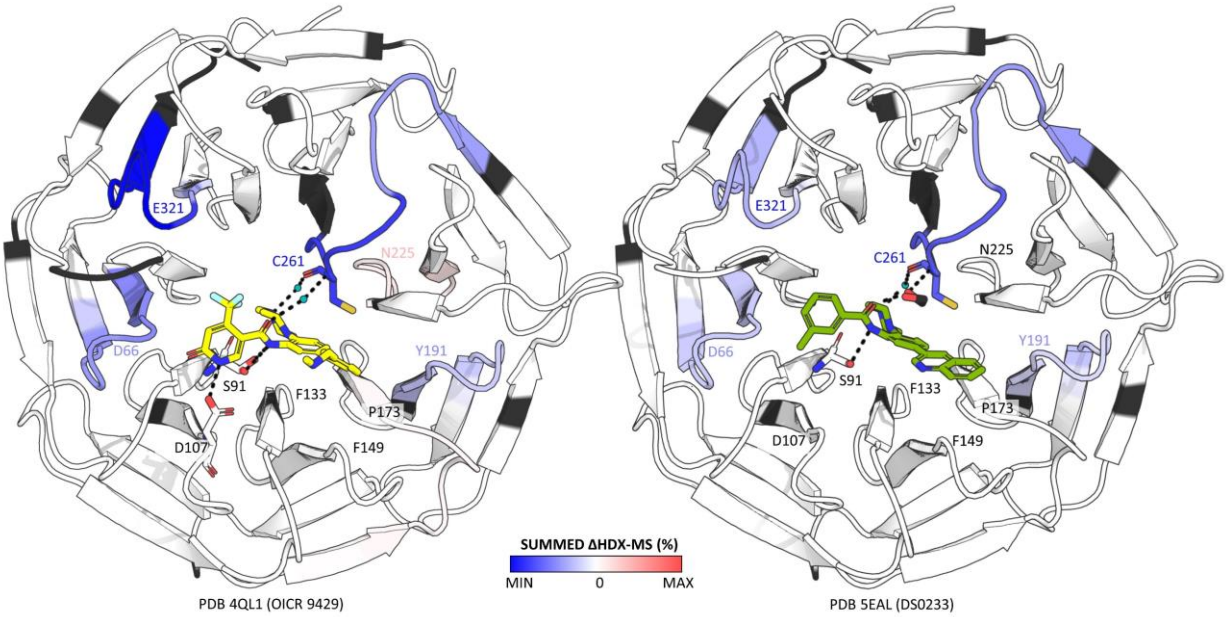


Figure C6. Summed HDX-MS Heatmaps of OICR-9429 (PDB 4QL1)[182] and DS0233 (PBD 5EAL)[195] bound WDR5. To be statistically significant, the summed signal must exceed the summed propagated error (dashed line) and show up as shades of blue (ΣΔHDX-MS < 0) or red (ΣΔHDX-MS > 0) in the WDR5 heatmaps. White corresponds to regions which are not statistically significant and therefore exhibit no change (ΣΔHDX-MS = 0), black indicates no peptide sequence coverage). The loop bearing Cys-261 undergoes HDX-detectable changes in dynamics whereas Ser-91 and Asp-107 are HDX-silent, despite crystal structures indicating their involvement in hydrogen bonds.

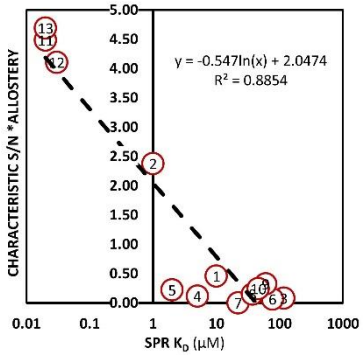


Figure C7. Example of Characteristic Peptide S/N approach using allosteric WDR5 peptide SAALENDKTIKL (318-329). The affinity ranking presents as follows (highest to lowest): WDR-0103 (6.14) > MM-102 (5.39) > OICR-9429 (4.69) > DS0413 (4.48) > DS0415 (4.11) > Win-10mer-Ala (3.71) > DS0233 (1.63) > MLL-Win (1.47) > H3K4Me3-B (1.29) > S/N = 1 (statistical significance cut-off, all S/N < 1 have insignificant allostery). The following molecules fall below the cut-off: DS0227 (0.46) > DS0334 (0.32) > RbBP5-DV (0.29) > OICR-0547 (0.24) > DS0278 (0.22) > DS0335 (0.15) > DS0273 (0.12) > DS0234 (0.08) > DS0271 (0.07) > DS0324 (0.00). Despite the presence of statistically significant differences in the bar plots (Figure 2&3), RbBP5-DV falls below the cut-off due to its distinct binding site, and DS0335, DS0273, and DS0334 are excluded due to undetected allostery.

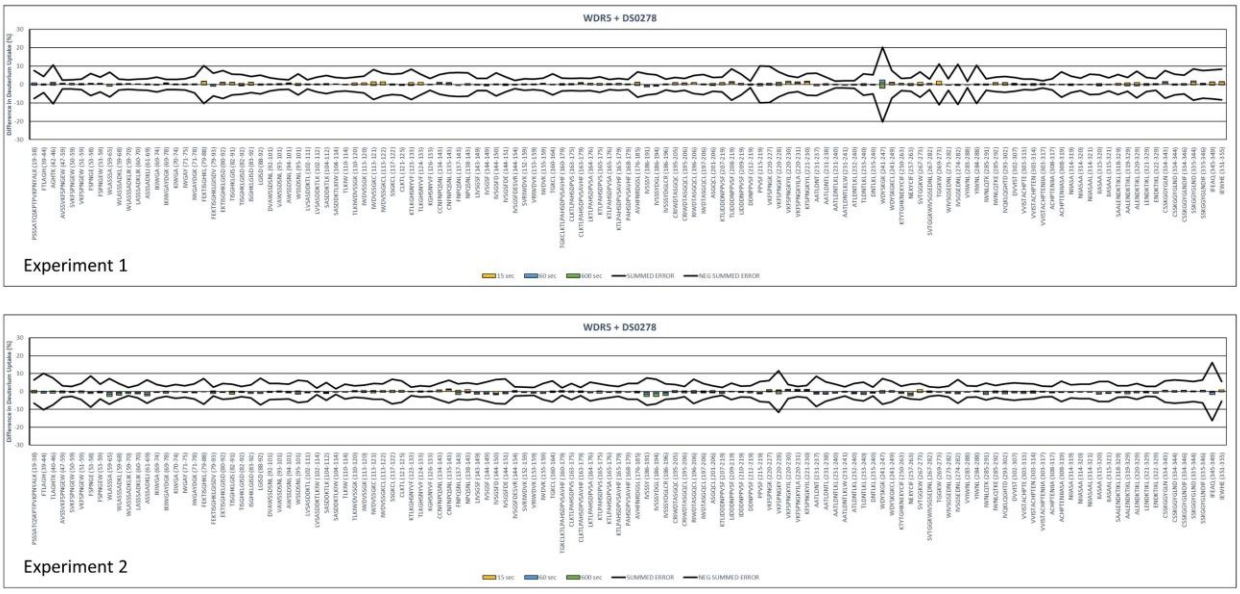


Figure C8. Summed Differences in Deuterium Uptake of DS0278 (Compound 6). WDR5 peptide sequences plotted against the extent of D uptake perturbation (%) for each small molecule. To be statistically significant, the summed signal must exceed the summed propagated error (black line).

Formula C1. Derivation of bound fraction.

Here, K_D is the small molecule's SPR measured dissociation coefficient, P_{eq} , L_{eq} and PL_{eq} are the protein, ligand (small molecule), and the ligand-bound protein concentrations at equilibrium, P_0 and L_0 are the initial concentrations of protein and ligand upon deuterium dilution of the sample stock. All units of variables are kept the same (i.e., μM).[191]

$$K_D = \frac{P_{eq} \times L_{eq}}{PL_{eq}} = \frac{(P_0 - PL_{eq})(L_0 - PL_{eq})}{PL_{eq}}$$

$$0 = (PL_{eq})^2 + (P_0 + L_0 + K_D)(PL_{eq}) + P_0 L_0$$

$$PL_{eq} = \frac{P_0 + L_0 + K_D \pm \sqrt{(P_0 + L_0 + K_D)^2 - 4P_0 L_0}}{2}$$

$$\theta_{eq} = \frac{PL_{eq}}{P_0} \times 100\%$$

Formula C2. Explanation of Summed Propagated Error Calculation.

Percent exchanged deuterium (x) of a peptide.

$$x_{technical\_measurement} = \frac{amide\ D\ exchanged}{exchangeable\ amide\ D} \times 100\ \%$$

Standard deviation (S) of a sample (s) at technical replicate of n=3.

Sample (s) is either unbound (apo) or bound state.

$$\sigma_s = \sqrt{\frac{\left(x_{technical_1} - \bar{x}\right)^2 + \left(x_{technical_2} - \bar{x}\right)^2 + \left(x_{technical_3} - \bar{x}\right)^2}{n}}$$

Propagated error (P) of the difference in deuterium uptake (ΔHDX) of the bound state and unbound state.

$$\Delta HDX_{timepoint} = x_{bound} - x_{unbound}$$

$$P_{\Delta HDX_{timepoint}} = \sqrt{{\sigma_{bound}}^2 + {\sigma_{unbound}}^2}$$

Summed signal ($S_t = \Sigma\Delta HDX$) and summed propagated error (N_t) where t = all timepoints measured, in this case there were three timepoints (15, 60, 600 sec). Note that propagated error is multiplied by 3 to achieve 3 deviations from the mean (99.7% confidence interval).

$$S_t = \Sigma\Delta HDX = \Delta HDX_{15sec} + \Delta HDX_{60sec} + \Delta HDX_{600sec}$$

$$N_t = 3(P_{\Delta HDX_{15sec}} + P_{\Delta HDX_{60sec}} + P_{\Delta HDX_{600sec}})$$

Summed propagated error (see *i*) offers the ability to quickly visualize the summed signal in a single bar plot, while retaining the significance of a signal from a single timepoint ($\Delta HDX_{timepoint}$).

$$(i)\ P_{\Delta HDX_{timepoint}} = 3\left(\sqrt{{S_{bound\_15sec}}^2 + {S_{unbound\_15sec}}^2} + \sqrt{{S_{bound\_60sec}}^2 + {S_{unbound\_60sec}}^2} + \sqrt{{S_{bound\_600sec}}^2 + {S_{unbound\_600sec}}^2}\right)$$

Alternatively, the practice of propagating all three timepoints together (all under one square root, see *ii*) rather than summing each propagated error (see *i*), artificially reduces the total error, making Summed signal ($S_t = \Sigma\Delta HDX$) significant when it should not be (i.e., the signals from the individual timepoints are not significant). Particularly, because these technical timepoint measurements are independent of one another (separate vials, separately loaded onto digestion/LC columns, separately measured by MS). Please note that the formulas *i* and *ii* are ***not*** equal.

$$(ii)\ P_{\Delta HDX_{timepoint}} = 3\sqrt{{S_{bound\_15sec}}^2 + {S_{unbound\_15sec}}^2 + {S_{bound\_60sec}}^2 + {S_{unbound\_60sec}}^2 + {S_{bound\_600sec}}^2 + {S_{unbound\_600sec}}^2}$$

Appendix D: Supplementary information for Chapter 4.

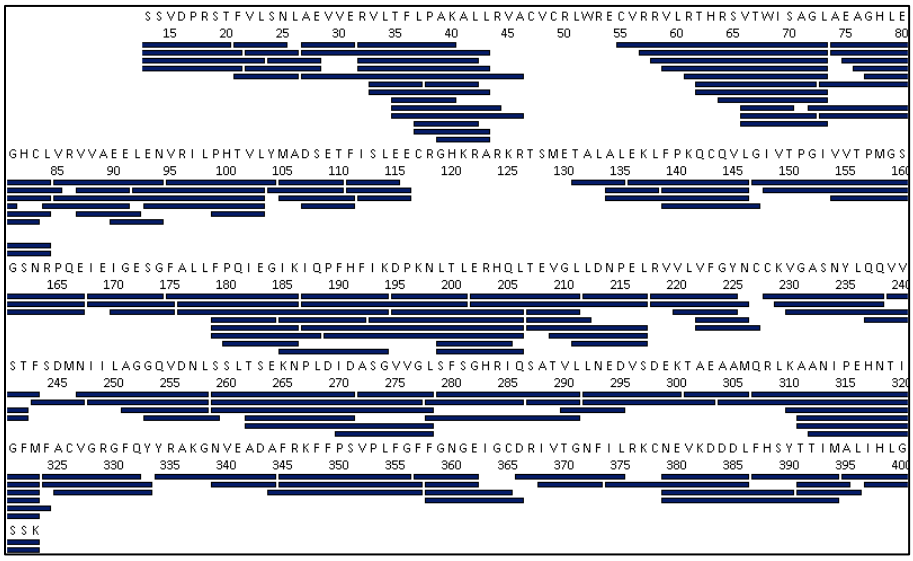


Figure D1. **FBXO22 Sequence Coverage**. Rectangles aligned to FBXO22 primary sequence represent peptides generated using Nep2-Pep proteolysis column. A total of 157 peptides were analyzed, resulting in 91.2% coverage of the protein construct and an average 4.17 peptides per residue.

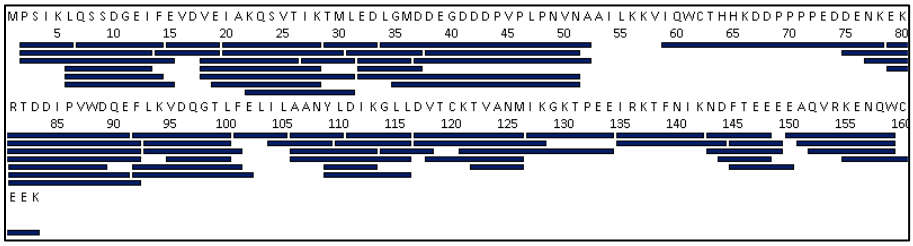


Figure D2. **SKP1 Sequence Coverage**. Rectangles aligned to SKP1 primary sequence represent peptides generated using Nep2-Pep proteolysis column. A total of 67 peptides were analyzed, resulting in 95.7% coverage of the protein construct and an average 4.09 peptides per residue.

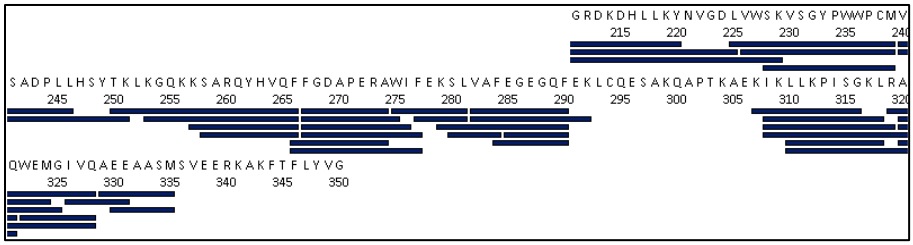


Figure D3. **NSD2_PWWP1 Sequence Coverage**. Rectangles aligned to NSD2 primary sequence represent peptides generated using Nep2-Pep proteolysis column. A total of 40 peptides were analyzed, resulting in 70.9% coverage of the protein construct and an average 3.49 peptides per residue.

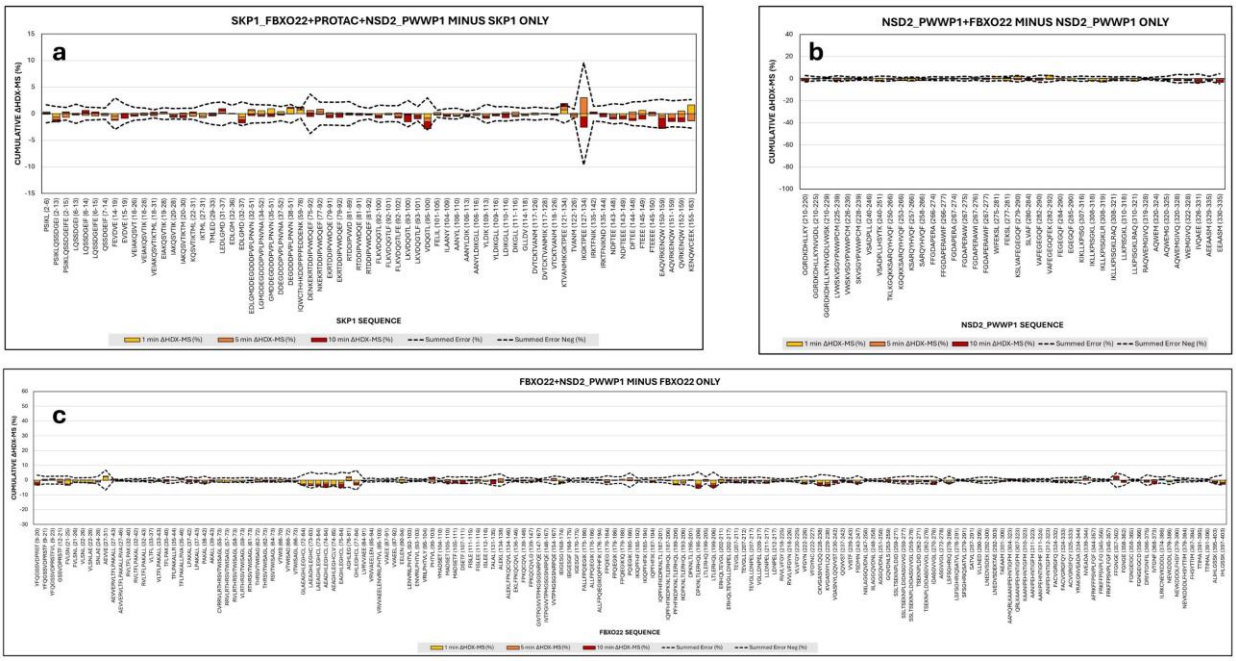


Figure D4. **Insignificant ΔHDX-MS.** (a) SKP1 vs. PWWP1/UNC10088, (b) PWWP1 vs. FBXO22, (c) FBXO22 vs. PWWP1. HDX was conducted at 1 (yellow), 5 (orange), and 10 minutes (red). To be statistically significant, the cumulative ΔHDX-MS signal must have exceeded the cumulative error (dashed line) by >1 %. Using this metric, no peptides in the above graphs undergo changes in dynamics, consistent with an absence of interaction.

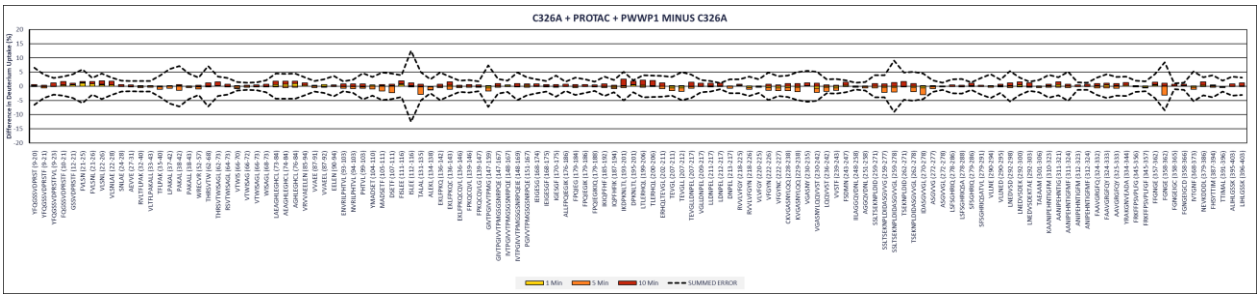


Figure D5. **UNC10088/C326A Mutant.** The absence of statistically significant ΔHDX-MS upon incubation of UNC10088 with FBXO22 C326A mutant is indicative of abrogated interaction.

Appendix E: Supplementary information for Chapter 5.

Table E1. PLGS filtering parameters. *

Parameter	Min. intensity	Min. sequence length	Max. sequence length	Min. products per amino acid	Minimum score	Max. MH+ Error (ppm)
Value	20000	4	26	0.12	7	5

*Additional manual inclusion of several peptides which were initially filtered out resulted in 60 peptides used out of 225 identified by PLGS. Manual inclusion was used to increase sequence coverage in regions missing peptides.

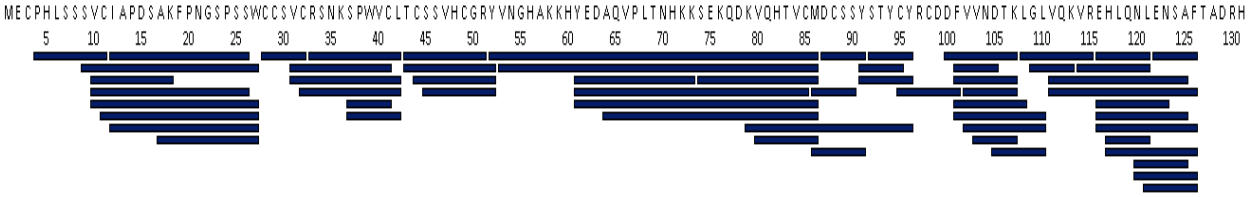


Figure E1. **Sequence coverage of USP3-ZnFBD (1-131).** A total of 60 peptides yielded 93.9% sequence coverage and a redundancy of 5.28. Black rectangles aligned to sequence represent peptides. Several peptides including 52-86 (length 35) and 53-86 (length 34) were manually added after PLGS filtering to obtain coverage from 53-60.

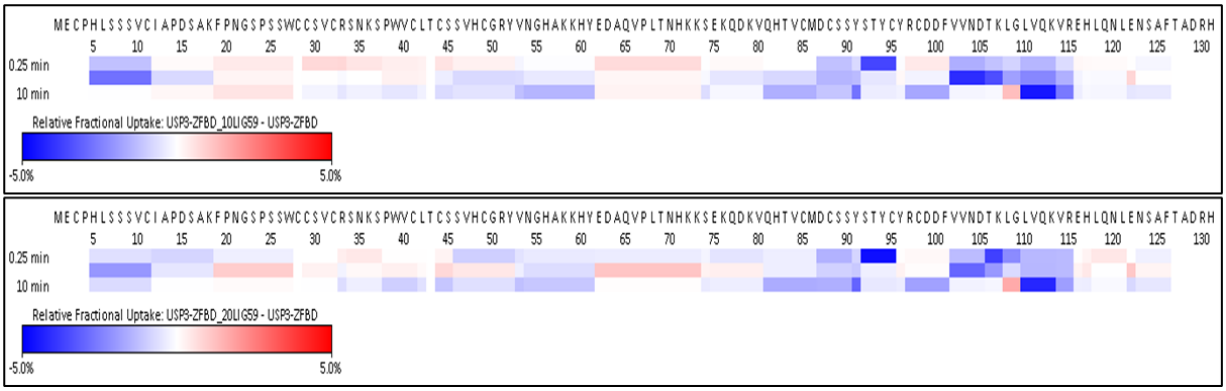
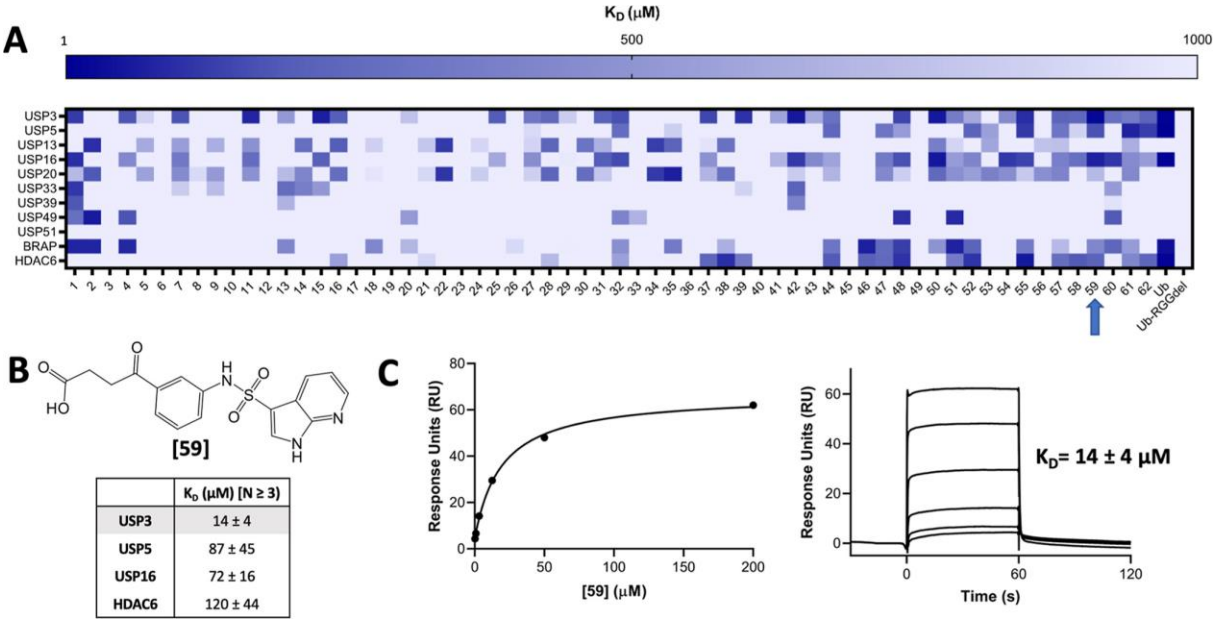


Figure E2. **Differential HDX Heatmaps.** Differential fractional uptake is displayed as a per-residue averaged heatmap aligned to USP3 primary sequence for USP3 Zf-UBD in complex with 59 at 1:10 (top) and 1:20 (bottom) . The highest magnitude D decreases (-5 %) are blue and D increases (+5%) are red. A limitation of this heatmap illustration (DynamX 3.0) is that is does not factor in the associated error for each residue. Please refer to Figure 4 for a peptide-level differential HDX bar-plot with associated error.



	K_D (μM) [N ≥ 3)
USP3	14 ± 4
USP5	87 ± 45
USP16	72 ± 16
HDAC6	120 ± 44

Figure E3. (Originally Figure 2 in publication, work performed by coauthors) **Small molecule screen identifies ZnF-UBD hits.** (A) Heat map showing binding of 64 ligands to 11 ZnF-UBDs by SPR. A 4-fold 6-point dilution series beginning at 200 μM was used for K_D determination with N = 1. (B) Summary of binding data for compound **59** by SPR, $N \geq 3$. (C) Representative SPR binding curve from steady-state fit analysis and sensorgrams for USP3 ZnF-UBD and **59**. A K_D of 14 ± 4 μM was obtained from the average of seven independent measurements. Reproduced with permission from Mann et al. (2023), open access under CC BY-NC-ND 4.0 DEED.[225]

Appendix F: Supplementary information for Chapter 6.

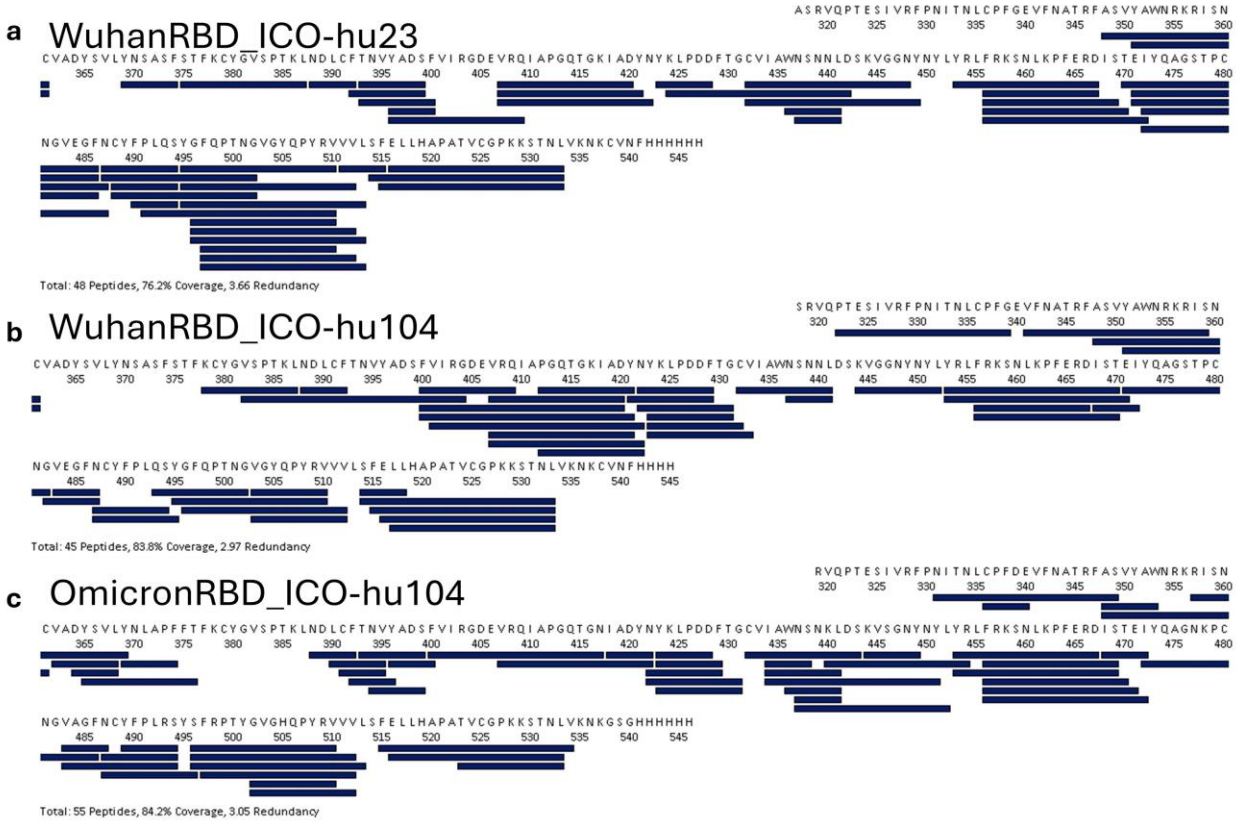


Figure F1. **Sequence coverage maps of RBDs.** (a) Wuhan RBD/ICO-hu23, (b) Wuhan RBD/ICO-hu104, and (c) Omicron RBD/ICO-hu104. These domains had sequence coverage (and redundancy) of 76% (3.66), 84% (2.97), and 84% (3.05), respectively.

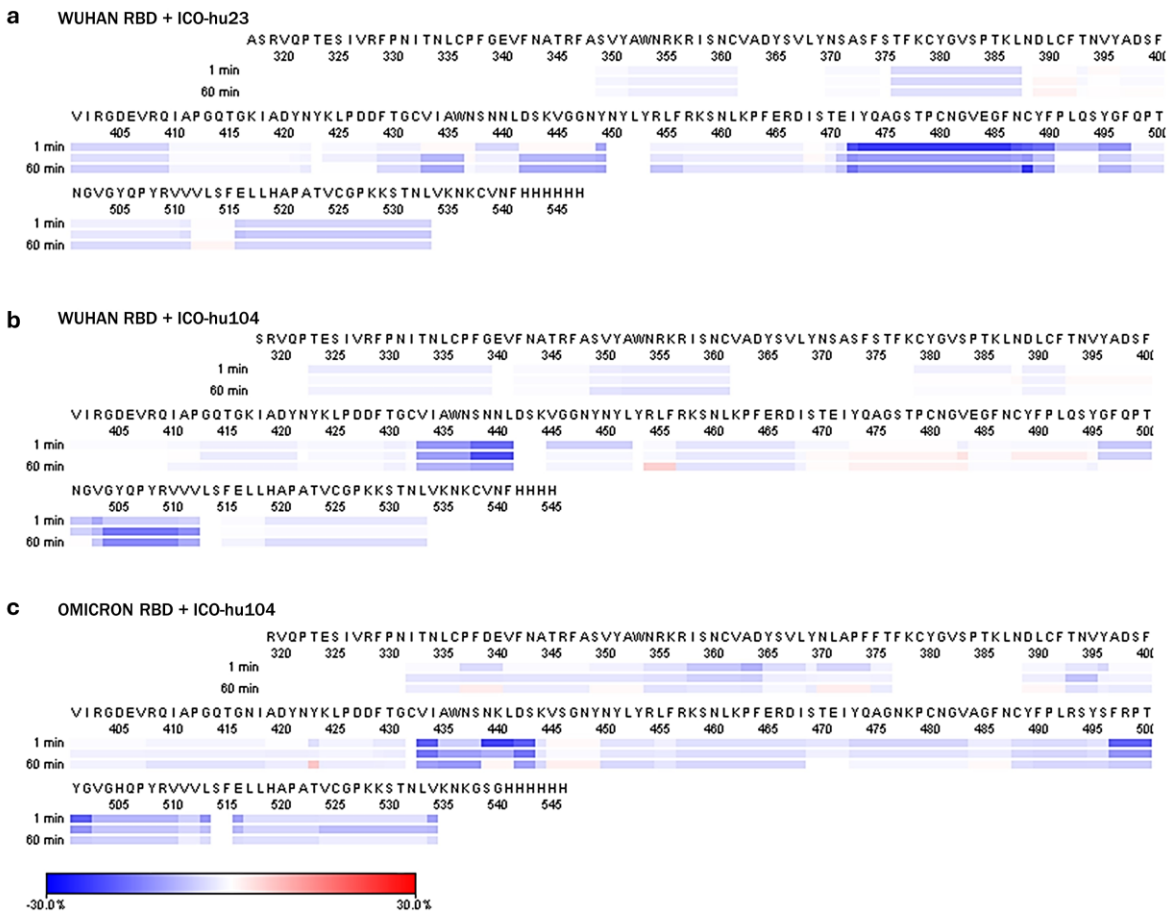


Figure F2. **Heatmaps of ICO-hu23 and ICO-hu104 impacts on RBDs.** The ΔHDX-MS heatmaps illustrate Wuhan RBD in the presence of ICO-hu23 (a), ICO-hu104 (b), and Omicron RBD in the presence of ICO-hu104 (c) with a gradient set to ± 30% difference in relative fractional uptake. Redundant peptides have been averaged and the timepoints (1, 10 , 60 mins) are stacked.

Figure F3. **Sequence coverage maps of Full-length Spike.** (a) Wuhan/ICO-hu23, (b) Wuhan/ICO-hu104, (c) Omicron/ICO-hu104, and (d) Delta/ICO-hu104. These proteins had sequence coverage (and redundancy) of 64% (2.74), 68% (2.92), 62% (2.07), and 65% (2.92), respectively.

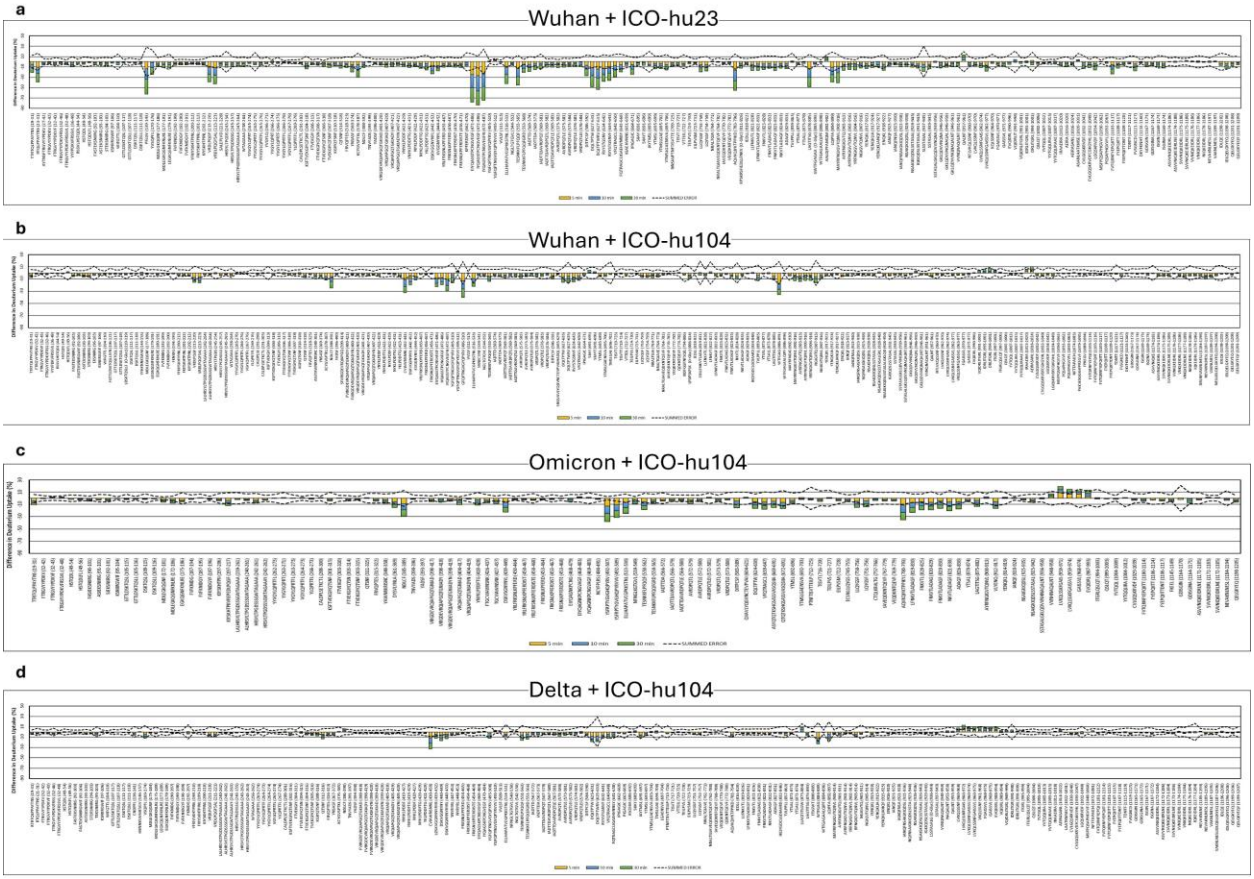


Figure F4. Δ**HDX-MS of Full-Length Spike in the presence of ICO-hu23 and ICO-hu104.** (a) Wuhan/ICO-hu23, (b) Wuhan/ICO-hu104, (c) Omicron/ICO-hu104, and (d) Delta/ICO-hu104. HDX-MS was conducted for 5 (yellow), 10 (blue), and 30 mins (green). To be considered statistically significant in **Fig. 6.4**, the cumulative signal must have exceeded three times the summed propagated error (dashed line).

Appendix G: Supplementary information for Chapter 7.

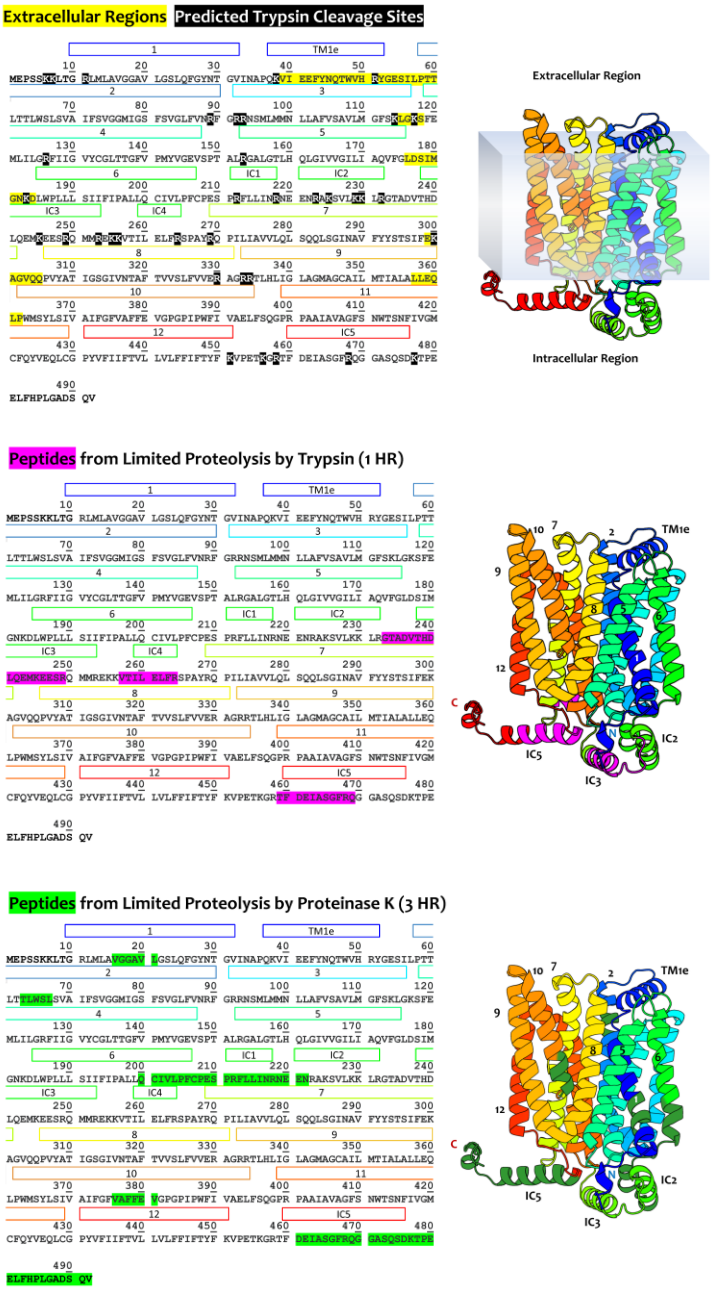


Figure G1. **Limited Proteolysis of POI GLUT1 embedded in VLP.** The GLUT1 sequence is colour-coded to show the extracellular regions in yellow, predicted trypsin digestion sites in black, the peptides generated after 1 hour of limited proteolysis with trypsin in pink, and in green, the peptides generated after 3 hours of proteolysis with Proteinase K. On the right, the GLUT1 transmembrane helices are labeled 1-12 and color coded according to the rectangles aligned to the sequence (AlphaFold structure AF-P11166-F1).[265,266]

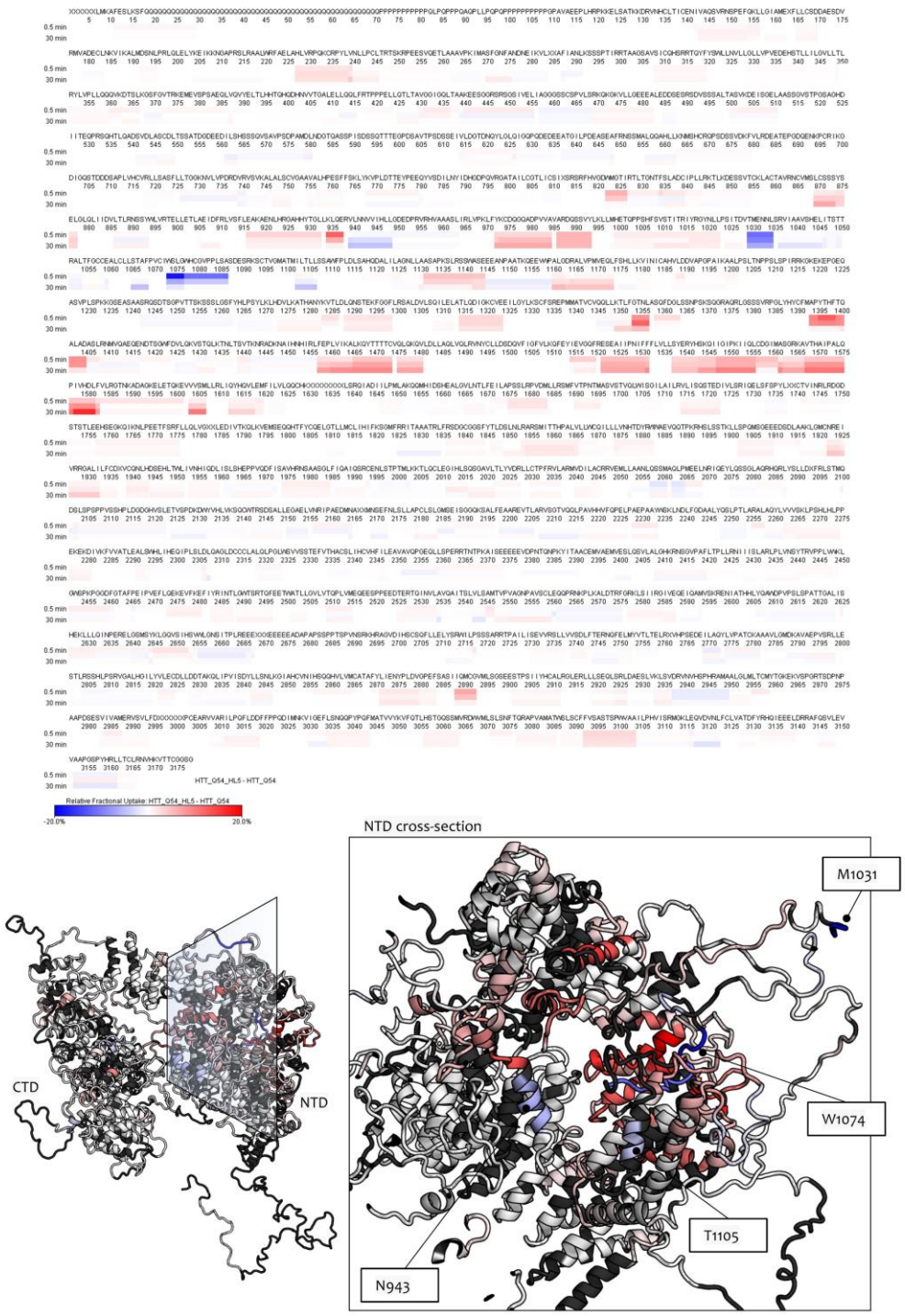


Figure G2. Δ**HDX-MS of macrocyclic peptide HL5 targeting HTT with a polyQ expansion of 54 repeats (HTT_Q54).** Labeling at 30 sec, 5 min, and 30 min. Yield of 451 peptides, 73.3% sequence coverage, and 2.10 redundancy. Summed differences were plotted on PDB 6X9O. HTTQ4 and macrocyclic peptides were prepared by the labs of Dr. Rachel Harding (Structural Genomics Consortium) and Dr. Hiroaki Suga (University of Tokyo).