Probing protein structure and dynamics through native MS and H/D exchange strategies

Shane A. Chandler

University College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy
in Physical and Theoretical Chemistry

Hilary 2019

Abstract

This thesis details the development and application of mass spectrometry (MS) labelling strategies for use in structural biology, and, specifically, how they offer insight into the quaternary structure and dynamics of protein complexes.

Chapter 3 first demonstrates the utility of isotopic H/D exchange (HDX) strategies for probing the in solution dynamics and assembly of two protein systems. Firstly, I use solution HDX to probe the light-induced structural change of the FAD-binding protein cryptochrome. Through quantitative analysis, I discover a region at the C-terminus undergoes a long-lived conformational change with high specificity to blue light, suggesting a potential role in the signalling cascade of circadian processes. Secondly, I use HDX-MS to probe the interface of a homo-dimeric ancestor of extant haemoglobin. Differential HDX-MS revealed protected peptide regions consistent with the α/β *packing* interface in the modern-day hetero-tetramer.

Studying proteins in the gas-phase by using native MS has the distinct advantage of high resolution separations, manipulations and accurate mass determination, afforded by the exquisite control possible in vacuum. Chapters 4, 5 and 6 describe the implementation of performing gas-phase HDX (gpHDX) inside the spectrometer and how the technique may be used to interrogate protein ion structure with high specificity. I first demonstrate how gpHDX can be used to probe native proteins and show high sensitivity to the surface chemistry of exposed residues. Secondly, I use gpHDX to aid our understanding of how protein complexes asymmetrically dissociate through collisional activation and charge precipitated unfolding. Importantly, I show that the deuterium uptake of a dissociated subunit can be used as a proxy for the intact precursor. Finally I use gpHDX coupled to ion mobility (IM) to reveal how the conformational landscape of an activated protein changes at a global and local level.

Probing protein structure and dynamics through native MS and H/D exchange strategies



Shane A. Chandler

University College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy
in Physical and Theoretical Chemistry

Hilary 2019

Acknowledgements

I am forever grateful to the many people have made the last four years such an enjoyable and stimulating experience of my life.

I am indebted to Justin Benesch for his endless support throughout this DPhil. He has always encouraged me to follow my interests and guided me through the intellectual challenges they presented. I have always felt inspired by the passion he holds for the technique and the research we do; observing his excitement has repeatedly rekindled my motivation throughout. I owe the way I can discuss, talk, and think about my work almost entirely to his attentive supervision. I also thank Mike Morris for all of the opportunities he has provided, both in Oxford and Wilmslow. I am deeply grateful to them both for enabling what has become a deep-rooted interest and source of enjoyment.

I thank Dale Cooper-Shepherd, though our overlap was brief, he provided the most enjoyable start to a DPhil a budding mass spectrometrist could ask for. To Erik Marklund and Georg Hochberg whose company in the office was a constant source of enjoyment and discussions, some of which have continued to this day. To Tim Allison and Max Hantke for their help and patience in teaching me the art of code. To Joseph Gault, to whom I could always turn to for advice, whether over a pint or not. I thank Shahid Mehmood for all of his advice surrounding HDX. To George Wright for the summer of hard work he spent in our lab. The Robinson, Benesch and Baldwin labs have made for an incredibly enjoyable environment full of both laughter and discussions, and for that I thank, Reid Alderson, Iva Pritišanac, Olya Tkachenko, Heidi Gastall, Weston Struwe, Jon Hopper, Idi liko, Jani Redi-bolla, Stephen Ambrose, Matteo Degiacomi, Mark Agasid, Denis Schutin, Anna van der Zaalm, and everyone else I have had the pleasure of working with. I thank Helena Hernandez, Andy Dolan, Vicky Cullen, and, in particular, Todd Mize for all of their help keeping the lab well organised and aiding our work. I additionally thank Carol Robinson for providing such a pleasurable research group family to belong, and whose vision and ambition has been a source of inspiration.

A special thank you to Miranda Collier, having started on the very same day, we shared many of the ups and downs of lab life together. I have always found her attitude and approach to science hugely impressive, something i'm not she

will allow me to say. To Gogs Karunanithy and Joe Donlan for sustained moral support during our writing efforts. I thank the other half of the gpHDX special interest group, Ulrik Mistarz and Kasper Rand, for all of the fun and advice gained from their company. I thank all of those at Waters, Kevin Giles, Jeff Brown, and Jakub Ujma and Mike Morris, for the advice over the years and exposure to such interesting research.

I am indebted to all of those I have collaborated with over the years for the exposure to exciting projects that have helped broaden my scientific development. For that I thank, Katie Buckton, Stuart Mackenzie, Peter Hore, Chris Timmel, and the Thornton, Conway, Mchaourab, and Smith laboratories. I owe Reid Alderson, Mark Agasid, Peter Hore and Joseph Gault for their diligent comments on this work.

I owe a big thank you to those friends outside of the lab, Giulio Mazzotta, James Barwick-Silk and Serjoscha Evers, that have been a source of relief and enjoyment in Oxford.

I am deeply grateful for the time spent at GSK with Bob Boughtflower and Tim Underwood, they first showed me what fun and creativity can be had with independent research, something that has helped see me to this point today.

My parents have had more of a part in this achievement than they may realise. They have always ensured my education has been nurtured and have given all they could to support me at every possible turn. I am always hugely appreciative of my weekly conversations with my Grandma often providing a welcome distraction from troubling thoughts of work. I thank all of my family for everything they have done over the years, it has gone far from unnoticed.

Lastly, I would like to thank Jo Williams for her unfailing support and encouragement when I have needed her most. I am very lucky to have had her at my side throughout this challenging and rewarding journey.

Statement of Originality

This thesis describes research conducted between October 2014 and January 2019. The results reported are all common to a research theme with each chapter focussing on a discrete project or discussion. All data presented was acquired and analysed by myself unless explicitly stated. Figures can be assumed to be of my own making, unless otherwise stated in the caption. Each results chapter has a corresponding Acknowledgements section further detailing any specific contributions. Chapter 5 was of a highly collaborative nature, and so additional efforts were made to delineate contributions. This work has not been submitted for a degree, diploma, or any other academic certification at any other institution, in whole or in part.

Shane A. Chandler
February 2019

Abstract

This thesis details the development and application of mass spectrometry (MS) labelling strategies for use in structural biology, and, specifically, how they offer insight into the quaternary structure and dynamics of protein complexes.

Chapter 3 first demonstrates the utility of isotopic H/D exchange (HDX) strategies for probing the in solution dynamics and assembly of two protein systems. Firstly, I use solution HDX to probe the light-induced structural change of the FAD-binding protein cryptochrome. Through quantitative analysis, I discover a region at the C-terminus undergoes a long-lived conformational change with high specificity to blue light, suggesting a potential role in the signalling cascade of circadian processes. Secondly, I use HDX-MS to probe the interface of a homo-dimeric ancestor of extant haemoglobin. Differential HDX-MS revealed protected peptide regions consistent with the α/β *packing* interface in the modern-day hetero-tetramer.

Studying proteins in the gas-phase by using native MS has the distinct advantage of high resolution separations, manipulations and accurate mass determination, afforded by the exquisite control possible in vacuum. Chapters 4, 5 and 6 describe the implementation of performing gas-phase HDX (gpHDX) inside the spectrometer and how the technique may be used to interrogate protein ion structure with high specificity. I first demonstrate how gpHDX can be used to probe native proteins and show high sensitivity to the surface chemistry of exposed residues. Secondly, I use gpHDX to aid our understanding of how protein complexes asymmetrically dissociate through collisional activation and charge precipitated unfolding. Importantly, I show that the deuterium uptake of a dissociated subunit can be used as a proxy for the intact precursor. Finally I use gpHDX coupled to ion mobility (IM) to reveal how the conformational landscape of an activated protein changes at a global and local level.

Contents

List of Figures **xiii**

List of Abbreviations **xvii**

1 Introduction **1**

1.1 Mass spectrometry for structural biology 1

1.2 Native MS exemplified 3

1.3 Mass spectrometry beyond the native state 6

1.3.1 Coulombic exploration of the free energy landscape 6

1.3.2 Structural interrogation of native and activated states 10

1.4 Concluding remarks 13

References **15**

2 Native mass spectrometry background **25**

2.1 Mass spectrometry instrumentation 25

2.1.1 Electrospray ionisation 27

2.1.2 Transmission of high molecular weight ions 32

2.1.3 Quadrupole mass analysers 32

2.1.4 Time-of-flight mass analyser 34

2.1.5 Orbitrap analysers 38

2.1.6 Ion mobility spectrometry 39

2.2 Gas-phase activation 45

2.2.1 Dissociation of protein complexes 46

2.3 Solution hydrogen/deuterium exchange MS 48

2.3.1 HDX kinetics 51

2.4 Gas-phase hydrogen/deuterium exchange 54

References **59**

3 Protein structure and solution dynamics through HDX-MS **73**

3.1 Abstract . . . 73

3.2 Introduction . . . 74

3.3 Part I : Structural dynamics of a light-induced conformational change 75

3.3.1 Biological context . . . 75

3.3.2 Results and Discussion . . . 78

3.3.3 Conclusion . . . 91

3.4 Part II : Protein-protein interface determination . . . 94

3.4.1 Biological context . . . 94

3.4.2 Results and Discussion . . . 96

3.4.3 Conclusion . . . 109

3.5 Concluding remarks . . . 109

3.6 Methods . . . 110

3.6.1 Materials . . . 110

3.6.2 HDX-MS . . . 110

3.6.3 Back-exchange controls of the system . . . 112

3.6.4 Native MS . . . 112

3.6.5 Data analysis . . . 113

3.7 Acknowledgments . . . 114

References **115**

4 Implementation and application of gas-phase HDX for probing protein surfaces **119**

4.1 Abstract . . . 119

4.2 Introduction . . . 120

4.3 Results and Discussion . . . 121

4.3.1 Implementation and modifications . . . 121

4.3.2 Application to biomolecules . . . 123

4.3.3 Effect of ion manipulation on uptake . . . 126

4.3.4 Considerations for ion mobility MS . . . 131
4.3.5 Gas-phase HDX control experiments . . . 133
4.3.6 Gas-phase labelling is sensitive to protein surface chemistry . . . 135
4.4 Conclusion . . . 142
4.5 Methods . . . 143
4.6 Acknowledgements . . . 147

References **149**

5 Probing the dissociation of protein complexes by means of gas-phase H/D exchange mass spectrometry **151**
5.1 Abstract . . . 151
5.2 Introduction . . . 152
5.3 Results and Discussion . . . 153
5.3.1 Setup for gas-phase HDX-MS and CID . . . 153
5.3.2 Gas-phase HDX-MS and CID of protein-protein dimers . . . 154
5.3.3 Gas-Phase HDX-MS and CID to inform on protein-protein multimers . . . 156
5.3.4 Large Protein-Protein Multimers . . . 158
5.3.5 Deuterium enrichment of ejected subunits during CID of large complexes . . . 162
5.4 Conclusion . . . 166
5.5 Methods . . . 170
5.6 Acknowledgements . . . 174

References **175**

6 Gas-phase HDX: collision induced unfolding **179**
6.1 Abstract . . . 179
6.2 Introduction . . . 180
6.3 Results and Discussion . . . 181
6.3.1 Gas-phase HDX reveals structural changes during activation 181

6.3.2 gpHDX is sensitive to surface side chain structure not observed in IM . . . 187
6.3.3 Validation of CIU-gpHDX behaviour . . . 190
6.3.4 Absolute CCS determination for the activated states of GFP 194
6.3.5 Gradual side-chain rearrangements during activation . . . 197
6.4 Conclusion . . . 199
6.5 Methods . . . 200
6.6 Acknowledgments . . . 202

References **203**

7 Concluding remarks **207**

List of Figures

1.1 **Mass determination of protein complexes** 4
1.2 **Dissecting protein complexes** 5
1.3 **Protein unfolding landscape** 8
1.4 **Gas-phase activation and labelling** 9
1.5 **Gas-phase protein imaging** 12

2.1 **Schematic of Waters synapt G1 ion mobility Q-ToF mass spectrometer** 26
2.2 **Electrospray ionisation of proteins** 29
2.3 **Average charge state of globular soluble proteins** 30
2.4 **Quadrupole mass analyser** 33
2.5 **Ion mobility spectrometry of protein complexes** 39
2.6 **Stacked ring travelling wave ion guide** 42
2.7 **Collision cross section calculation methods** 43
2.8 **Collision induced dissociation of protein complexes** 47
2.9 **Theoretical rate of intrinsic H/D exchange** 50
2.10 **Solution HDX-MS workflow** 52
2.11 **HDX-MS kinetics** 53
2.12 **Gas-phase HDX reactivity** 55

3.1 **Photo-activation of *Drosphila melanogaster* Cryptochrome** 77
3.2 **Dark state dynamics of *Dm*-Cryptochrome** 80
3.3 **Light-induced conformational change in *Dm*-Cryptochrome** 83
3.4 **Conformational specificity towards blue light** 86
3.5 **EX1 fitting of C-terminal peptide spectra** 87

3.6 **Kinetics of the light induced conformational change** 89
3.7 **Oligomeric origins of extant haemoglobin** 95
3.8 **Native mass spectra of ancestral haemoglobin and mutants** 98
3.9 **Pepsin digest of ancestral haemoglobin** 100
3.10 **Uptake spectra for dimeric Anc haemoglobin and monomeric mutants** 101
3.11 **Differential HDX-MS reveals interfacial residues** 103
3.12 **Titration of ancestral α/β haemoglobin supports interfacial regions determined by HDX-MS** 105
3.13 **HDX-MS data suggest residues buried at the ancestral haemoglobin dimer interface** 107

4.1 **Gas-phase HDX modifications to an ion mobility mass spectrometer** 122
4.2 **Gas-phase HDX of biomolecules** 124
4.3 **Gas-phase HDX of sugars.** 125
4.4 **Manipulation of peptide deuterium incorporation.** 127
4.5 **Travelling wave parameter control over protein labelling** . . 128
4.6 **Travelling wave effects on the gpHDX labelling efficiency in a trapped ion mode.** 130
4.7 **Travelling wave impact on transmission and subsequent mobility** 132
4.8 **Undeuterated control experiments** 133
4.9 **Effect of ion current on labelling** 134
4.10 **Surface electrostatics of model GFP variants** 135
4.11 **Native MS spectra of model GFP variants** 137
4.12 **Gas-phase HDX discriminates between surface chemistries** 138
4.13 **Impact of negative ion generation on GFP mutant exchange ability** 141

5.1 **Gas-phase HDX-MS for studying protein-protein complexes** 153
5.2 **Gas-phase HDX and subsequent CID of a dimeric complex** 155

5.3 **Deuterium distribution of after symmetric and asymmetric charge partitioning.** . . . 156
5.4 **Dissociation of tetrameric transthyretin studied by gas-phase HDX-MS** . . . 157
5.5 **Gas-phase HDX and subsequent CID of sHSP oligomers** . . . 159
5.6 **Reproducibility of gas-phase HDX-CID on sHSP oligomers** 160
5.7 **Collision energy dependency on m/z for gas-phase labelled sHSP oligomers** . . . 161
5.8 **Migration of deuterium upon dissociation of protein complexes.** . . . 163
5.9 **Deuterium levels partition symmetrically during CID of native protein complexes** . . . 169

6.1 **Gas-phase HDX modifications to allow conformer specific labelling** . . . 182
6.2 **Obtaining CIU-gpHDX native mass spectra** . . . 183
6.3 **Gas-phase HDX and ion mobility report on conformer structural changes** . . . 184
6.4 **Collision induced unfolding and labelling of GFP 9+ ions** . 185
6.5 **Step and gradient extractions from IM and gpHDX dimensions** . . . 189
6.6 **Reproducibility of CIU-gpHDX** . . . 190
6.7 **CIU-gpHDX unfolding plots resulting from an inter-laboratory study** . . . 192
6.8 **Trap activation of GFP variants with simultaneous and subsequent labelling** . . . 193
6.9 **Absolute CCS determination of unfolded intermediates** . . 196
6.10 **Protein surface inversion during gas-phase activation** 198

List of Abbreviations

APR . . . . . Ancestral protein reconstruction

BB . . . . . Backbone

CCS . . . . . Collision cross section

CID . . . . . Collision induced dissociation

CIU . . . . . Collision induced unfolding

CRM . . . . . Charged residue model

CRY . . . . . Cryptochrome

CTT . . . . . C-terminal tail

DC . . . . . Direct current

Dm . . . . . *Drosophila melanogastor*

DT-IM . . . . Drift tube ion mobility

ECD . . . . . Electron capture dissociation

EHSS . . . . . Elastic hard sphere scattering

EM . . . . . Electron microscopy

(n)ESI . . . . . (nano)Electrospray ionisation

ETD . . . . . Electron transfer dissociation

FAD . . . . . Flavin adenine dinucleotide

FWHM . . . . Full width half maximum

FRET . . . . . Förster resonance energy transfer

GFP . . . . . Green fluorescent protein

GPA . . . . . Gas-phase proton affinity

GPB . . . . . Gas-phase basicity

gpHDX . . . . Gas-phase hydrogen/deuterium exchange

(s)HSP . . . . (small)heat-shock protein

HDX . . . . . Hydrogen/deuterium exchange

IEM . . . . . Ion ejection model

IM . . . . . Ion mobility

IRPD . . . . . Infrared photo dissociation

JET . . . . . Jetlag

LC . . . . . Liquid chromatography

LED . . . . . Light emitting diode

MD . . . . . Molecular dynamics

NMR . . . . . Nuclear magnetic resonance

oa . . . . . Orthogonal acceleration

PA . . . . . Projection approximation

PDB . . . . . Protein data bank

PER . . . . . PERIOD

PHR . . . . . Photolyase homology region

RF . . . . . Radio frequency

RP . . . . . Reverse phase

SA . . . . . Surface accessible

SaBRe . . . . . Salt-bridge rearrangement

SB . . . . . Salt-bridge

SD . . . . . Standard deviation

SEC . . . . . Size exclusion chromatography

SID . . . . . Surface induced dissociation

TIM . . . . . TIMELESS

TM . . . . . Trajectory method

ToF . . . . . Time-of-flight

TTR . . . . . Transthyretin

TW . . . . . Travelling wave

TIWG . . . . . Travelling wave ion guide

UVPD . . . . . Ultra-violet photo dissociation

WT . . . . . Wild type

1 Introduction

1.1 Mass spectrometry for structural biology

The role of mass spectrometry (MS) in structural biology is now commonplace for assessing all levels of protein structure, from determining residue-level composition of proteoforms, to the characterisation of large intact heterogeneous protein complexes.[1,2] Such prevalence hinges on the ability to introduce intact biomolecules into the gas-phase, from where exquisite control over an ion's motion can lead to an array of manipulations,[3] and, ultimately, accurate mass determination. The diverse range of MS methodologies for probing the structure of biomolecules can broadly be grouped into two approaches: those in which non-covalent protein interactions are maintained on transferral to vacuum and those where peptidic fragments are generated in solution and introduced into the spectrometer. The focus of this thesis is to develop and utilise MS strategies from both of these approaches to inform on the surface topology, higher order structure, and dynamics whilst in, and away from, a protein's native state.

Peptide-centric approaches, often referred to as "bottom-up", use traditional biochemical methodologies to modify or manipulate the protein complex in solution; an enzymatic digestion generates peptides that report on their intact environment.

Mass spectrometry can then be used to accurately determine the identity of each peptide and quantitatively characterise the species present and any incorporated solution modifications.[4,5] Such strategies include chemical cross-linking,[6] limited proteolysis,[7]and labelling strategies. A multitude of different chemical labels[8] can be used to assess the surface of an intact protein assembly, ranging from oxidative or carbene footprinting[9,10] to the focus of this thesis, hydrogen/deuterium exchange (HDX).[11] HDX is a highly sensitive probe of a protein's structural and conformational dynamics through the isotopic exchange of labile, solvent exposed sites associated with the backbone of each residue (§2.3). Mass spectrometry may then be used to characterise the heavy isotope incorporation for each of the peptidic fragments generated and give a read-out of the ensemble's solvent accessibility and local interaction network. In Chapter 3, we utilise HDX-MS to report on the conformational dynamics of an inducible signalling state and determine the likely binding interface of an ancestral dimeric complex.

Native MS allows for a protein-centric approach in which intact proteins, and other biomolecules, are studied in the vacuum of a mass spectrometer, having preserved their non-covalent interactions upon transferral from solution.[12,13] A variety of experimental data and simulations have shown that the native state of proteins and their complexes can persist on the millisecond time-scale of MS experiments.[14–16] That many structural elements remain, with some electrostatic interactions becoming even stronger,[17,18] allows one to capture quaternary structure of protein assemblies.[19,20]

1.2 Native MS exemplified

At its core, native MS can accurately report on the mass of the species present in solution with high sensitivity, as low as nano - micro molar concentrations.[15,21] The following section demonstrates some of the wide ranging applications a mass measurement can provide using examples from my own work. Figure 1.1A shows the native mass spectrum for a monomeric metal binding protein, NDM-1, where a narrow distribution of low charge state ions are observed (§2.1.1). The binding stoichiometry and relative abundance of each state can be accurately determined for a range of Zn^{2+} concentrations. Native MS is uniquely placed in its ability to distinguish each association event and determine the relative affinity. Further evidence that the gas-phase measurement provides an accurate representation of solution dynamics comes from the ability to accurately quantify dissociation constants, K_d.[22,23] Such a strategy is shown for the homo-dimeric KefB protein binding to a small molecule inhibitor, Figure 1.1B. The relative abundance of each species was determined over the course of a ligand titration, and in doing so, extracted data can be fit to an appropriate binding model and the step-wise dissociation constants ascertained.

Native MS has real strength when posed with increased complexity through sample heterogeneity, polydisperse assemblies and dynamic systems. The electrospray ionisation process allows one to quantitatively capture a snap-shot of the solution distributions inside the spectrometer. Previous work in our group has shown native MS to be particularly well suited to probing the large dynamic assemblies formed by the small heat-shock proteins (HSP).[24] Stoichiometric determination of engineered HSP 16.5 shows how the structural assembly can be tuned by small differences in the primary structure of the individual subunits.[25] The highly quantitative nature of MS can precisely determine the species identity across a wide ranging distribution and inform on the relative abundances, Fig 1.1C.

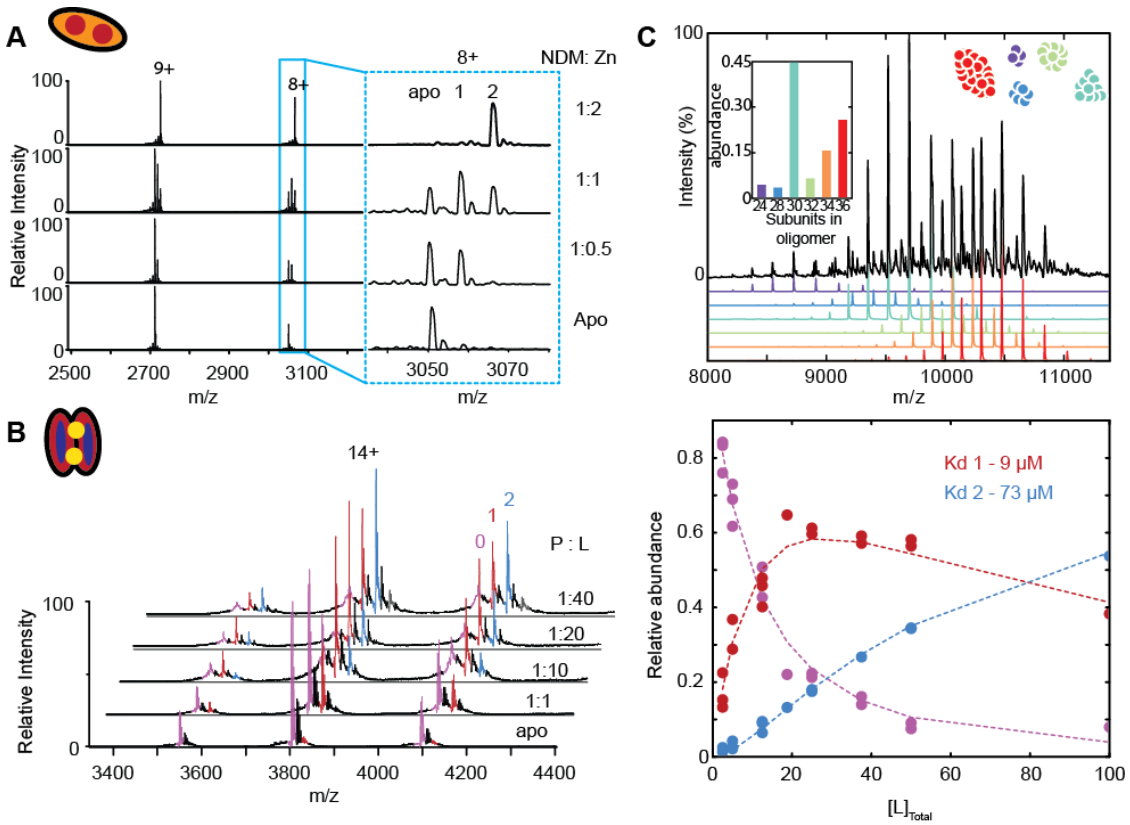


Figure 1.1. Mass determination of protein complexes **A** Native mass spectra of the 24.4 kDa monomeric zinc binding New Delhi metallo-beta-lactamase 1 (NDM-1). The apo protein (bottom) was electrosprayed at a concentration of 2.5 µM under non-activating conditions. High affinity zinc binding was observed after titration of $Zn(OAc)_2$. Two specific binding sites were observed with no charge state variation. **B** The dimeric protein KefB was electrosprayed at 1.4 µM to give a narrow charge state distribution, exclusively found with two endogenous AMP molecules bound (See my previous work[26]). The derivitised glutathione ligand, ESG (432.5 Da), was titrated to give the indicated ratios to reveal two specific binding sites. Peak areas were extracted to determine the abundance of each state across all charge states. The binding curves are shown (right) for the apo protein (purple), the singly bound (red), and doubly bound (blue). Fits suggest negative cooperativity and the dissociation constant, K_d, is found to be ~9 µM and ~73 µM for the first and second binding events respectively. Data was fit using Unidec[27] and validated accounting for non-specific binding interactions.[28] **C** Native mass spectrum of engineered HSP 16.5 (Var-3) revealed the co-population of multiple oligomeric states. The raw spectrum is shown in black with deconvoluted distributions shown in colour. Only even number stoichiometries were observed ranging from a 24mer to a 36mer (monomeric mass = 17.5 kDa). The relative abundances of the different states (inset) were extracted from the signal intensity using Unidec.[27] Panels **A** & **B** experiments were performed on an orbitrap (Thermo) platform[29] modified for high mass transmission.[30] Panel **C** was performed on a UHMR-orbitrap (Thermo) with in-source trapping activation.[31] NDM-1 protein was provided by Jürgen Brem (Uni. of Oxford). KefB protein and inhibitor was provided by Anthony Chan (Uni. of Oxford). Engineered HSP 16.5 samples were provided by Sanjay Mishra (Uni. of VanderBilt). Panel **C** was previously published in Mishra *et al.*[25]

The control available inside the spectrometer allows for highly specific ion manipulations in the form of mass-to-charge (m/z) or size selections, charge

state manipulations, and, importantly, activation. Having preserved many of the non-covalent interactions found in solution, it can be advantageous to disrupt these in the gas-phase and cause ligand or subunit loss. Where a complex mass cannot be determined unambiguously, tandem MS, isolation according to m/z followed by dissociative activation, can eject subunits allowing for more accurate characterisation(§2.2.1). Figure 1.2A shows how a tandem MS workflow can be used to verify the identity of a bound ligand.[32] Tandem MS, using slow activation methods, has the additional advantage of decongesting spectra, causing peaks to spread over a greater m/z range (Fig. 1.2B).

Detailed mechanistic processes and technical instrument background of the above discussion can be found in Chapter 2.

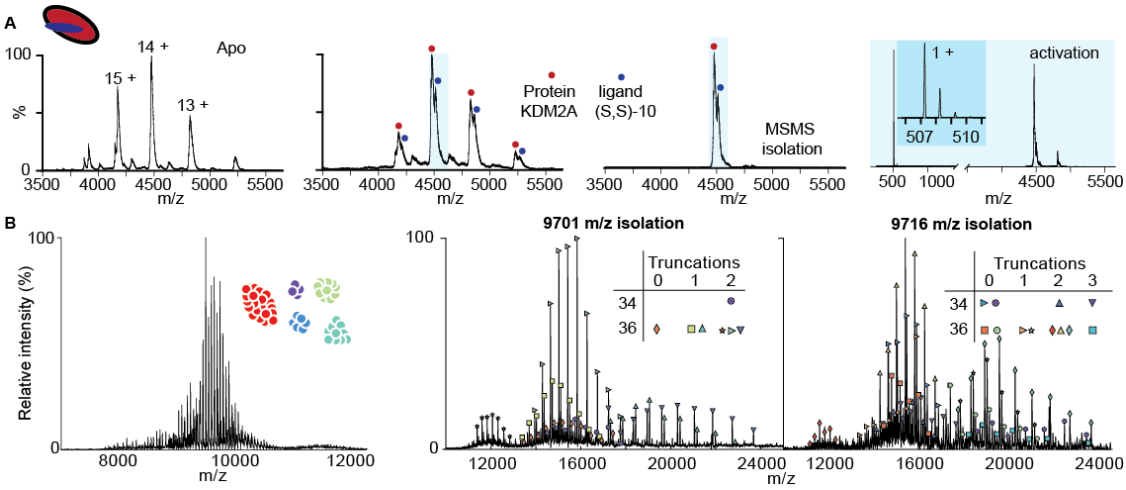


Figure 1.2. Dissecting protein complexes A Validation of 1:1 ligand binding of the inhibitor (S,S)-6 to the monomeric protein KDM2A through tandem MS. *Apo* mass spectrum at 2.5 µM (left) obtained on a Waters Synapt G1 Q-ToF instrument(§2.1). Addition of 12.5 µM inhibitor shows ligand binding (blue) in a 1:1 stoichiometry. The 14+ charge state of the *apo/holo* peak was isolated in the quadrupole and collisionally activated (right). The dissociated (S,S)-6 ligand is shown as a singly charged ion in the inset. The spectrum intensity has been magnified 1.5-fold above 3500 m/z. **B** Tandem MS of polydisperse engineered HSP 16.5 (Var-6) oligomers. The low activation native mass spectrum results in a complex spectrum with overlapping charge state series. To identify the underlying species tandem MS was performed to give singly, doubly, and triply stripped oligomers separated to a greater extent in m/z(§2.2.1). Representative spectra are shown (middle & right) for the indicated m/z isolation with a transmission window of ± 10 m/z. The oligomeric identity was determined and revealed the incorporation of truncated monomeric species. The oligomers present for each quadrupole selection is given including the number of truncated subunits. Spectra were acquired using a UHMR-orbitrap (Thermo) with in-source trapping activation.[31] KDM2A and inhibitor were provided by Philip Gerken (Uni. of Oxford). HSP 16.5 (Var-6) was provided by Sanjay Mishra (Uni. of VanderBilt). Adapted from Gerken *et al.* and Mishra *et al.*[25,32]

1.3 Mass spectrometry beyond the native state

The time-scale of "native-like" gas-phase structures provides ample opportunity for interrogation of not just native structure, but also the free-energy landscape accessible to the protein.[33,34] This can be viewed in terms of a canonical protein (un)folding funnel, where protein ions can populate states differing in their atomic coordinates and energy, Fig. 1.3. While the native states of proteins are relatively accessible to structural biologists and biophysicists, probing states populated only to low levels at equilibrium, or fleetingly during assembly and folding, remains extremely challenging.[35] In this introduction I discuss how and why MS-based experiments can contribute to this endeavour.

I first consider how the protein folding funnel can be explored within the mass spectrometer, discussing the role played by protein charge state in populating different minima. I evidence how these represent starting points for exploring higher levels of the protein folding funnel by activating ions within the mass spectrometer. In recent years a diverse range of gas-phase experiments have helped contribute to our understanding of the level at which "solution-like" interactions and structures persist and where the differences lie.[14] In this thesis I describe our use of gas-phase labelling, Chapters 4-6, as a means of understanding gas-phase structure to complement ion mobility, and emerging spectroscopic and imaging methods that are being integrated with native MS. These studies provide additional levels of detail regarding the structure of proteins in vacuum, including during excursions from the native state.

1.3.1 Coulombic exploration of the free energy landscape

Protein ions generated by (nano)electrospray ionisation (ESI) each populate a distribution of charge states. Oftentimes the gross structure, evidenced by collision cross-section (CCS) values obtained from ion mobility (IM) measurements, appears to be lowly affected by this charge state variation for folded states.[36] However, in cases where the range of charge states populated is broad relative to the average, or

when the protein has an intrinsic tendency towards disorder, significant variations in CCS are observed.[37,38] This demonstrates how, as in solution,[39] different charge states of a protein in vacuum correspond to different free-energy minima in the free energy landscape, Fig. 1.3.

To investigate how global protein structure depends on charge state, a number of recent studies have performed IM-MS on protein ions that are selected within the mass spectrometer. Charge manipulation has been performed via ion-ion collisions with a reaction partner of opposite polarity, most commonly performed by abstraction of a proton,[40,41] or transfer of an electron.[42,43] A number of studies have shown that reduction of charge on native conformers of ubiquitin and cytochrome C results in a step-wise compaction.[41–46] In other cases, increases in CCS upon charge reduction are observed.[42,47] This demonstrates the impact of charge on the initial position on the folding funnel, and that other local minima can be occupied by charge modulation. These minima therefore represent alternative starting points for exploration of the conformational landscape.

In order to navigate the protein folding funnel, the ions can be activated within the mass spectrometer (Fig.1.3), typically by energetic collisions with a bath gas. By varying the amount of energy supplied, the types of bonds broken and level of structural information obtained can be controlled. At a low level of gas-phase activation, proteins first undergo structural rearrangements and unfolding due to breakage of intramolecular non-covalent bonds.[48] IM-MS measurements have shown that this occurs via a number of discrete transitions in CCS (Fig. 1.4), diagnostic of the trajectory taken by the protein as it escapes the folding funnel. The structural evolution of a protein depends strongly on its charge state,[43,45,49] and tandem IM-MS experiments reveal complex conformational inter-dependencies,[50] consistent with the view of different minima in the folding funnel being populated and leading to alternative pathways out, Fig. 1.3 .

As a consequence, the transitions observed and the energies at which they occur represent a characteristic signature for the protein charge state being investigated.

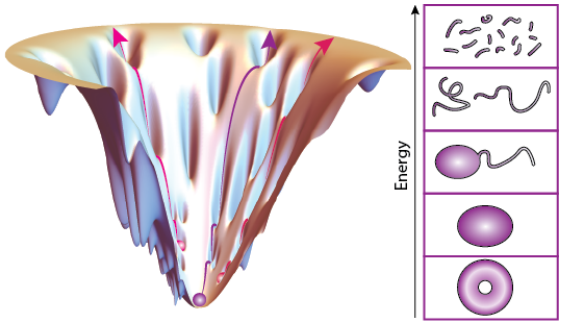


Figure 1.3. Protein unfolding landscape. Proteins exist on a rugged free-energy landscape in vacuum,[33] that can be framed in terms of a protein folding funnel with the native state at the global minimum.[35] Different charge states of the protein can be considered as other minima on this surface, and thereby represent alternative starting points for exploration of the funnel. This can be achieved by gas-phase activation of the protein ions,[48] enabling access to different unfolded states. At high levels of activation, dissociation of non-covalently bound units and even fragmentation of the polypeptide backbone can occur, providing information as to protein composition and identity. Originally published in Chandler & Benesch.[3]

This can be exploited as a fingerprint for comparison, enabling, for instance, distinction between different antibody structures and interactions.[51–54] Unfolding has also been actively pursued as a measure of ligand-stabilisation of protein structure,[55] revealing measures consistent with solution data,[56] as well as reflecting interactions that are strengthened in vacuum.[57] The number of unfolding transitions appears to reflect the number of independent structural domains seen in solution.[58] However, assignment of the transitions observed to individual regions of the protein is difficult. One way to address this is to use specific small-molecule probes, and correlate their release upon activation with the transitions in CCS.[59] Potentially a more general strategy will come from performing electron-mediated fragmentation of the protein during its unfolding pathway.[60] Increasing activation further results in the dissociation of non-covalently attached units, protein subunits or other ligands. In general, dissociation of multimeric assemblies appears to proceed via the expulsion of single, highly charged subunits.[48] The identity of the dissociation products can yield compositional information on the complex, while the asymmetric partitioning of charge between products can be exploited to mine spectra not resolvable by MS alone.[61–63] However, fast activation, in particular by collision with a surface, can lead to information regarding the assembly and architecture of the complex.[64,65] This can be rationalised by the activation allowing access to the higher reaches of the

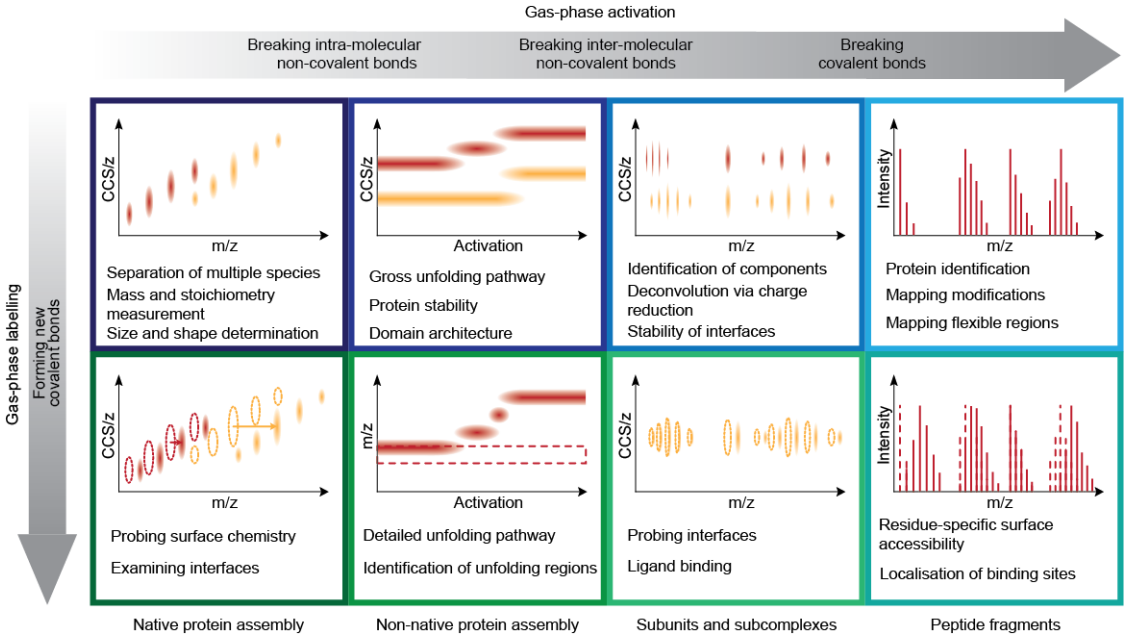


Figure 1.4. Gas-phase activation and labelling. Gas-phase activation and labelling are two orthogonal and complementary dimensions for interrogating the conformational states of proteins and trajectories in their folding funnel (Fig. 1.3). Activation provides structural information via the breaking of intra- and inter-molecular bonds. This leads to unfolding, dissociation of non-covalently bound ligands or subunits, and fragmentation of covalent bonds (upper row, left to right, and Fig. 1.3). These processes can provide information on the identity, composition and structure of native and non-native states. Labelling instead relies on the formation of new bonds at surface-exposed sites. MS is sensitive to these reactions due to a shift in mass from the unlabelled species (dotted lines) upon introduction of the labelling agent. Labelling efficiency is sensitive to the nature of the species present and so differences can provide another means of informing on structure, across the activation spectrum (upper row, left to right). Some of the insights that can be obtained from gas phase activation, in the absence of presence of labelling agents, are inset, together with idealised schematics of anticipated IM-MS data. Adapted from Chandler & Benesch.[3]

protein energy landscape without transitioning through lower-energy intermediates via polypeptide unfolding.[66] Remarkably, this information appears to be retained on the timescale of seconds.[67]

At the highest energies accessible in the mass spectrometer, covalent fragmentation of the polypeptide backbone can be induced, Fig. 1.4.[48] This means that "top-down" proteomics is possible from the starting point of a native protein assembly.[68] Exploiting this capability enables the attribution of properties of individual protein assemblies to differences in their primary sequence.[69] A particularly attractive means for reducing native proteins into peptide fragments comes from irradiation of a protein complex using UV photons.[70] This has been shown to

operate on a fast timescale, providing comprehensive sequence coverage,[71,72] while retaining non-covalent interactions[71] and promoting a more symmetric dissociation pathway.[73] Interestingly, regions of enhanced fragmentation correlate reasonably well with solvent exposure in solution suggesting that aspects of the native fold are retained.[74] Similar correlations have also been found when using electron-mediated fragmentation.[60,75]

1.3.2 Structural interrogation of native and activated states

Aside from mass and charge, the information most frequently obtained from native MS experiments is CCS, through the implementation of an orthogonal IM dimension.[76] While the CCS contains significant information on the global shape of proteins,[77,78] and can be used to aid structural modelling,[79,80] complementary approaches are needed to probe structure on a local level. An attractive strategy is to perform labelling experiments analogous to those in solution,[8] in the vacuum of the mass spectrometer. This provides a means for probing the surface of a native protein, mass-separated from other species in solution. Furthermore, it can be used to interrogate the structures populated along the activation pathway, allowing for a detailed understanding of surface interactions and interfaces (Fig. 1.4, lower panels).

Gas-phase labelling strategies have mainly involved HDX reactions by collisions between the protein ion and a deuterated vapour.[81,82] The level of deuterium uptake, which manifests itself as a mass increase, has been shown to be sensitive to protein conformation in solution,[83–85] charge state,[84,86] and structural changes due to collisional activation.[87] Differences in deuterium uptake between a complex and its constituents are a measure of the occlusions and interactions at the interface. By dissociating the labelled complex to examine the partitioning of uptake over the constituents, information as to the associations made can be obtained.[88] This experiment relies on the activation not causing significant "scrambling" of deuterium labels between the dissociation products, a phenomenon that poses a challenge for increasing the spatial resolution of solution-phase HDX.[89] However, it appears that

scrambling across non-covalent interfaces is relatively minimal for small systems.[90] To extract the maximal amount of information from a labelling experiment requires localisation of the labels to individual amino acids. This is in principle possible using a top-down strategy, with the proviso that the activation that produces the peptide fragments outpaces any proton scrambling. The prospect of achieving this comes from the coupling of fast activation techniques with gpHDX,[91–95] but has yet to be explored fully on native proteins and assemblies.

An alternative means to probing protein structure in vacuum directly is afforded by spectroscopic studies. Notably, gas-phase IR spectra display distinguishing features that mirror those obtained from measurements in solution. This has been exploited to reveal the retention of protein secondary structure in native charge states,[96] and the formation of new structures upon activation,[97] while measurements on green fluorescent protein allowed the role of solvent on the chromophore to be probed.[98] Spectroscopic methods also have the potential for reporting intra-molecular distances. For instance, recent data has shown that Förster resonance energy transfer experiments in the gas phase are feasible.[99,100] A related means to obtain distance measures, which has the benefit of not requiring the addition of an extrinsic fluorescent label, comes from taking advantage of excitation energy transfer.[101] By studying disulphide bond cleavage, the proximity of aromatic amino acids can be determined through residue-specific absorption behaviour.[102] Though many of the techniques discussed above push MS-based methodologies towards revealing more structural information, the ideal would be to directly image individual proteins in the gas phase, rather than inferring structure from an ensemble. The attraction from an imaging perspective is the isolation of the protein, and the control afforded in the gas phase. Recent studies show that this ideal is not unrealistic (Fig. 1.5). Intact protein and protein complex ions can be mass selected for "soft landing" onto a surface for imaging.[107] While imaging the native structure is of obvious interest, the technique can also report on the conformational space accessible to the proteins as a function of charge state (Fig. 1.5A).[103] Landing native ions can also preserve intermolecular non-covalent

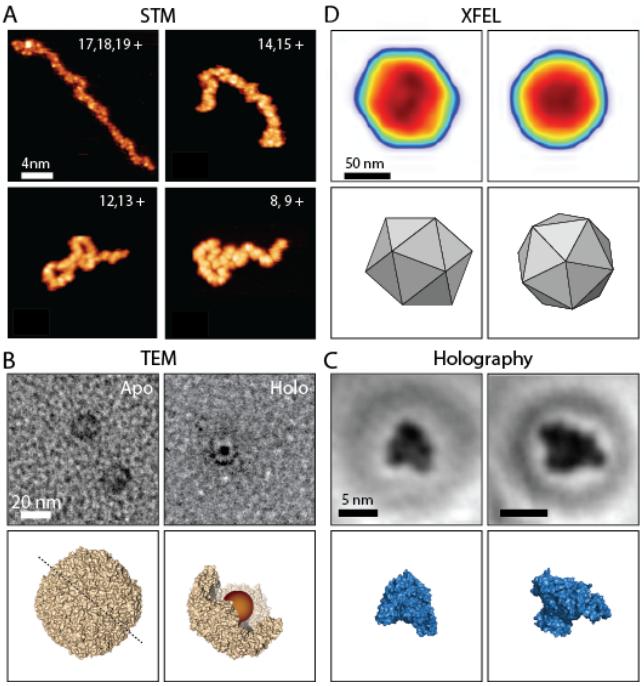


Figure 1.5. Gas-phase protein imaging. Imaging methods can be coupled to native MS to provide high-resolution views of proteins separated in vacuum. **A** Scanning tunnelling microscopy images of deposited cytochrome C ions shows a difference in unfolding as a function of charge state, with the highest charge states being almost completely extended.[103] **B** Transmission electron microscopy images of mass-selected and soft-landed apo- and holo-ferritin.[104] The crystal structure of the spherical ferritin cage is shown for comparison, and the cavity occupied with the iron core is illustrated. **C** Low-energy electron holograms of bovine serum albumin after soft-landing, reveals images of individual molecules that correspond to the native structures.[105] **D** Reconstruction projection images of carboxysomes generated from single-particle diffraction patterns from an X-ray free electron laser. Corresponding icosahedra are shown alongside in projection-matching orientations.[106] All panels have been modified from the original source for which I kindly thank Stephan Rauschenbach (**A & C**) and Max Hantke (**D**). Panel **B** was previous work performed by Benesch *et al.*[104] This figure was originally published in Chandler & Benesch 2018.[3]

interactions, with electron microscopy revealing structural differences between mass-selected species (Fig. 1.5B).[104] Recently, low-energy electron holography has achieved sub-nanometre resolution on single molecules, soft-landed onto ultra-clean graphene surfaces. Strikingly, native orientations of the landed proteins can be clearly seen (Fig. 1.5C).[105] An alternative approach, that does not require deposition on a surface, is to use an X-ray free electron laser which allows for the diffraction data from individual molecules to be obtained before their destruction. The current state-of-the-art is 20-nm resolution, but nonetheless different orientations of reconstructed single particle images of carboxysomes show good agreement with predicted geometric models (Fig. 1.5D).[106] These studies highlight how the

integration of imaging approaches with the selection capabilities, and potentially ion alignment,[108] of MS could enable novel methods for high-resolution structure determination.

1.4 Concluding remarks

In the 25 years since its inception, native MS has become an important approach for structural biology, and enabled many investigations into the structure that proteins populate at equilibrium. Here I have presented the case that this approach has great potential, some of which has already been realised, for probing the conformational space of proteins. However, while the preservation of native structure in the gas phase is clear, it remains to be seen to what extent the free-energy landscape resembles that in solution. In the absence of water, and consequently any competition between protein-solvent and intra-molecular solvation, the landscape will be more rugged, with higher barriers to conformational inter-conversion.[33] However, an important question is to what extent gas-phase trajectories in the protein folding funnel resemble the functional excursions from the native state in solution. Interestingly, it has been shown that, in protein crystals, mechanics stimulated by the application of external electric fields that act on local dipoles resemble the biologically relevant motions.[109] This suggests that exploration of the protein folding funnel in vacuum, where Coulombic interactions are dominant, will likely encompass some functional trajectories.

The work highlighted in this introduction reveals a key aspect of native MS, namely that the control afforded by ion manipulation in vacuum allows the implementation of multiple experiments on the same ion. The versatility, and multi-pronged nature of the approach will be key in probing the free-energy landscape of proteins. To this end, this Thesis focusses on probing the structure of both native and "non-native" species using mass spectrometric methods, and, for much of the work presented here, addresses the need to probe gas-phase structure *in situ* at a more local level.

References

[1] Lössl, P.; van de Waterbeemd, M.; Heck, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. *The EMBO Journal* **2016**, *35*, 2634–2657.

[2] Kaltashov, I. A.; Eyles, S. J. *Mass spectrometry in structural biology and biophysics: Architecture, dynamics, and interaction of biomolecules*, 2nd ed.; Wiley, 2012.

[3] Chandler, S. A.; Benesch, J. L. Mass spectrometry beyond the native state. *Current Opinion in Chemical Biology* **2018**, *42*, 130–137.

[4] Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422*, 198–207.

[5] Cox, J.; Mann, M. Quantitative, High-Resolution Proteomics for Data-Driven Systems Biology. *Annual Review of Biochemistry* **2011**, *80*, 273–299.

[6] Sinz, A. Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein-Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angewandte Chemie International Edition* **2018**, *57*, 6390–6396.

[7] Fontana, A.; De Laureto, P. P.; De Filippis, V.; Scaramella, E.; Zambonin, M. Probing the partly folded states of proteins by limited proteolysis. *Folding and Design* **1997**, *2*, 17–26.

[8] Konermann, L.; Vahidi, S.; Sowole, M. A. Mass spectrometry methods for studying structure and dynamics of biological macromolecules. *Analytical Chemistry* **2014**, *86*, 213–232.

[9] Hambly, D. M.; Gross, M. L. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *Journal of the American Society for Mass Spectrometry* **2005**, *16*, 2057–2063.

[10] Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C. High-resolution mapping of carbene-based protein footprints. *Analytical Chemistry* **2012**, *84*, 4411–4418.

[11] Englander, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical Perspective. *Journal of the American Society for Mass Spectrometry* **2006**, *17*, 1481–1489.

[12] Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 5–13.

[13] Winston, R. L.; Fitzgerald, M. C. Mass spectrometry as a readout of protein structure and function. *Mass Spectrometry Reviews* **1997**, *16*, 165–179.

[14] Breuker, K.; McLafferty, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase, 10^{-12} to 10^{2} s. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18145–52.

[15] Hilton, G. R.; Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of The Royal Society Interface* **2012**, *9*, 801–816.

[16] Meyer, T.; Gabelica, V.; Grubmüller, H.; Orozco, M. Proteins in the gas phase. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 408–425.

[17] Kitova, E. N.; Seo, M.; Roy, P. N.; Klassen, J. S. Elucidating the intermolecular interactions within a desolvated protein-ligand complex. An experimental and computational study. *Journal of the American Chemical Society* **2008**, *130*, 1214–1226.

[18] Sheu, S.-Y.; Yang, D.-Y.; Selzle, H. L.; Schlag, E. W. Energetics of hydrogen bonds in peptides. *Proceedings of the National Academy of Sciences* **2003**, *100*, 12683–12687.

[19] Ruotolo, B. T.; Robinson, C. V. Aspects of native proteins are retained in vacuum. *Current Opinion in Chemical Biology* **2006**, *10*, 402–408.

[20] Benesch, J. L. P.; Robinson, C. V. Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Current opinion in structural biology* **2006**, *16*, 245–51.

[21] Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nature Methods* **2012**, *9*, 1084–1086.

[22] Jørgensen, T. J.; Roepstorff, P.; Heck, A. J. Direct Determination of Solution Binding Constants for Noncovalent Complexes between Bacterial Cell Wall Peptide Analogues and Vancomycin Group Antibiotics by Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **1998**, *70*, 4427–4432.

[23] Hagan, N.; Fabris, D. Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 ψ-recognition element. *Biochemistry* **2003**, *42*, 10736–10745.

[24] Aquilina, J. A.; Benesch, J. L. P.; Bateman, O. A.; Slingsby, C.; Robinson, C. V. Polydispersity of a mammalian chaperone: Mass spectrometry reveals the population of oligomers in B-crystallin. *Proceedings of the National Academy of Sciences* **2003**, *100*, 10611–10616.

[25] Mishra, S.; Chandler, S. A.; Williams, D.; Claxton, D. P.; Koteiche, H. A.; Stewart, P. L.; Benesch, J. L.; Mchaourab, H. S. Engineering of a Polydisperse Small Heat-Shock Protein Reveals Conserved Motifs of Oligomer Plasticity. *Structure* **2018**, *26*, 1116–1126.e4.

[26] Pliotas, C. et al. Adenosine Monophosphate Binding Stabilizes the KTN Domain of the Shewanella denitrificans Kef Potassium Efflux System. *Biochemistry* **2017**, *56*, 4219–4234.

[27] Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Analytical chemistry* **2015**,

[28] Shimon, L.; Sharon, M.; Horovitz, A. A method for removing effects of nonspecific binding on the distribution of binding stoichiometries: Application to mass spectroscopy data. *Biophysical Journal* **2010**, *99*, 1645–1649.

[29] Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S. Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* **2011**, *10*, M111.011015.

[30] Gault, J. et al. High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nature Methods* **2016**, *13*, 333–336.

[31] van de Waterbeemd, M.; Fort, K. L.; Boll, D.; Reinhardt-Szyba, M.; Routh, A.; Makarov, A.; Heck, A. J. R. High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles. *Nature Methods* **2017**, *14*, 283–286.

[32] Gerken, P. A. et al. Discovery of a Highly Selective Cell-Active Inhibitor of the Histone Lysine Demethylases KDM2/7. *Angewandte Chemie - International Edition* **2017**, *56*, 15555–15559.

[33] Clemmer, D. E.; Russell, D. H.; Williams, E. R. Characterizing the Conformationome : Toward a Structural Understanding of the Proteome. *Accounts of Chemical Research* **2017**, *50*, 556–560.

[34] Wyttenbach, T.; Pierson, N. A.; Clemmer, D. E.; Bowers, M. T. Ion Mobility Analysis of Molecular Dynamics. *Annual Review of Physical Chemistry* **2014**, *65*, 175–196.

[35] Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338*, 1042–1046.

[36] Allen, S.; Schwartz, A.; Bush, M. Effects of polarity on the structures and charge states of native-like proteins and protein complexes in the gas phase. *Analytical chemistry* **2013**, *85*, 12055–12061.

[37] Borysik, A. J.; Kovacs, D.; Guharoy, M.; Tompa, P. Ensemble Methods Enable a New Definition for the Solution to Gas-Phase Transfer of Intrinsically Disordered Proteins. *Journal of the American Chemical Society* **2015**, *137*, 13807–13817.

[38] Jhingree, J. R.; Bellina, B.; Pacholarz, K. J.; Barran, P. E. Charge Mediated Compaction and Rearrangement of Gas-Phase Proteins: A Case Study Considering Two Proteins at Opposing Ends of the Structure-Disorder Continuum. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 1450–1461.

[39] Gitlin, I.; Carbeck, J. D.; Whitesides, G. M. Why are proteins charged? Networks of charge-charge interactions in proteins measured by charge ladders and capillary electrophoresis. *Angewandte Chemie (International ed. in English)* **2006**, *45*, 3022–60.

[40] Laszlo, K. J.; Bush, M. F. Analysis of Native-Like Proteins and Protein Complexes Using Cation to Anion Proton Transfer Reactions (CAPTR). *Journal of the American Society for Mass Spectrometry* **2015**, *26*, 2152–2161.

[41] Zhao, Q.; Soyk, M. W.; Schieffer, G. M.; Fuhrer, K.; Gonin, M. M.; Houk, R. S.; Badman, E. R. An Ion Trap-Ion Mobility-Time of Flight Mass Spectrometer with Three Ion Sources for Ion/Ion Reactions. *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 1549–1561.

[42] Jhingree, J. R.; Beveridge, R.; Dickinson, E. R.; Williams, J. P.; Brown, J. M.; Bellina, B.; Barran, P. E. Electron transfer with no dissociation ion mobility–mass spectrometry (ETnoD IM-MS). The effect of charge reduction on protein conformation. *International Journal of Mass Spectrometry* **2017**, *413*, 43–51.

[43] Lermyte, F.; Łącki, M. K.; Valkenborg, D.; Gambin, A.; Sobott, F. Conformational Space and Stability of ETD Charge Reduction Products of Ubiquitin. *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 69–76.

[44] Laszlo, K. J.; Buckner, J. H.; Munger, E. B.; Bush, M. F. Native-Like and Denatured Cytochrome c Ions Yield Cation-to-Anion Proton Transfer Reaction Products with Similar Collision Cross-Sections. *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 1382–1391.

[45] Laszlo, K. J.; Munger, E. B.; Bush, M. F. Folding of Protein Ions in the Gas Phase after Cation-to-Anion Proton-Transfer Reactions. *Journal of the American Chemical Society* **2016**, *138*, 9581–9588.

[46] Zhao, Q.; Schieffer, G. M.; Soyk, M. W.; Anderson, T. J.; Houk, R. S.; Badman, E. R. Effects of ion/ion proton transfer reactions on conformation of gas-phase cytochrome c ions. *Journal of the American Society for Mass Spectrometry* **2010**, *21*, 1208–1217.

[47] Laszlo, K. J.; Bush, M. F. Interpreting the Collision Cross Sections of Native-like Protein Ions: Insights from Cation-to-Anion Proton-Transfer Reactions. *Analytical Chemistry* **2017**, *89*, 7607–7614.

[48] Benesch, J. L. P. Collisional activation of protein complexes: picking up the pieces. *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 341–8.

[49] Wagner, N. D.; Clemmer, D. E.; Russell, D. H. ESI-IM-MS and Collision-Induced Unfolding That Provide Insight into the Linkage-Dependent Interfacial Interactions of Covalently Linked Diubiquitin. *Analytical Chemistry* **2017**, *89*, 10094–10103.

[50] Koeniger, S. L.; Clemmer, D. E. Resolution and structural transitions of elongated states of ubiquitin. *Journal of the American Society for Mass Spectrometry* **2007**, *18*, 322–331.

[51] Botzanowski, T.; Erb, S.; Hernandez-Alba, O.; Ehkirch, A.; Colas, O.; Wagner-Rousset, E.; Rabuka, D.; Beck, A.; Drake, P. M.; Cianférani, S. Insights from native mass spectrometry approaches for top- and middle- level characterization of site-specific antibody-drug conjugates. *mAbs* **2017**, *9*, 801–811.

[52] Huang, Y.; Salinas, N. D.; Chen, E.; Tolia, N. H.; Gross, M. L. Native Mass Spectrometry, Ion mobility, and Collision-Induced Unfolding Categorize Malaria Antigen/Antibody Binding. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 2515–2518.

[53] Pacholarz, K. J.; Barran, P. E. Use of a charge reducing agent to enable intact mass analysis of cysteine-linked antibody-drug-conjugates by native mass spectrometry. *EuPA Open Proteomics* **2016**, *11*, 23–27.

[54] Tian, Y.; Han, L.; Buckner, A. C.; Ruotolo, B. T. Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding Patterns, Glycosylation, and Structures. *Analytical Chemistry* **2015**, *87*, 11509–11515.

[55] Allison, T. M.; Reading, E.; Liko, I.; Baldwin, A. J.; Laganowsky, A.; Robinson, C. V.

Quantifying the stabilizing effects of protein–ligand interactions in the gas phase. *Nature Communications* **2015**, *6*, 8551.

[56] Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V. Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **2014**, *510*, 172–5.

[57] Landreh, M.; Marklund, E. G.; Uzdavinys, P.; Degiacomi, M. T.; Coincon, M.; Gault, J.; Gupta, K.; Liko, I.; Benesch, J. L.; Drew, D.; Robinson, C. V. Integrating mass spectrometry with MD simulations reveals the role of lipids in Na^+/H^+ antiporters. *Nature Communications* **2017**, *8*, 1–9.

[58] Zhong, Y.; Han, L.; Ruotolo, B. T. Collisional and Coulombic unfolding of gas-phase proteins: high correlation to their domain structures in solution. *Angewandte Chemie (International ed. in English)* **2014**, *53*, 9209–12.

[59] Eschweiler, J. D.; Martini, R. M.; Ruotolo, B. T. Chemical probes and engineered constructs reveal a detailed unfolding mechanism for a solvent-free multidomain protein. *Journal of the American Chemical Society* **2017**, *139*, 534–540.

[60] Lermyte, F.; Sobott, F. Electron transfer dissociation provides higher-order structural information of native and partially unfolded protein complexes. *Proteomics* **2015**, *15*, 2813–2822.

[61] Shepherd, D. A.; Marty, M. T.; Giles, K.; Baldwin, A. J.; Benesch, J. L. Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins. *International Journal of Mass Spectrometry* **2015**, *377*, 663–671.

[62] Wang, G.; de Jong, R. N.; van den Bremer, E. T. J.; Parren, P. W. H. I.; Heck, A. J. R. Enhancing Accuracy in Molecular Weight Determination of Highly Heterogeneously Glycosylated Proteins by Native Tandem Mass Spectrometry. *Analytical Chemistry* **2017**, *89*, 4793–4797.

[63] Marty, M. T.; Hoi, K. K.; Gault, J.; Robinson, C. V. Probing the Lipid Annular Belt by Gas-Phase Dissociation of Membrane Proteins in Nanodiscs. *Angewandte Chemie - International Edition* **2016**, *55*, 550–554.

[64] Quintyn, R. S.; Yan, J.; Wysocki, V. H. Surface-induced dissociation of homotetramers with D_2 symmetry yields their assembly pathways and characterizes the effect of ligand binding. *Chemistry and Biology* **2015**, *22*, 583–592.

[65] Song, Y.; Nelp, M. T.; Bandarian, V.; Wysocki, V. H. Refining the structural model of a heterohexameric protein complex: Surface induced dissociation and ion mobility provide key connectivity and topology information. *ACS Central Science* **2015**, *1*, 477–487.

[66] Zhou, M.; Wysocki, V. H. Surface induced dissociation: Dissecting noncovalent protein complexes in the gas phase. *Accounts of Chemical Research* **2014**, *47*, 1010–1018.

[67] Harvey, S. R.; Yan, J.; Brown, J. M.; Hoyes, E.; Wysocki, V. H. Extended Gas-Phase Trapping Followed by Surface-Induced Dissociation of Noncovalent Protein Complexes. *Analytical Chemistry* **2016**, *88*, 1218–1221.

[68] Belov, M. E.; Damoc, E.; Denisov, E.; Compton, P. D.; Horning, S.; Makarov, A. A.; Kelleher, N. L. From protein complexes to subunit backbone fragments: A multi-stage approach to native mass spectrometry. *Analytical Chemistry* **2013**, *85*, 11163–11173.

[69] Ben-Nissan, G.; Belov, M. E.; Morgenstern, D.; Levin, Y.; Dym, O.; Arkind, G.; Lipson, C.; Makarov, A. A.; Sharon, M. Triple-Stage Mass Spectrometry Unravels the Heterogeneity of an Endogenous Protein Complex. *Analytical Chemistry* **2017**, *89*, 4708–4715.

[70] R. Julian, R. The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 1823–1826.

[71] O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. Characterization of native protein complexes using ultraviolet photodissociation mass spectrometry. *Journal of the American Chemical Society* **2014**, *136*, 12920–12928.

[72] Theisen, A.; Yan, B.; Brown, J. M.; Morris, M.; Bellina, B.; Barran, P. E. Use of Ultraviolet Photodissociation Coupled with Ion Mobility Mass Spectrometry to Determine Structure and Sequence from Drift Time Selected Peptides and Proteins. *Analytical Chemistry* **2016**, *88*, 9964–9971.

[73] Tamara, S.; Dyachenko, A.; Fort, K. L.; Makarov, A. A.; Scheltema, R. A.; Heck, A. J. R. Symmetry of Charge Partitioning in Collisional and UV Photon-Induced Dissociation of Protein Assemblies. *Journal of the American Chemical Society* **2016**, *138*, 10860–10868.

[74] Morrison, L. J.; Brodbelt, J. S. 193 nm Ultraviolet Photodissociation Mass Spectrometry of Tetrameric Protein Complexes Provides Insight into Quaternary and Secondary Protein Topology. *Journal of the American Chemical Society* **2016**, *138*, 10849–10859.

[75] Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. Native top-down electrospray ionization-mass spectrometry of 158 kDa protein complex by high-resolution fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry* **2014**, *86*, 317–320.

[76] Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature Chemistry* **2014**, *6*, 281–294.

[77] Ewing, S. A.; Donor, M. T.; Wilson, J. W.; Prell, J. S. Collidoscope: An Improved Tool for Computing Collisional Cross-Sections with the Trajectory Method. *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 587–596.

[78] Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. Collision Cross Sections for Structural Proteomics. *Structure* **2015**, *23*, 791–799.

[79] Ahdash, Z.; Pyle, E.; Politis, A. Hybrid Mass Spectrometry: Towards Characterization of Protein Conformational States. *Trends in Biochemical Sciences* **2016**, *41*, 650–653.

[80] Thalassinos, K.; Pandurangan, A. P.; Xu, M.; Alber, F.; Topf, M. Conformational states of macromolecular assemblies explored by integrative structure calculation. *Structure* **2013**, *21*, 1500–1508.

[81] Green, M. K.; Lebrilla, C. B. Ion-molecule reactions as probes of gas-phase structures of peptides and proteins. *Mass spectrometry reviews* **1997**, *16*, 53–71.

[82] Maurer, M. M.; Donohoe, G. C.; Valentine, S. J. Advances in ion mobility-mass spectrometry instrumentation and techniques for characterizing structural heterogeneity. *The Analyst* **2015**, *140*, 6782–6798.

[83] Beeston, H. S.; Ault, J. R.; Pringle, S. D.; Brown, J. M.; Ashcroft, A. E. Changes in protein structure monitored by use of gas-phase hydrogen/deuterium exchange. *Proteomics* **2015**, 2842–2850.

[84] Rand, K. D.; Pringle, S. D.; Murphy, J. P.; Fadgen, K. E.; Brown, J.; Engen, J. R. Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Analytical chemistry* **2009**, *81*, 10019–28.

[85] Zhu, S.; Campbell, J. L.; Chernushevich, I.; Le Blanc, J. C.; Wilson, D. J. Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution. *Journal of the American Society for Mass Spectrometry* **2016**, *27*, 991–999.

[86] Donohoe, G. C.; Khakinejad, M.; Valentine, S. J. Ion mobility spectrometry-hydrogen deuterium exchange mass spectrometry of anions: part 1. Peptides to proteins. *Journal of the American Society for Mass Spectrometry* **2015**, *26*, 564–76.

[87] Valentine, S. J.; Clemmer, D. E. H/D Exchange Levels of Shape-Resolved Cytochrome c Conformers in the Gas Phase. *Journal of the American Chemical Society* **1997**, *119*, 3558–3566.

[88] Mistarz, U. H.; Brown, J. M.; Haselmann, K. F.; Rand, K. D. Probing the Binding Interfaces of Protein Complexes Using Gas-Phase H/D Exchange Mass Spectrometry. *Structure* **2016**, *24*, 310–318.

[89] Rand, K. D.; Zehl, M.; Jørgensen, T. J. Measuring the hydrogen/deuterium exchange of proteins at high spatial resolution by mass spectrometry: Overcoming gas-phase hydrogen/deuterium scrambling. *Accounts of Chemical Research* **2014**, *47*, 3018–3027.

[90] Modzel, M.; Stefanowicz, P.; Szewczuk, Z. Hydrogen scrambling in non-covalent complexes of peptides. *Rapid Communications in Mass Spectrometry* **2012**, *26*, 2739–2744.

[91] Khakinejad, M.; Kondalaji, S. G.; Tafreshian, A.; Valentine, S. J. Gas-Phase Hydrogen-Deuterium Exchange Labeling of Select Peptide Ion Conformer Types: a Per-Residue Kinetics Analysis. *Journal of The American Society for Mass Spectrometry* **2015**, 1115–1127.

[92] Mistarz, U. H.; Brown, J. M.; Haselmann, K. F.; Rand, K. D. A Simple Setup for Gas-Phase H/D Exchange Mass Spectrometry Coupled to Electron Transfer Dissociation and Ion Mobility for Analysis of Polypeptide Structure on a Liquid Chromatographic Timescale. *Analytical chemistry* **2014**, *86*, 11868–11876.

[93] Pan, J.; Heath, B. L.; Jockusch, R. a.; Konermann, L. Structural interrogation of electrosprayed peptide ions by gas-phase H/D exchange and electron capture dissociation mass spectrometry. *Analytical Chemistry* **2012**, *84*, 373–378.

[94] Rand, K. D.; Pringle, S. D.; Morris, M.; Brown, J. M. Site-Specific Analysis of Gas-Phase Hydrogen/Deuterium Exchange of Peptides and Proteins by Electron Transfer Dissociation. *Analytical Chemistry* **2012**, *84*, 1931–1940.

[95] Mistarz, U. H.; Bellina, B.; Jensen, P. F.; Brown, J. M.; Barran, P. E.; Rand, K. D. UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Analytical Chemistry* **2018**, *90*, 1077–1080.

[96] Seo, J.; Hoffmann, W.; Warnke, S.; Bowers, M. T.; Pagel, K.; von Helden, G. Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study. *Angewandte Chemie - International Edition* **2016**, *55*, 14173–14176.

[97] González Flórez, A. I.; Mucha, E.; Ahn, D. S.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; Von Helden, G. Charge-Induced Unzipping of Isolated Proteins to a Defined Secondary Structure. *Angewandte Chemie - International Edition* **2016**, *55*, 3295–3299.

[98] Frankevich, V.; Barylyuk, K.; Chingin, K.; Nieckarz, R.; Zenobi, R. Native Biomolecules in the Gas Phase? The Case of Green Fluorescent Protein. *ChemPhysChem* **2013**, *14*, 929–935.

[99] Czar, M. F.; Zosel, F.; König, I.; Nettels, D.; Wunderlich, B.; Schuler, B.; Zarrine-Afsar, A.; Jockusch, R. A. Gas-Phase FRET Efficiency Measurements to Probe the Conformation of Mass-Selected Proteins. *Analytical Chemistry* **2015**, *87*, 7559–7565.

[100] Daly, S.; Knight, G.; Halim, M. A.; Kulesza, A.; Choi, C. M.; Chirot, F.; MacAleese, L.; Antoine, R.; Dugourd, P. Action-FRET of a Gaseous Protein. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 38–49.

[101] Hendricks, N. G.; Lareau, N. M.; Stow, S. M.; McLean, J. A.; Julian, R. R. Bond-specific dissociation following excitation energy transfer for distance constraint determination in the gas phase. *Journal of the American Chemical Society* **2014**, *136*, 13363–13370.

[102] Hendricks, N. G.; Julian, R. R. Two-step energy transfer enables use of phenylalanine in action-EET for distance constraint determination in gaseous biomolecules. *Chem. Commun.* **2015**, *51*, 12720–12723.

[103] Rinke, G.; Rauschenbach, S.; Harnau, L.; Albarghash, A.; Pauly, M.; Kern, K. Active Conformation Control of Unfolded Proteins by Hyperthermal Collision with a Metal Surface. *Nano Letters* **2014**, *14*, 5609–5615.

[104] Mikhailov, V. A.; Mize, T. H.; Benesch, J. L. P.; Robinson, C. V. Mass-selective soft-landing of protein assemblies with controlled landing energies. *Analytical Chemistry* **2014**, *86*, 8321–8328.

[105] Longchamp, J.-N.; Rauschenbach, S.; Abb, S.; Escher, C.; Latychevskaia, T.; Kern, K.; Fink, H.-W. Imaging proteins at the single-molecule level. *Proceedings of the National Academy of Sciences* **2017**, *114*, 1474–1479.

[106] Hantke, M. F. et al. High-throughput imaging of heterogeneous cell organelles with an X-ray laser. *Nature Photonics* **2014**, *8*, 943–949.

[107] Rauschenbach, S.; Ternes, M.; Harnau, L.; Kern, K. Mass Spectrometry as a Preparative Tool for the Surface Science of Large Molecules. *Annual Review of Analytical Chemistry* **2016**, *9*, 473–498.

[108] Marklund, E. G.; Ekeberg, T.; Moog, M.; Benesch, J. L.; Caleman, C. Controlling Protein Orientation in Vacuum Using Electric Fields. *Journal of Physical Chemistry Letters* **2017**, *8*, 4540–4544.

[109] Hekstra, D. R.; White, K. I.; Socolich, M. A.; Henning, R. W.; Šrajer, V.; Ranganathan, R. Electric-field-stimulated protein mechanics. *Nature* **2016**, *540*, 400–405.

2

Native mass spectrometry background

2.1 Mass spectrometry instrumentation

Since the inception of native mass spectrometry (MS), many types and architectures of instrumentation have been used depending on what challenges and questions the protein complex requires. Common to all spectrometers are three principal stages: ionisation, ion transmission and manipulation, and mass analysis and detection. The work presented in this thesis used one particular architecture throughout, a nano-electrospray quadrupole time-of-flight spectrometer with ion mobility (IM), Q-IM-ToF, (Waters Synapt G1); here I describe each aspect of the instrument in turn with a particular emphasis on its role in native MS.

An instrument schematic is provided in Figure 2.1A showing the progression of the ion beam from the electrospray source to the time-of-flight (ToF) mass analyser. Native protein ions enter the source region and then become dampened and focussed before entering the quadrupole mass analyser for m/z selection or filtering. Additional beam focussing and attenuation, if required, then occurs before entering the first of the instrument's two collision cells. In ToF operating mode, without an IM separation, ions pass through three stacked ring radio-frequency (RF) ion guides propelled through by moving electric field pulses. If ion activation

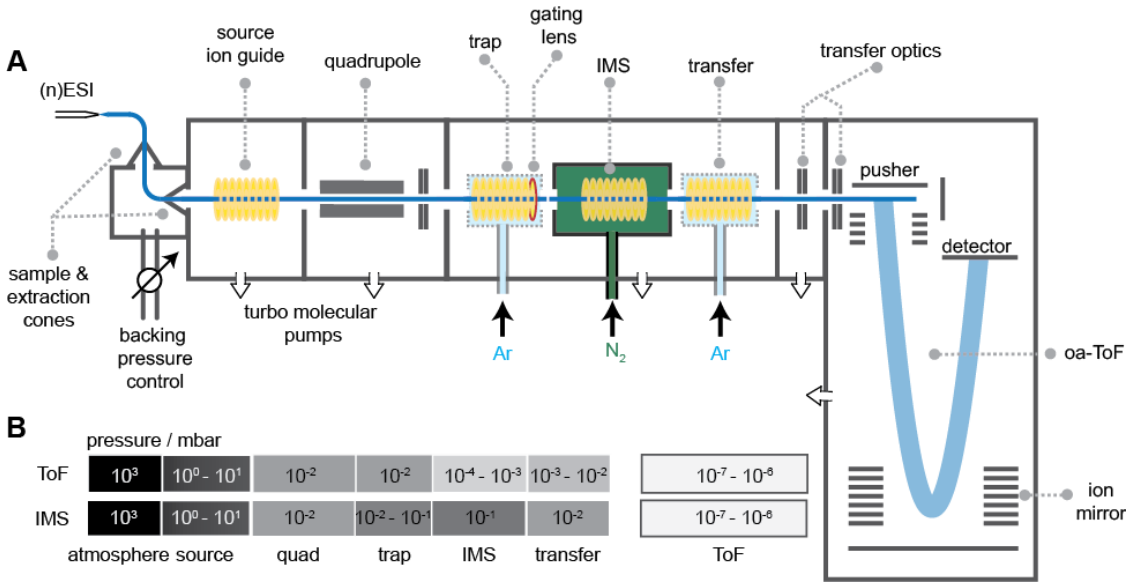


Figure 2.1. Schematic of a Waters Synapt G1 ion mobility Q-ToF mass spectrometer **A** Proteins are ionised from solution using a nano-electrospray source at atmosphere before becoming kinetically dampened and focussed in a stacked ring RF ion guide with a tunable backing pressure. A quadrupole then allows ion gating or selection according to *m/z*. Collisional activation is possible in both the trap and transfer ion guides typically filled with argon. In IM mode, ions are gated in the trap cell before being separated according to their size by ion mobility(IM). Ions are then focussed before entering the time-of-flight (ToF) drift tube. A pusher and a series of conducting grids accelerate ions orthogonally into a drift tube before ion detection. The m/z can then be calculated according to the flight time. **B** The pressure of each region of the spectrometer is indicated (mbar) when in ToF or ion mobility modes. A range is provided which reflects the conditions used for low and high *m/z* species. Each region is discussed in greater detail in the following sections.

is required, ions may be accelerated into the trap and/or the transfer collision cells typically filled with argon. In mobility mode, ions are gated in the trap ion guide for the duration of the IM duty cycle. Ions are then injected into the travelling wave (TW) IM cell, at an elevated pressure of nitrogen, before passing through the transfer cell to the ToF region. Ions enter the low pressure ($10^{-7} - 10^{-6}$) ToF orthogonal drift tube for mass analysis and detection. The pressure range for each region, in ToF and IM mode, is provided in Figure 2.1B for reference throughout the thesis. Ranges reflect the tunable operating pressures possible for transmission of low and high mass species. Greater detail on each stage is provided in the following sections.

2.1.1 Electrospray ionisation

The advent of electrospray (ESI) for the ionisation of intact biological macromolecules came about in the 1980s, having previously been employed for macromolecular organic systems,[1,2] and provided a means for the gentle transferral of proteins into the gas phase as charged ions. Seminal work[3,4] showed that ESI could comfortably surpass the mass limitations for existing ionisation methods[5] whilst maintaining intact structure, and by the 1990s it was clear that the technique could be used to preserve and study non-covalent protein complexes.[6]

Native MS often utilises ESI as a means of transferring intact biomolecules in their native structural states to the vacuum of the gas-phase. Native MS at its very core is the act of preserving the delicate, often non-covalent, interactions present in bimolecular systems throughout the analysis: from bulk solution conditions through to the ESI process and ultimately manipulation throughout the spectrometer. Native MS firstly requires an aqueous sample buffer that is both sufficiently volatile that electrolyte ions are lost through evaporation during electrospray, reducing adduction to the sample, and is able to provide the ionic strength required to maintain stabilising interactions.[7] Ammonium acetate or ammonium bicarbonate fulfil this requirement and are the most commonly employed buffers.[8] The buffer composition is chosen to mimic the physiological environment whilst maintaining ESI-MS compatibility, as such physiological pH and avoidance of organic solvents are important. It has been noted that ammonium acetate is able to restrict the pH fluctuations throughout electrospray but at neutral pH does not have any buffering capacity, rather its buffering action only begins at higher or lower pH.[9] Nonetheless, ammonium acetate buffers remain commonplace and provide the desired ionic strength and volatility.

Once the native protein has been loaded into an ESI emitter one also has a degree of control over maintaining structure. The temperature of the instrument source may be increased to encourage droplet evaporation, though a compromise must be made as to how this will effect both the bulk solution in the needle tip[10,11]

and how protein structure may be effected as a droplet and down-stream naked ion.[12,13] Emitter design may help limit temperature requirements and is discussed later.[14] Akin to increasing the source temperature one can also elevate declustering potentials in the source region to help with efficient evaporation. Care must also be taken when applying such voltages as to how this may also activate the protein with different instrumental architectures and platforms having differing impacts.[12] The control over the parameters discussed can often be borne out through low charge states and compact collisional cross sections as described in this section and § 2.1.6.

A protein solution is sprayed from a glass capillary and becomes charged through liquid contact with a conductive coating at the capillary tip, or by insertion of a metal electrode directly into the solution from the rear of the tube. Emitter diameters in electrospray are typically in the hundreds of μm range and require high μL/min flow rates to maintain droplet formation. NanoESI (nESI) was developed to overcome the relatively high consumption rate and temperatures required in ESI; in reducing the tip diameter to ∼5 μm, or less, a lower sample consumption is achieved and importantly elevated temperatures are not required, even facilitating the addition of low volatility solvents and salts. As the desire to spray proteins from ever more physiological buffers increases, there is now a resurgence in research around emitter design. Moving towards smaller tip diameters aids with reducing salt adducts and emitter optimisation, specifically reduction of the tip orifice diameter, is required to maintain quality spectra. Sub-micron emitters can tolerate high concentrations of non-volatile salts, >100 mM, owing to fewer salt adducts forming from smaller droplets.[15,16]

A high voltage capillary, 1 - 3 kV, is positioned within a few millimetres of the source orifice,[17] a schematic is shown of the process in Figure 2.2A. Once a stable spray has been established a Taylor cone will develop forming a jet from which a charged droplet will be formed and break away. During positive ionisation, charged ions, both proteins and small cationic species, will remain in the interior of the droplet to remain charge-solvated. This will result in a net positive surface caused by the unsolvated positive dipole of surface solvent molecules.[18] As the

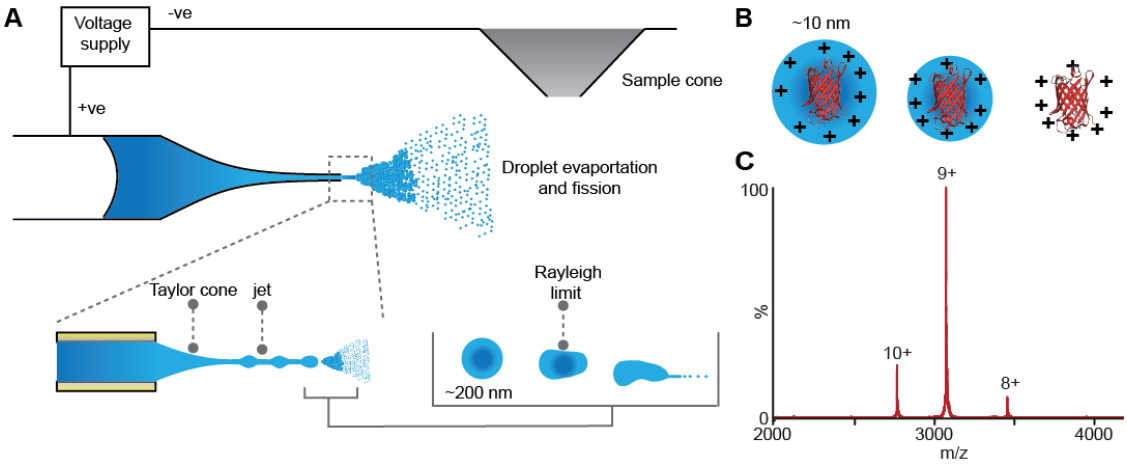


Figure 2.2. Electrospray ionisation of proteins. **A** A protein solution is sprayed from the orifice of a small diameter conducting capillary. The solution is charged using a high voltage applied to the solution typically via the capillary coating at the tip. A Taylor cone is formed which tapers into a jet and becomes unstable forming discrete droplets (~200 nm). As solvent evaporates the surface charge increases until the droplet becomes unstable around the Rayleigh limit, at which point asymmetric distortion and fission occurs expelling many more smaller droplets. **B** This process occurs until the droplet size becomes comparable to the protein. The charged residue model (CRM) suggests that in the final stages of electrospray layers of solvent evaporate until the remaining droplet charge is deposited onto the surface of the protein. **C** A native mass spectrum generated from nano-electrospray for a 28 kDa globular soluble protein. A narrow charge series is observed due to stochastic deposition of charge throughout the described process.

solvent evaporates the droplet will shrink, increasing the charge density until the surface charge repulsion overcomes the surface tension and a Coulombic explosion or fission event occurs. The ejection of small cations or small molecules may be ionised in this way—the ion ejection model (IEM)[19]—acting to carry charge away from the parent droplet. This process will be repeated until the final layer of solvent around the globular protein has evaporated depositing the remaining charge—the charged residue model (CRM)—as first described by Dole,[20] Fig. 2.2B. Decades of experimental and simulated evidence now supports this mechanism,[21] whilst unfolded or intrinsically disordered systems are thought to progress by alternative routes.[18] The onset of fission was predicted by the Rayleigh limit[22] and shown to scale with increasing droplet radius according to equation 2.1:

$$Q_{Ry} = 8\pi(\epsilon_0 \gamma R^3)^{1/2} \tag{2.1}$$

where Q_{Ry} is the charge on the droplet, R is the radius of the droplet, γ is the surface tension of the solvent and ϵ_0 is the vacuum permittivity. Experiments have

shown that for an aqueous droplet the onset of instability remains within 10 % of the Rayleigh limit,[23] with experimentally deduced final charge states falling in the range of 75 - 95 % of this limit.[24]

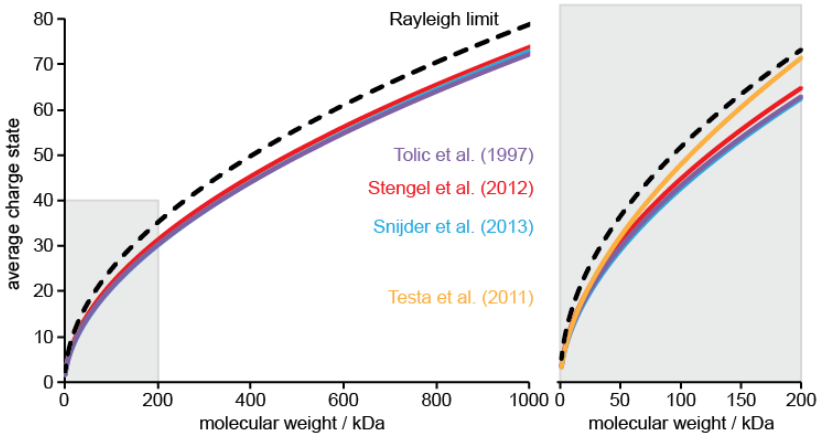


Figure 2.3. Average charge state of globular soluble proteins. The relationship between average charge state, z, generated through nESI and the molecular weight, M_W, is given by the general empirical relation $z_{avg} = aM_W{}^x$. Where the exponent, x, is typically between 0.53 - 0.55 and a is a scaling constant.[25–28] The Rayleigh limit is shown for comparison, dashed line, having been calculated for an assumed spherical droplet of constant density (1 g cm^{-3}) and a surface tension of 0.072 N/m. A zoom is shown in the grey box for small molecular weight complexes. The simulated average charge state data determined by Testa *et al.* is omitted on the full scale as it exceeds the Rayleigh limit above ~200 kDa and is inconsistent with the other three studies.

The surface area of the protein is a large determinant of the resulting charge deposition and can be used to predict the average charge state resulting from electrospray, and indeed the inverse can often help inform on solution structure.[29] As native protein structures have relatively low surface areas for their mass, charge states are typically low, resulting in a charge series at high m/z.[30] Multiple charge states are observed due to the stochastic nature of charge deposition in the final stages of electrospray; the narrow charge series for a native 28 kDa globular soluble protein is shown in Figure 2.2C. Treating a globular compact protein as a sphere of constant density, ρ, means the ion radius can be estimated from the mass, M_W, giving:

$$R^3 = \frac{3M_W}{4\pi\rho} \tag{2.2}$$

This would result in the following relationship between surface area, S_A, and molecular weight: $S_A \sim M_W{}^{2/3}$. This relationship was predicted based on the

relationship between a progeny droplet volume and charge[31] with the exponent scaling term highly consistent with experimental determinations, 0.60 - 0.69.[26,29,32–34] As surface areas can often be difficult to calculate experimentally many groups have explored the relation between average charge and the molecular weight of native soluble proteins with their relationships plotted in Figure 2.3.[25–28] Most studies have shown that the average charge scales with ${M_W}^{0.53-0.55}$ with the exception of the work by Testa *et al.*, whose fit would exceed the Rayleigh limit for m/z greater than 200 kDa. Experiments come close but remain lower than the Rayleigh limit, estimated for a sphere of density 1 g/cm^3 and shown as a dashed line, which may reflect the assumptions made in predicting R from the molecular weight or that additional charge loss is caused through competing mechanisms.[35] Recent molecular dynamics simulations support the CRM of protein charging with the ejection of solvated cations, IEM, as the charge density approaches the Rayleigh limit.[24,36] It is widely accepted that the main determinant of the charge state resulting from nESI is the surface area though this does not remain undisputed with some arguing gas-phase basicity of the surface has an important role,[37] though the number of basic sites will often correlate with the surface area. Others argue against this model[38] and indeed our data presented in Chapter 4 provide evidence against this view.

Charge state distributions provide some indication as to the conformational flexibility[39,40] and where charge states are consistent with those predicted one often has immediate confirmation as to whether a protein has remained folded, though this may be misleading for intrinsically disordered systems[33,34] and membrane proteins whose charge can be influenced by the membrane mimetic.[41] Charge states also play an important role in preserving non-covalent complexes. As the charge on the protein surface decreases the Coulombic strain will diminish, this will act to maintain native interfaces but also limit the Coulombic repulsion between binding partners.[42–45] Charge has been shown to be reduced in many ways either through solution additives,[46–48] ion-ion reactions inside the spectrometer[49,50] and in-source

additives/reactions.[51,52] The ability to manipulate, and in particular, reduce charge can have important consequences on detecting intact complexes.

2.1.2 Transmission of high molecular weight ions

As ions enter the source region of the spectrometer they experience a pressure drop from atmosphere to $\sim 10^{-3}$ mbar over a distance of a few centimetres, this causes ions and remaining droplets to travel at hundreds of $\mathrm{m\,s^{-1}}$. Such high velocities pose a transmission issue for large molecules as they gather dozens of electron volts of kinetic energy, not all of which is aligned along the ion path axis, causing difficulties for beam focussing, ion trapping/gating and mass analysis.[53,54] Increased pressures, 10^{-4} - 10^{-2} mbar, in the source region and initial guides help to reduce the axial and radial components of the ion's translational velocity through what was termed collisional cooling,[55] or sometimes collisional focussing/dampening. For Q-ToF instruments the benefits are two fold: firstly as ions become thermalised they become more effectively trapped by electric potentials in quadrupole ion guides leading to greater radial confinement aiding transmission, and secondly, the axial energy becomes dampened meaning that in an orthogonal time-of-flight analyser ions do not *overshoot* the fixed detector. Increasing the pressure is a very effective method of improving the transmission of high mass protein ions,[56,57] though it is not without its limitations; the axial stopping distance scales with $\sqrt[3]{M_w}$ and so simply increasing the pressure higher and/or lengthening the cooling cell may reach practical limitations. Alternative methods less commonly used include the use of retarding electrostatic lenses to slow ions down or ion traps to thermalise ions over time before release.[53]

2.1.3 Quadrupole mass analysers

The ability to accurately select a given species according to its m/z and subsequently perform a selective gas-phase reaction, ion-mobility separation or activation is of huge importance to native mass spectrometry studies and was heavily utilised in this thesis. Quadrupole mass analysers, or filters, are often employed to perform

this role due to their continuous mode of operation and relatively mild pressure requirements; meaning ions can be selected according to their m/z at an early stage of the instrument. Tandem MS instruments capable of MS/MS analyses typically consist of an initial selection using a quadrupole, followed by ion activation or manipulations, and finally, a second often high resolution mass analyser.

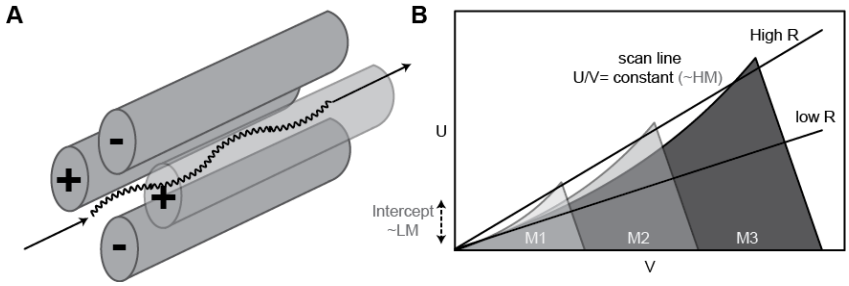


Figure 2.4. Quadrupole mass analyser. A Schematic of a quadrupole mass analyser with a stable ion trajectory. Consists of four parallel cylindrical electrodes with opposite pair-wise potentials applied and switched at a given frequency. **B** A cartoon stability diagram for three different mass ions, three shades of grey, as a function of the RF (V) and DC (U) amplitudes. Scan lines of constant U/V ratios are shown which result in high resolution (R) or low resolution separations. In Waters instrumentation the gradient may be controlled by a "HM" quadrupole parameter and the intercept by "LM".

A quadrupole mass analyser discriminates ions according to their m/z by selectively varying the electric field to allow a stable ion trajectory through the guide. Quadrupoles typically consist of four perfectly parallel cylindrical electrodes with the same voltage applied to opposite rods, in phase, separated by a distance $2r_0$, with adjacent rods out of phase but with the same amplitude, Fig. 2.4A. An ion will migrate towards the electrode of opposite polarity and if the polarity is switched before the ion discharges on contact, the ion will begin to migrate in another direction. The magnitude of the voltage will determine whether a given m/z is able to survive a number of oscillations and exit the guide. The applied electric field constitutes an RF voltage, V, superimposed on to a time independent direct current (DC) voltage, U.[58] By solving the equations of motion for a given m/z and frequency, Mathieu's equations, one can determine the values of U and V for which the ion has a stable trajectory, Fig. 2.4B. If the quadrupole operates entirely in RF-only mode, i.e. $U = 0$, then it acts as an ion guide transmitting a broad m/z range of ions and acting to focus the beam.[55] To act as an m/z filter

U is varied to bring some ions into unstable regions whilst maintaining stability for others; the greatest resolving power can be achieved when U and V are set to be the apex of a given m/z stability region. When defining a mass selection or region of required transmission it is typical to scan over U and V at a fixed ratio, shown in Figure 2.4B as the gradient. The steeper the scan line the narrower the m/z transmission window for a given RF amplitude, achieving a stringent isolation. In these studies the quadrupole in a Synapt Q-ToF allows control of the gradient and intercept of the *scan line* using two parameters termed "LM res" and "HM res" respectively, these are often provided in the methods sections of later chapters.

The maximum transmittable m/z is determined by:[57]

$$M_{max} = \frac{7 \times 10^6 V}{f_2 r_0^2} \tag{2.3}$$

Where, the voltage applied to each rod is $\Theta = \pm(U - V cos(2\pi f t))$, f is the frequency, $2V$ is the peak-to-peak RF amplitude, t is time and r_0 is the distance between paired rods. Practical limitations in instrument design mean only the frequency can be used to allow greater transmission of significantly larger ions. Q-ToF platforms were developed with low RF frequencies, 300 kHz as opposed to ∼1 MHz, to allow the transmission of ions as high as 32,000 m/z,[57,59] allowing for tandem MS strategies to be applied to large bimolecular complexes. Operating at such low frequencies does coincide with a loss in resolution but as modern tandem MS strategies typically utilise a high resolution analyser downstream, the effects are often inconsequential.[7]

2.1.4 Time-of-flight mass analyser

ToF spectrometers accurately measure the time taken for a pulsed ion packet to traverse a low pressure drift region, of known length.[58] Ions are accelerated into the drift region over a short distance using a high potential difference between a *pusher* electrode and a set of conducting grids through which ions can pass. The drift region is electric field-free and under high-vacuum ($\sim 10^{-7}$ mbar) meaning an ion's speed is determined principally by the initial velocity upon entering the drift tube. The accelerating potential is almost entirely converted into kinetic energy,

E_k, with ideally equal gain across all ion species. The m/z can be related to the time taken, t, to reach the detector if the acceleration voltage and effective drift length, L_{eff}, are known for a linear tube.

$$\begin{aligned} E_k = qU_0 = ezU_0 = \frac{1}{2}m\nu^2 \\ m/z = 2eU_0\Big(\frac{t}{L_{eff}}\Big)^2 \end{aligned} \tag{2.4}$$

Where, q is the ion charge, U_0 is the acceleration potential, ν is the ion velocity, e is the charge of an electron, and z is the ion charge state.

Early ToF developement focussed on a linear drift tube in line with the axis of the source beam.[60] In order to measure the flight time, however, ions are required to be pulsed into the drift region as discrete accelerated packets. Continuous ion beams must therefore be either deflected or trapped for the duration of the current acquisition. To avoid this most continuous beams, such as ESI, are rapidly sampled, µs range, using an orthogonal acceleration (oa) ToF.[61] The ToF region is positioned orthogonal to the ion beam and by applying a high accelerating voltage across the ion path ions are injected into the field-free region; ions have a velocity, in the ToF axis, determined by the *pusher* acceleration, whilst maintaining the beam velocity in the orthogonal axis.[62,63]

In principle there is no upper limit to the m/z detectable as one need simply wait for longer times for ions to traverse the drift region. Practical limitations make this difficult however. We previously discussed the pressure dependencies large ions require for efficient transmission in the source region and during ion-manipulation but building an instrument capable of isolating high pressure regions remains challenging. ToF pressures are typically maintained at low 10^{-7} mbar to ensure few, if any, collisions occur preventing ion scattering or fragmentation.[64] Scattering will act to change the trajectory and thus a broadening of arrival times will be detected, lowering the resolution, or even causing the ion not to be detected at all. The mean free path of a native protein, ∼30 kDa, can be calculated to be ∼10 m.

Resolution

The ability to resolve two species in the m/z dimension is of importance to both bottom-up and native strategies. Spectral complexity is commonplace in *omics* workflows particularly where a prior solution separation is limited or when gas-phase fragmentation can result in numerous isotope envelopes. The resolving power of a mass analyser, R, is defined as $R = (m/z)/(\Delta m/z)$ where, $\Delta m/z$ is the peak width, often referenced at half maximum, *FWHM*.

Isotopic resolution is achieved, at half height, when the instrument resolving power is equal to the ion mass, M_{ion}, of the protonated peptide or protein, $R = M_{ion}$; making the isotope patterns of peptides and small proteins readily resolvable for most platforms. Whilst isotopic resolution is rarely required in native MS methodologies, increased instrument resolving power is desirable for characterising low mass binding partners or identifying charge states in heterogeneous or complex samples.[65]

The next consideration is the ability to resolve a charge state series which is crucial in native mass determination. The absolute m/z difference between the most abundant charge state z and z+1 (the next charge state closest in m/z) can be estimated using the relationship between average charge state and molecular weight, page 31. It can be shown, however, that the charge state separation remains greater than 180 m/z even for a 1.5 MDa protein. This corresponds to a resolving power of only ∼300 to theoretically resolve these charge states.

Whilst resolved charge states for an isolated homogeneous protein presents little challenge for modern instrumentation the experimental reality is quite different. Firstly, as the mass of a protein increases, the isotopic distribution broadens resulting in an intrinsic mass width, Δm_{iso}, which has been shown to scale almost exactly with the square root of the molecular weight, $\Delta m_{iso} = 0.063\sqrt{M_w}$.[28] Though as the average charge state, page 31, scales with a greater exponential, 0.533 - 0.550, the intrinsic peak width will narrow at higher m/z and above ∼7 kDa will always be smaller than 1 m/z. Secondly, and more importantly, experimentally observed peak widths are almost certainly greater than these minimum values owing to a lack of

desolvation, leading to the addition of remaining solvent and salt adducts. There is often a strong mass dependency associated with adduct formation and as such the theoretical resolution is very rarely a reality. These effects can be reduced prior to electrospray and sometimes *in vacuo* through gentle activation in prior collision cells and the source region with vastly different instrument dependencies.[66]

For a ToF instrument it can be readily shown that resolution is theoretically independent of m/z: $R = t/2\Delta t = |U_0/\Delta U_0|$.[67] As a result, resolution is most commonly determined by the width in time of an ion packet as it strikes the detector. Variations in the acceleration voltage applied, ΔU_0, are relatively minor in modern instrumentation due to accurate and stable electronics, though electric field inhomogeneities do still persist.

One of the greatest impacts on the time dispersion is caused during the initial acceleration; as ions are not perfectly focussed, off-axis displacement will result in a distribution of received electric potential and ultimately orthogonal injection velocity. A second delayed acceleration pulse can help limit these effects—delayed extraction—yet an undesirable distribution of kinetic energies still exist[60] for ions of the same m/z and as a result most high resolution ToF analysers utilise an ion mirror or reflectron.[65] An electrostatic mirror, whose reflection plane is parallel to the ion beam, not only improves the ability to discriminate between different m/z species by extending the path length inside the spectrometer but importantly helps compensate for variability in the initial drift velocity. Ions with a greater velocity are able to penetrate the reflectron field to a greater depth and thus extending the flight path and time in a compensatory manner. If the reflectron field is fully optimised ions can be kinetically focussed to within 5 % of the maximum velocity.[68]

Additional sources acting to decrease the resolution that are broadly independent of the reflectron include increased pressures and grid perturbations. To achieve a field-free region a conducting grid or mesh is placed at the start of the flight tube preventing significant penetration of the high acceleration field. However, the grids can produce inhomogeneities in the boundary electric field potential which may

effect the trajectory and ultimately the arrival time of the ions.[61,69] Other impacts on the resolution may arise from mechanical defects, detector timing limitations and space-charge effects on the ion cloud.[68]

Efforts to increase ToF resolution have mainly focussed around improving the effects discussed above. Novel applications which act to increase the path length by an order of magnitude include the development of multi-mirror flight tubes[70,71] and an oscillating orbital path around a central spindle, not too dissimilar to an orbitrap,[72] though these applications have not been realised fully.

2.1.5 Orbitrap analysers

The last decade has seen the rise of an alternative mass analyser for native mass spectrometry studies, the orbitrap.[73] An orbitrap measures the ion current produced by a trapped ion processing parallel to a central spindle electrode whilst orbiting radially.[74,75] The frequency of such oscillations can be related to the *m/z* of the ion and Fourier transformed to provide a mass spectrum. To obtain sufficient data in frequency space requires long acquisitions, 10-100 ms, which also necessitates very low pressures to prevent scattering. Despite this the transient signal decays over time in a *m/z* dependent manner though recent commercial platforms (QExactive, Thermo Scientific)[76] have produced well-resolved spectra as high as 70,000 *m/z*. The high resolutions afforded, $10^4 - 10^5$, from an orbitrap detector are due to very low pressures and are not impacted by spatial dispersion of ion packets which allow for significantly more complex spectra to be analysed. Whilst orbitraps have a $1/\sqrt{m/z}$ dependence with mass-to-charge, resolution for high *m/z* ions remains impressively high, ∼530 at 70,000 *m/z*.[76] The improved resolution on a QExactive platform, relative to commercial ToF instruments, is not simply due to the instrument analyser but also the degree of desolvation observed prior to analysis.[77,78] Whilst this can often mean well-resolved spectra the energy of the system should be treated with some caution when studying protein-ligand interactions; loss of cationated peaks simplifies spectra but can result in gas-phase ligand loss also. The relatively long acquisition times required for large ions makes orbitraps not immediately amenable

to be coupled to IM, a further comment is made in Section 2.1.6, and as such these instruments were not utilised for much of the data presented in this thesis.

2.1.6 Ion mobility spectrometry

Ion mobility spectrometry provides a means of separating ions, on a millisecond time scale, according to size and shape. Ions migrate through a pressurised cell (~0.1 mbar) under the influence of a relatively weak electric field and are impeded by tens of thousands of non-activating collisions.[79] The mobility, K, of an ion is determined by the speed in which the species drifts through the cell under a given electric field with compact structures travelling faster than more extended ions of the same charge Fig. 2.5A & B. In a linear drift cell instrument a constant electric field is applied over the cell length, L_d, allowing ions to traverse the cell at a near constant velocity, which can be determined from the measured drift time, t_d, as described by Equation 2.5.

$$K = \frac{\nu_d}{E} = \frac{L_d}{t_d.E} \tag{2.5}$$

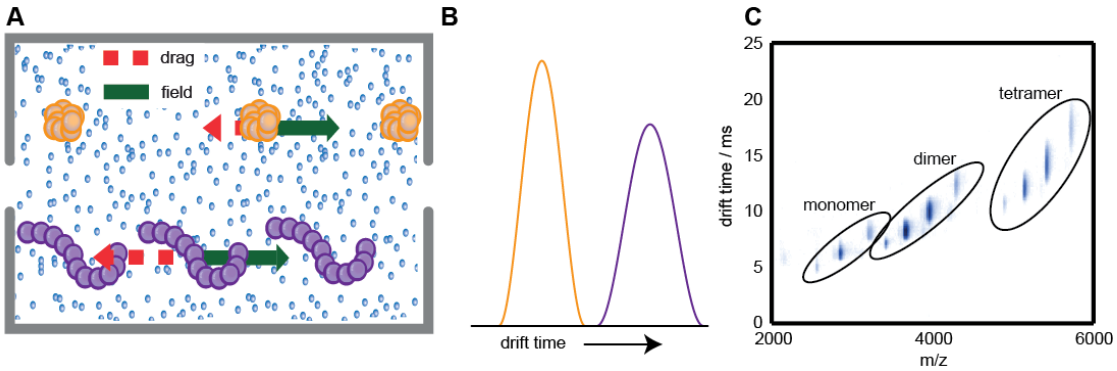


Figure 2.5. Ion mobility spectrometry of protein complexes. A Ions are gated and released into a mobility cell containing a relatively high pressure of a neutral gas. In a linear drift cell, ions migrate through under the presence of a constant electric field and are impeded through collisions with the bath gas. For the same charge, more extended ions undergo a greater number of collisions resulting in increased *drag* and emerge at the detector later (**B**). **C** Ion mobility provides an orthogonal separation technique to a ToF mass analysis resulting in a 2D spectrum of mobility and m/z. Different stoichiometries of concanavalin A can be seen to cluster having been separated by mobility. Panel **A** was modified from a figure provided by Justin Benesch.

The number of collisions will be heavily determined by the collision cross section (CCS) of the ion and the associated buffer gas. The relationship between

ion mobility and the CCS is given by the Mason-Schamp equation,[80] Eq. 2.6. The drift time is also proportional to the charge state of the ion as this determines the forward migration force; structural differences between charge states may also be present though this component is reflected within the CCS term.

$$K = \frac{3}{16}\sqrt{\frac{2\pi}{\mu k_B T}}.\frac{q}{N.\Omega} \approx \frac{3}{16}\sqrt{\frac{2\pi k_B T}{\mu}}.\frac{q}{p.\Omega} \tag{2.6}$$

Where Ω is the CCS, k_B is the Boltzmann constant, T the temperature, q the charge, μ the reduced mass of the ion and gas, N is the number density of the gas, which can be approximated to the pressure, p, through the ideal gas law. Increasing the pressure will result in a greater number of collisions decreasing the K, whilst increasing the temperature will result in a greater mean-free path and will act to increase K. It should be noted, however, that the CCS of large proteins is known to have a temperature dependence owing to the thermal velocity of the gas and the resulting long-range interaction time, Fig 2.7B.[79,81]

In order to experimentally measure the drift time, ions are trapped behind a raised gating lens before being released into the drift cell and detected. ToF mass analysers lend themselves as appropriate detectors for a mobility separation as ions can be pulsed into the ToF tube as frequently as every few hundred microseconds, even for very large ions. As such, IM separated ion packets can be pulsed into the mass analyser for detection with the mobilogram consisting of binned[i] data sets defined by the frequency of the ToF pusher. IM drift cells have long been coupled to Q-ToF instruments and can obtain a two dimensional IMS-MS analysis which can provide spectral deconvolution and importantly for structural studies can directly report on the CCS of a given m/z species. Figure 2.5C shows a 2D mobilogram-mass spectrum for the protein complex, concanavalin A. Monomers, dimers and tetramers are all observed and are separated according to both IM and m/z, forming clusters for each stoichiometry.

[i]The ion mobility instrument used in this thesis, Waters Synapt G1 or G2, has a mobility scan cycle defined as 200 pusher bins.[82]

Travelling wave ion mobility

The use of travelling wave ion mobility allows for greater IM resolving power and transmission over drift time (DT) IM-MS for the same drift distance.[83] On a Waters Synapt platform, electrostatic pulses move through a stacked ring RF ion guide, Fig. 2.6A, at a tunable velocity creating a moving or *travelling* wave.[84] On a first generation instrument this is generated by raising the voltage on two adjacent lenses, separated by 1.5 mm, and moving the voltage pair in the direction of the beam. At the centre of the stacked ring axis the amplitude of the wave is ∼0.63 of the applied voltage and the waveform can be treated as a near sin^2 potential,[85] plotted in Figure 2.6B. As such the electric field, E, is not constant as a function of time nor axial distance; the maximum electric field for a wave height of 10 V is $E_{max} \approx 21$ V/cm, which is higher than for a DT-IM experiment ($E_{max} < 11$) though ions are exposed to these fields for a much shorter period of time. Some studies have shown that these elevated electric fields are sufficient to cause activation of small molecule systems,[86,87] though the local activation effects on large protein ions are limited through energy dispersion.

There is no simple relationship between the drift time in a TW-IM device and the CCS, although empirical,[88] and now mathematical,[89] relationships do exist. As such, TW-IM drift times are often calibrated against standards, of known CCS.[90] TW-IM exploits the difference in ion mobility as it traverses the front and rear of the wave.[91] If assumptions are made that $K << s/E_{max}$, where s is the wave speed and E_{max} is the maximum electric field, then the drift time can be approximated as follows:

$$t_d \approx \frac{sL_{DT}}{(K\bar{E})^2} \tag{2.7}$$

Where, $\bar{E}$ is the average electric field experienced by the ion. Note, there is a squared dependency on mobility which will also mean the charge state dependency will be a squared.

When the mobility of the ion is high, ions will simply *surf*[83] the wavefront and be pushed through according to the wave velocity, this is particularly true

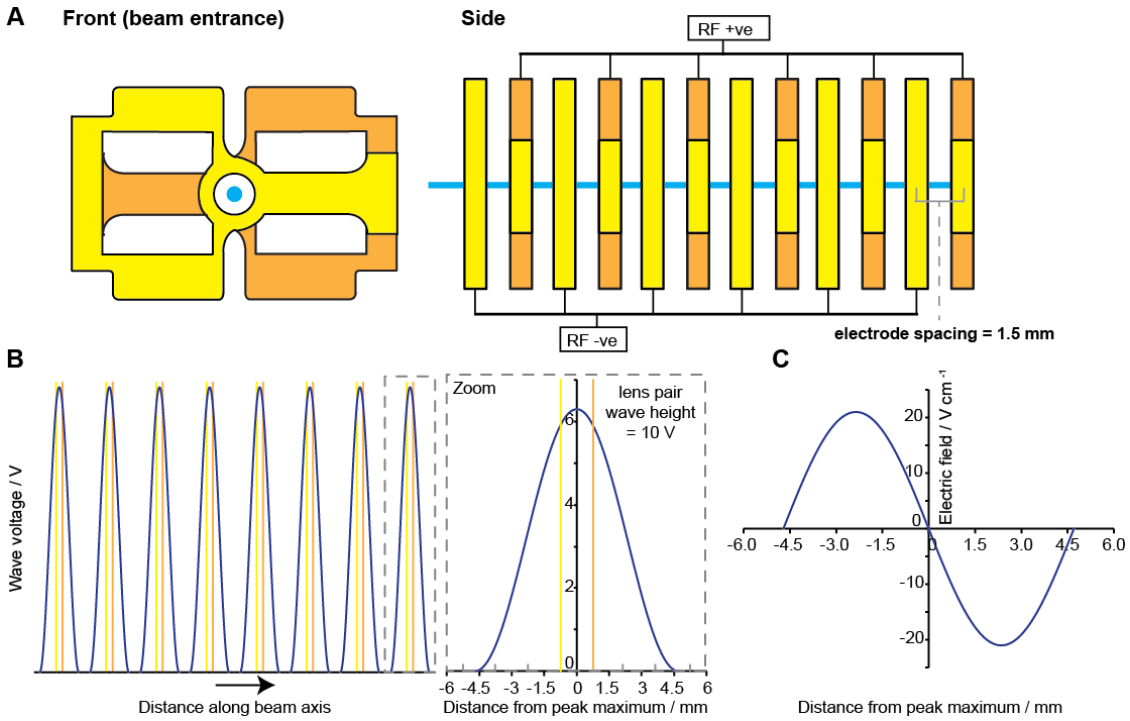


Figure 2.6. Stacked ring travelling wave ion guide. **A** Drawing, not to scale, of the stacked ring lenses on a Waters Synapt G1 ion guide. The beam axis is shown in blue with alternating lenses shown in yellow and orange. A schematic side view is also shown demonstrating the alternating application of a radio frequency confining electric field. Lenses are separated by 1.5 mm, centre to centre. **B** A plot of the waveform of a Synapt G1 travelling wave. Pulses are applied by raising the wave height on a pair of lenses and repeat every 6 lenses. Applied voltages are then moved along in the beam direction one plate at a time. The waveform of the wave adopts a near sin^2 function with the voltage at the ring centre plotted against distance, for an applied wave height of 10 V (zoom). The raised lenses are shown in yellow and orange. **C** The electric field can be estimated through differentiation of the potential with respect to distance. A maximum electric field of $\sim$21 V cm^{-1} is achieved for a wave height of 10 V as expected.[85] Lens drawings were made from CAD drawings provided by Todd Mize.

for high wave heights and whilst useful in aiding transmission or maintaining a prior mobility separation[ii], a poor separation will be obtained. It is desirable that ions are able to *roll-over* the wave front for repeated pulses to allow discrimination between ions of different mobility.

Interpreting the CCS

Ion mobility in its simplest analytical form allows one to reduce the spectral complexity of the sample by separating according to mobility in another dimension. What is most interesting, however, for structural biology is the ability to determine

[ii]High wave heights are used for this very reason in the transfer ion guide to preserve the prior mobility separation before detection

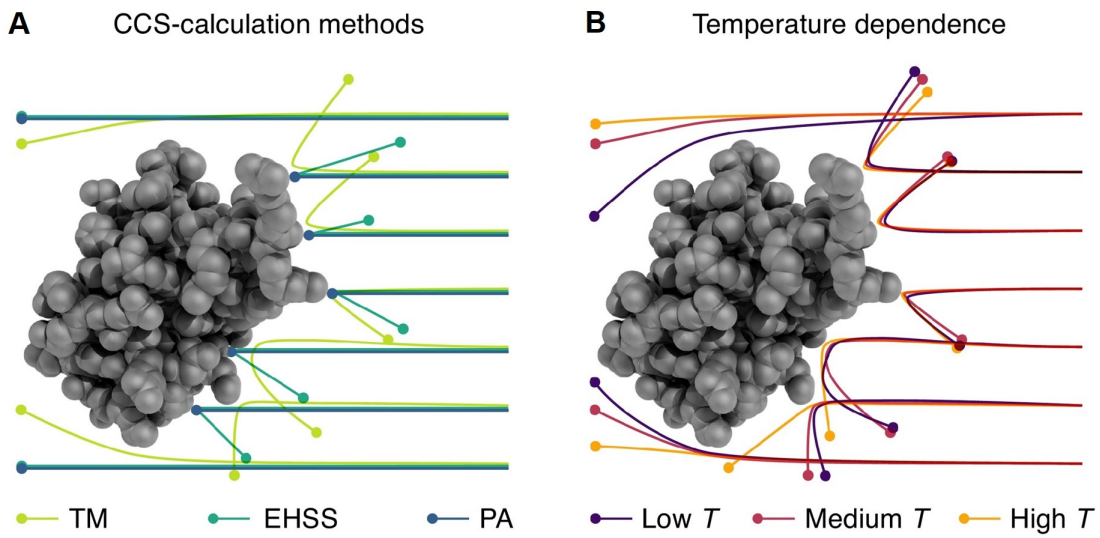


Figure 2.7. Collision cross section calculation methods. A A depiction of the treatment of different calculation methods upon collision with a gas partner. The trajectory method (TM) calculates the trajectory for each collision including long-range interactions that effect the CCS. Elastic hard sphere scattering (EHSS) and the projection approximation (PA) ignore such interactions and only consider *direct* collisions. **B** There is a temperature dependence on the collision cross section due to the long-range interactions present. Lower gas temperatures, with a lower velocity, are more influenced by the passing ion. These grazing collisions account for decreased CCS values with increased temperatures. This figure was published in Gabelica & Marklund 2018.[79] Reprinted from Fundamentals of ion mobility spectrometry, 42, V. Gabelica & E. Marklund, Current Opinion in Chemical Biology, 51-59, Copyright (2017), with permission from Elsevier.

the CCS and relate this back to relevant solution structures. CCS values can be used to verify a folded structure where a structure is known or can be used to aid modelling of unknown assemblies.[92–94] Many methods exist to predict the CCS of a 3D structure with a compromise made between accuracy and the computational cost.[95] These methods of calculation are illustrated in Figure 2.7. The most accurate method is to calculate the trajectory of each colliding gas partner, trajectory method (TM), including those influenced by long range interactions capable of perturbing the momentum but without any scattering, though this is computationally expensive and for large systems becomes unfeasible when many candidate structures and orientations must be probed. Using this approach one is even able to account for temperature dependencies observed in long-range *grazing* interactions where slower gas molecules are influenced to a greater extent.[79,81] In ever large systems the peripheral long-range interactions diminish as a proportion of all collisions

making alternative approximations appropriate for large protein ions. Collisions can be treated as elastic hard spheres scattering (EHSS) which neglect any long range interactions and multiple collisions. This can be simplified further by only determining whether ions have a direct hit or not, and calculating the rotationally projected area of the structure, projection approximation (PA). Whilst the PA method is known to be inaccurate for proteins when compared to crystal structures a correction can be applied, $CCS_{experimental} = 1.14\ CCS_{PA}$.[96]

Alternative IM devices

All of these instruments utilise relatively high gas pressures, high number density *N*, with a low electric field E, resulting in what is termed a low field regime with a small E/N ratio. Under such conditions limited activation is observed and electric fields are not sufficient to cause dipole alignment,[97] resulting in a rotationally averaged measure. Alternative methods of IM[98] that exploit the high field regime include high-field asymmetric IM (FAIMS) or differential IM (DMS), though due to the activating nature these techniques tend not to be compatible with native MS. DT-IM and TW-IM operate using a *stagnant* gas with no net flow on any axis. Trapped ion IM operates on the inverse principle as there is a gas flow away from the detector end of the IM device and ions are held by a trapping field; only ions with a particular charge and mobility will remain trapped whilst others will migrate away.[99–103]

The resolution of an IM separation can be improved by simply extending the length of the drift tube, so long as transmission and gas-phase stability will allow. Extended linear drift tubes with multi-stage IM have previously been developed by the Clemmer group[104] in addition to a cyclical drift tube, which in principle has infinite length.[105] Orthogonal cyclic drift devices are also under development by Waters Corp.[106,107] with additional multi-stage IM capabilities. One of the most promising developments has been the use of miniaturised structures for lossless ion manipulation (SLIM) devices.[108] Efforts have also enabled IMS instruments to be coupled to high resolution orbitraps through the use of novel ion gating[109] whilst

others have used the intrinsic decay of the image current to determine the CCS of simple systems,[110] though these methods are far from appropriate for most complex native MS analyses.

2.2 Gas-phase activation

When ions are accelerated into a collision cell they undergo many thousands of inelastic collisions with the neutral bath gas molecules. As they collide a proportion of the ion's kinetic energy is transferred to internal energy through excitation of the vibrational and rotational modes of the macromolecule, occurring over 10s of microseconds.[111,112] The slow nature of collisional heating allows the energy to redistribute across the considerable number of degrees of freedom that exist in a large biomolecule. The gradual build up of internal energy may eventually lead to the breaking of non-covalent interactions, complex dissociation, and ultimately fragmentation. The maximum amount of internal energy gained is given by:[113]

$$\Delta E_{int} = E_k \left[1 - \frac{M_i^2 + M_g^2}{(M_i + M_g)^2} \right] \tag{2.8}$$

where the gain in internal energy, ΔE_{int}, and the kinetic energy, E_k, prior to the collision are in the laboratory frame, M_i and M_g correspond to the molecular weights of the ion and gas collision partner respectively.[54] These provide approximate levels of translational energy conversion but do not address the variable levels of collision elasticity depending on the identity of each collision partner nor the degrees of freedom of each species. It is evident that as the molecular weight becomes very large, $M_i >> M_g$, the scaling term of the kinetic energy becomes very small. Assuming argon to be the collision partner, a 100 Da molecule will lose up to 40 % of its kinetic energy whilst anything above 8 kDa will transfer at most 1 %. Increasing the mass of the gas collision partner, typically Ar, Xe or SF_6, will also increase the amount of internal energy gained[iii].

[iii]This is always true for proteins but for small molecules this only holds up to the point at which the mass of the gas is equal to the ion mass.

Alternative methods of activation that have found utility within native mass spectrometry include the use of ion-ion collisions (electron-transfer, ETD, or electron capture, ECD), photo-activation (ultraviolet, UV, or infra-red, IR) and single collisions with a surface (SID). IRPD, also a slow heating method, has been shown to display similar dissociation properties to CID.[114] Of significant interest to the native community are those capable of rapid activation, pico-femto second, for probing top-down structure or maintaining labile moieties or labels, UVPD and ET(D/C).[115] SID utilises an immovable collision partner to transfer kinetic energy but unlike CID the efficiency of energy transfer is significantly greater over a single collision for large ions. This rapid, somewhat non-ergodic, process has a tendency to cause ions to dissociate symmetrically at an interface.[116]

2.2.1 Dissociation of protein complexes

Whilst much of native MS relies on maintaining non-covalent interactions, dissociation of a protein complex can be hugely advantageous in deciphering the components in a subunit assembly.[117] A protein complex may be activated using an acceleration potential into a collision cell to induce the subunit or ligand dissociation aiding accurate stoichiometric identification, and, determination of the subunit connectivity and quaternary topology.[118,119] The utility of such a strategy has been shown through my own published work, from identifying endogenous ligand binding stoichiometries,[120] validation of inhibitor drug affinity[121] through to the more complex example of quantification of polydisperse oligomeric distributions.[122]

As the internal energy of a protein or ligand complex increases, through the collisional processes described previously, it will reach a point where the Coulombic repulsions of each subunit overcome the remaining interfacial interactions holding the complex together. This results in the charged loss of a subunit or ligand (though ligands are commonly lost as neutrals). During CID of large protein complexes charge is very rarely symmetrically partitioned with respect to mass,[35] this idea is considered in much greater detail in Chapter 5, with the ejected subunit leaving with as much as 50 % of the available charge. This is demonstrated for the monodisperse

oligomeric, 24mer, small heat shock protein, *Mj*-HSP 16.5, in Figure 2.8. As the energy is increased, subunits are sequentially lost taking a disproportionate amount of charge, this has the fortunate consequence of severely reducing the charge carried by the *stripped* oligomers causing them to be separated to a greater extent in *m/z*.

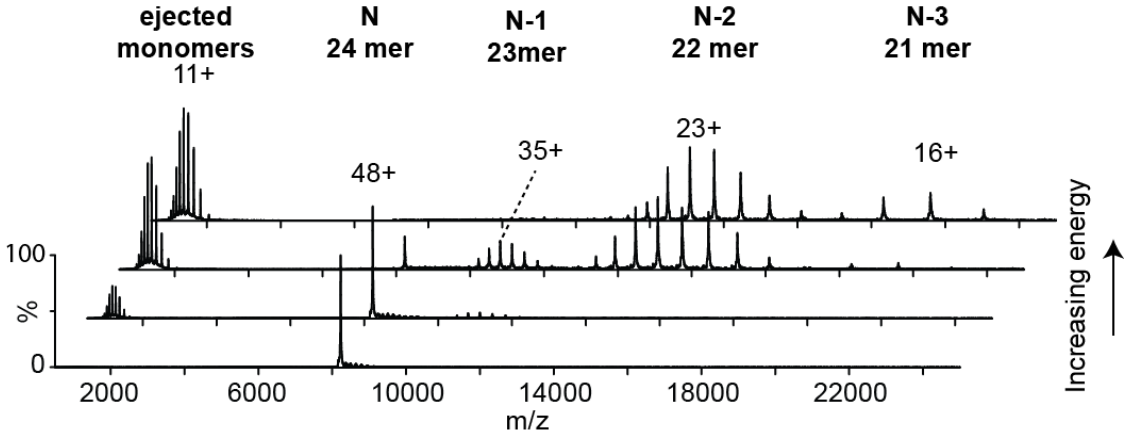


Figure 2.8. Collision induced dissociation of protein complexes. Mass spectra showing the collisional activation of the quadrupole isolated 48+ 24mer HSP16.5. Increasing the acceleration energy into the collision cell sees the loss of highly charged monomers at low *m/z* and N-1, 23mer, at higher *m/z*. At higher collision voltages sequential loss of subunits sees the N-2 and finally N-3 stripped oligomers. Charge is lost in a highly asymmetric manner with the monomers leaving with almost one quarter of the initial charge.

The commonly held mechanism for this behaviour is that as native protein ions become activated the interactions holding the solution-like fold together become weakened and a local unfolding event can occur.[123] Coulombic repulsion of mobile charges will try to minimise the electrostatic energies by distancing charge, when a protein unfolds and the surface area increases this provides sites for charge to migrate to and reduce the Coulombic strain.[124,125] This process is self-reinforcing as migrating charge causes additional unfolding through localised repulsions. This idea of collisional activation precipitating an increase in surface area, which is further propagated by charge migration is supported by observations that the charge density of the stripped subunit is comparable to that of the intact *stripped* oligomer.[126] Molecular dynamics, MD, simulations also suggest proteins undergo propagated Coulombic unfolding as they are ejected from the intact oligomer. MD simulations, activated by elevated temperatures, show that positively charged subunits are able to unravel, typically from the free N-terminus, leaving enriched with charge.[127]

Some of these studies even account for charge migration with a pseudo-mobile charge model.[128]

It should be noted that there is still a degree of debate as to the mechanism that explains asymmetric charging. Historically, some have argued that a protein behaves somewhat like a charged liquid droplet undergoing asymmetric fission.[129] Treating a desolvated protein as a droplet is likely to be a gross oversimplification and cannot fully explain observed data, though the authors noted that partial unfolding of a subunit could account for some of the asymmetry, bearing some similarities with the Coulombically driven unfolding and ejection described above. One final mechanism proposed suggests that salt-bridges located at a subunit interface asymmetrically cleave, with one subunit leaving with excess charge, salt bridge rearrangement (SaBRe). Whilst this mechanism is conceivable for small oligomers it struggles to explain the magnitude of charge migration associated with large numbers of subunits as with small heat-shock proteins.[130]

2.3 Solution hydrogen/deuterium exchange MS

Hydrogen/deuterium exchange (HDX) is a reversible method of chemical labelling in which the exchange of hydrogens for deuterium can be monitored as a function of time or under different environmental conditions. Exchange of labile sites on a protein is sensitive to the hydrogen bonding network and steric exclusion of the bulk solvent by the protein structure.[131] The use of isotopic exchange to probe the solvent accessibility of biomolecules was realised as early as the 1950s by Hvidt and Linderstrøm-Lang to probe polypeptide structure.[132] In the absence of structure, a random coil, hydrogens associated with the side-chain heteroatoms exchange rapidly, millisecond time-scale,[133] whilst backbone amide exchange is slower at a second timescale. As amide hydrogens are universal to all residues except proline, and the newly formed N-terminus in a cleaved peptide product, HDX provides a non-selective means of reporting on the entire structure. Amide hydrogens are particularly sensitive to the local structural environment, specifically the involvement

of hydrogen bonding interactions associated with secondary and tertiary structure, which can slow the rate of exchange to minutes or even hours.[134] As such, following the exchange rate of these sites allows one to probe protein conformational dynamics and structural interactions.[135,136]

HDX is both base and acid catalysed and proceeds via a resonance structure over the amide; side-chains associated with the residue can either help stabilise the resonant structure or hinder it, altering the intrinsic rate of exchange.[137] There is a strong pH dependence on exchange due to catalysis (Eq. 2.9) and results in a chevron shaped exchange rate curve when plotted against pH, Fig. 2.9A.

$$k_{int} = k_{acid}[\mathrm{H^+}] + k_{base}[\mathrm{OH^-}] + k_W \tag{2.9}$$

where, $k_{int}, k_{acid}, k_{base}$ and k_W correspond to the total intrinsic rate of HDX and the component rates for the acid, base and water catalysed reactions, respectively.[65] In this plot each line corresponds to the theoretical exchange for each individual amino acid, X, in a poly-alanine peptide, AAXAA, based on rates and theory derived by Bai *et al.*[138] There can be as much as two orders of magnitude difference depending on the residue, with histidine the fastest exchanging site and isoleucine the slowest. A minimum is observed for each between pH 2.4 and 3.4 with those experiencing faster exchange minima closer to the former. Additionally reducing the temperature from 30 - 0 °C results in a wholesale decrease in the rate of exchange by almost two orders of magnitude, Fig. 2.9C. These two properties allow the H/D reaction to be effectively quenched at a pH minimum and at 0 °C.

Mass spectrometry and nuclear magnetic resonance (NMR) can both be used to monitor the rate of exchange of labile hydrogens after exposure to a deuterated solvent, though this thesis focuses exclusively on the role of HDX-MS.[139–141] The exchange patterns of HDX-MS offer the most structural information when the deuterium incorporation is localised to just a few residues, typically achieved through prior digestion. A typical HDX-MS experiment, Fig. 2.10A, consists of the dilution of a protein sample from its native H_2O based buffer into D_2O of the same

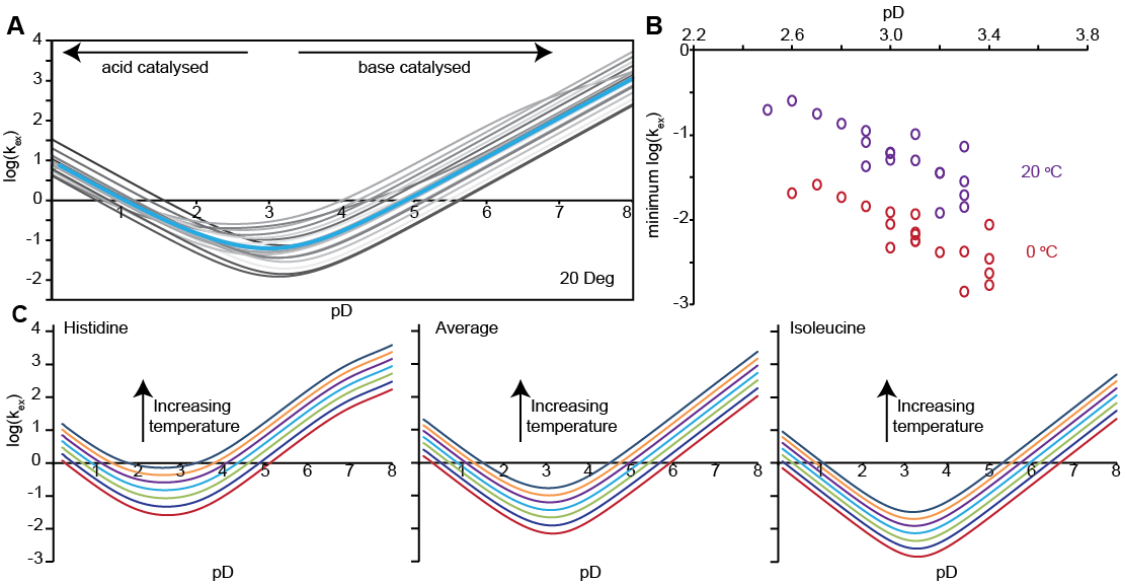


Figure 2.9. Theoretical rate of intrinsic H/D exchange. **A** Theoretical H/D exchange curves as a function of pD for the peptide AAXAA, where X is a variable amino acid. Each line corresponds to a different amino acid with the average shown in blue. **B** The minimum rate is plotted against the corresponding pD value for each residue. The data is shown at 0 °C (red) and 20 °C (purple). **C** The temperature dependence on exchange is shown for the fastest exchanging residue, histidine, the slowest, isoleucine, and the average. 0 °C is shown as a red curve with each line above corresponding to 5 °C increments until 30 °C in dark blue. The inflection observed for histidine corresponds occurs around the pKa of histidine. Data was plotted using the equations described in Bai *et al.*[138]

composition resulting in a solution >90 % D[iv]. The exposure time is typically varied from a few seconds to over an hour, with day-long exposures sometimes desirable, followed by a quenching step at 0 °C and a partial dilution into a H_2O based low pH (∼2.3 - 2.5) buffer.[139] Low pH is often sufficient to cause unfolding though chaotropic salts or reducing agents may be necessary.[143] In this highly unfolded state exchange rates increase and so back-exchange with the H_2O is inevitable, limiting this exchange with the quench conditions preserves the isotopic incorporation for analysis.[144,145] The back-exchange of the labile side chains is so great, even under quench conditions, that all deuterated labels are lost from these sites; this can be seen as beneficial as it removes the convolution of the back-bone and side-chain exchange regimes. If localisation of uptake is required, an enzymatic digestion step follows[v], at a slightly elevated temperature for enzyme activity, before downstream

[iv] Some studies have shown the benefit of limiting the D % to <50 % to restrain isotope envelopes, though this is not typical practice.[142]

[v] Pepsin is typically used owing to its tolerance to low pH conditions and non-specificity of

desalting and a reverse-phase separation of the peptide mixture under quench-like conditions. Peptide products are then rapidly analysed using a mass spectrometer[vi] to determine the deuterium incorporation. Commercial platforms such as that used in this thesis enable all stages of liquid handling and separations to be performed in one on-line workflow, the schematic for the handling unit is shown in Figure 2.10B.

2.3.1 HDX kinetics

The kinetics of exchange can often be reduced to a simple two-state kinetic model with an equilibrium between exchange competent and incompetent sites,[65] Eq. 2.10. Regions highly protected in the native structure are rendered exchange incompetent and rely on a either localised *unfolding/breathing* of the structure or large conformational changes to populate an exchange competent state.[148,149]

$$\mathrm{NH(incompetent)} \underset{k_{\mathrm{close}}}{\overset{k_{\mathrm{open}}}{\rightleftharpoons}} \mathrm{NH(competent)} \xrightarrow{k_{\mathrm{int}}} \mathrm{ND(competent)} \rightleftharpoons \mathrm{ND(incompetent)} \tag{2.10}$$

Where k_{open}, k_{close} and k_{int} correspond to the rate constants for opening, closing and the intrinsic rate as defined above. An exchange event can only occur if the probed region escapes the potential well of the native state and populates a more disordered minima, from where exchange can proceed according to the intrinsic rates described above. For most proteins under near-native conditions the unfolded state is only ever transiently populated as the refolding rate, k_{close}, tends to be fast and certainly faster than the rate of exchange, $k_{close} >> k_{int}$, meaning multiple unfolding events must occur for a single exchange event to take place.[65] For regions displaying this behaviour, termed EX2 exchange kinetics, the rate of HDX is given by:

$$k_{HDX} = k_{open}(k_{int}/k_{close}) \tag{2.11}$$

cleavage. A small number of studies utilise alternative proteases with high success.[146]

[vi]The main requirement for the spectrometer is that it is capable of quantitatively determining the isotope distribution over what is often a rapid chromatographic separation. For this reason ToF instruments have historically been used with some arguing that distorted isotope patterns can occur in trapping instruments due to isotopic beating and resonance effects.[147]

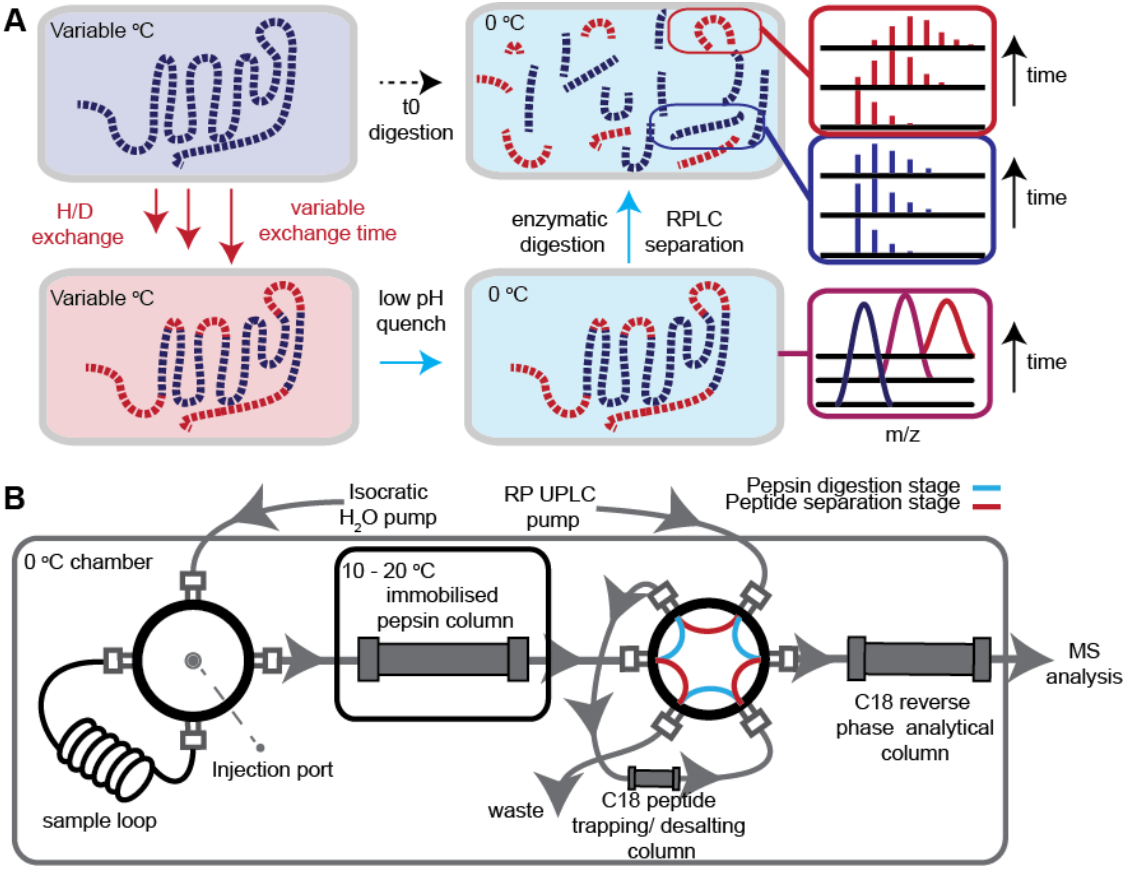


Figure 2.10. Solution HDX-MS workflow. A Schematic of a continuous labelling HDX-MS experiment. Protein in a native H_2O buffer is diluted into a D_2O buffer of the same composition. The sample is exposed to D_2O for a variable time for exchange to occur before being quenched in a low pH H_2O buffer at low temperature, 0 °C. At this point a global measure of uptake could be obtained by taking a mass spectrum of the intact protein. Uptake can be localised by performing an enzymatic digestion and rapidly separating the products. Spectra can then be obtained for each peptide. **B** A schematic of a Waters HDX platform on-line digestion unit. The intact protein is loaded into a sample loop at low temperature before passing through a column containing immobilised pepsin (10 - 20 °C). Peptidic products are collected on a C18 trapping column and desalted. A switching valve changes the flow direction and washes the trapping column using a reverse phase gradient. Peptides are then separated using the gradient on a C18 analytical column before being analysed through ESI-MS.

If, however, there is a significant return barrier between the unfolded state and the native state, then amides remain exposed for much longer, increasing the probability of exchange after a single unfolding event, $k_{close} << k_{int}$. This is termed EX1 kinetics and the rate of HDX is simply the rate of opening/unfolding:

$$k_{HDX} = k_{open} \tag{2.12}$$

These two kinetic exchange regimes result in differences in the isotopic spectral distributions, Fig. 2.11. EX2 type peptides see the isotope envelope gradually

move to towards higher mass with time, reflecting the gradual population of exchange competent states. EX1 kinetics results in two distinct isotope envelopes corresponding to slow and fast exchanging populations, with the relative intensities reflecting the number of molecules that have accessed the open state. The abundance of the closed state spectra decreases over time as it populates the open state and becomes labelled, it may of course close and reopen but labelling will have already been incorporated after the initial event. Two-state kinetics are not a perfect model for protein dynamics as they treat exchange as a binary event and neglect intermediate motions and rate regimes (i.e. $k_{close} \approx k_{int}$)[150] but do help describe much of the observed behaviour.

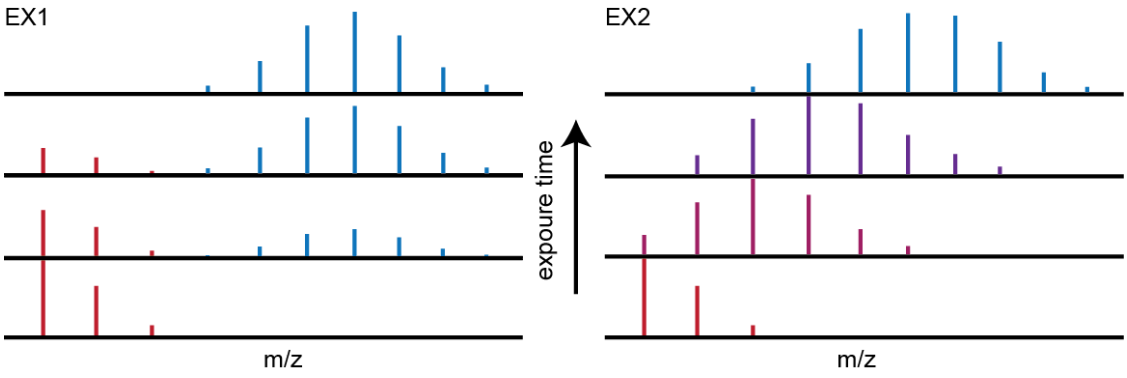


Figure 2.11. HDX-MS kinetics The evolution of peptide HDX-MS spectra as a function of exposure time. EX1 kinetics (left) observe two distinct states corresponding to an "open" exchange competent state and a "closed" exchange incompetent state. The rate of depletion in the closed state is equal to the rate of populating the open state. EX2 spectra are more typical for a native protein and see the gradual shift in spectra towards higher m/z with exposure time. See text for further details.

The continuous labelling method described above, over *long* time-scales, are by far the most commonly exploited for protein structural studies. Pulsed labelling can provide a means of monitoring structural changes after an external stimulus by introducing a short labelling and quenching step.[151] This can be readily employed using millisecond labelling in microfluidic devices, which also have the added advantage of negating back exchange.[152] Bottom-up strategies are an effective means of localising uptake to short peptide regions but the desire for residue level localisation has led to the development of middle-down and top-down strategies. Due to the lability of the D label gas-phase fragmentation methods must be sufficiently

fast to not induce deuteron migration/scrambling. Fast activation methods such as ETD have been proven to fulfil this role,[153,154] though its use is still limited. Photoactivation using UVPD has also been developed in recent years and shows greater promise owing to its increased efficiency and coverage, particularly in top-down methodologies.[155,156]

2.4 Gas-phase hydrogen/deuterium exchange

Since the inception of electrospray and the first observations of intact proteins in the gas-phase, researchers have used labelling strategies in an attempt understand gas-phase structure. As the previous section describes, exchange of labile hydrogens on a protein's surface can help inform on local interactions and accessibility to the labelling reagent. Many studies have performed a gas-phase H/D exchange (gpHDX) reaction inside a spectrometer and followed the global deuterium incorporation by directly measuring the intact mass.[157] As early as 1992, deuterated vapours of D_2O and ND_3 were utilised to probe proteins[158] and peptides[159] respectively, and these have remained the most commonly used reagents since.

Much of the previous work have sought to understand the underlying mechanism of exchange in the gas-phase by focussing on peptide labelling. Exchange occurs at the labile hydrogen sites associated with the side-chain heteroatoms; the N/C-termini may also undergo rapid exchange, Fig. 2.12A. Backbone amides have been shown to be very slowly exchanging in small peptide systems.[160] The exchange mechanisms proposed divide gpHDX into two camps, those with a high gas-phase basicity (GPB) such as ND_3 and those with a lower GPB including D_2O and MeOD.[157] Important work by Beauchamp and co-workers[161,162] studying the reactivity of polyglycine oligomers towards D_2O and ND_3 led to the postulation that D_2O proceeds via a bridging "relay" mechanism, whilst ND_3 progresses through the "onium" mechanism, Fig. 2.12B & C, and experimental and computational studies have supported these findings since.[157,163] Both mechanisms rely on the initial interaction with a positively charged site and formation of an encounter complex.[162] The relay

mechanism involves two sites, bridged by the exchanging D_2O; one of the sites is a proton accepter and the other is a charged proton donor. In a single concerted reaction, the proton accepter abstracts a deuterium from the reagent whilst the reagent abstracts a proton from the charged donor. After the D_2O reagent, now HDO, is lost from the protein surface, intramolecular migration of the deuteron can occur re-charging the most basic site. Most studies involving D_2O have utilised ion trap spectrometers or long drift tubes to store ions on the order of seconds.[164,165]

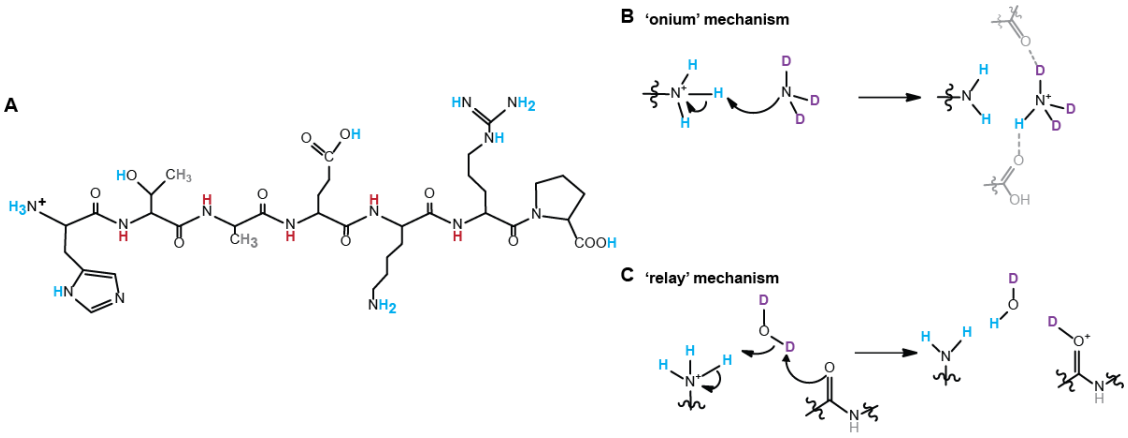


Figure 2.12. Gas-phase HDX reactivity. **A** Labile hydrogens prone to gas-phase HDX on an arbitrary peptide sequence. Labile hydrogens are shown in blue associated with the N/C-termini, basic residues (R, K, H), acidic residues (D & E) and other heteroatom sites (S, T, C, Y, N, Q). Amide hydrogens (red) are slow exchanging in the gas-phase and rarely observed. Aliphatic hydrogens are shown in grey and do not exchange. **B** The "onium" mechanism is proposed for basic deuterated exchange partners such as ND_3 used in these studies. Direct interaction and exchange between ND_3 and a positively charged site. The ammonium intermediate formed can then be stabilised by neighbouring electronegative sites, though not required. The reverse abstraction of a deuteron will then follow. **C** The "relay" mechanism is the accepted as the mode of exchange for D_2O, and alcohols, requiring a "bridging" interaction. The D_2O interacts with both a positively charged site and a basic electronegative site. Exchange occurs in a concerted step as depicted. After the loss of the exchange partner deuteron migration can occur returning charge to the more basic site.

Protein and peptide reactions with ND_3 progress via a different mechanism owing to the increased basicity of the reagent and occurs on a much faster millisecond timescale.[162] gpHDX reactions are seen to occur at a faster rate when the difference between the gas-phase basicity for the labile residue and the exchange reagent is small. For reference the gas-phase proton affinity (GPA) of ammonia is ~850 $kJmol^{-1}$ [166] and 1026, 978, 971 $kJmol^{-1}$ for the Arg, Lys, and His residues respectively.[167] The onium mechanism suggests that the basic ammonia forms a favourable interaction

with a positively charged site, which then directly abstracts a proton, an endothermic process, to result in an ammonium ion ($[ND_3H]^+$) intermediate.[168] The intermediate may be stabilised by neighbouring electronegative moieties such as carboxylic acids and backbone carbonyls.[169] The basic site on the amino acid re-abstracts a deuteron and the encounter complex is broken. Exchange at such sites have also shown to be hindered by capping basic sites, demonstrating the influence non-covalent interactions may have on the exchange rate.[170,171] That exchange competent sites may be stabilised or restricted by the surrounding environment is what makes gpHDX a sensitive probe of a protein's surface. Conversely it is also this feature that has made predicting/modelling gpHDX very difficult for large systems.[172]

Whilst gpHDX is sensitive to the exchange competency of surface residues and the local structure, it should also be noted that the kinetic energy of the exchange collision may have an impact on the rate of exchange. Previous gpHDX studies involving small molecule and peptide ion-neutral collisions have shown that an increase in kinetic energy can result in a decrease in uptake.[157,159,162,173] This has been attributed to the disruption of the ion-neutral encounter complex that is required for exchange, where excess energy is expected to reduce the lifetime of the reagent-ion interaction. As such, for structural interpretations to be made under differing collision energies care must be taken to thermalise ions with translationally cooling collisions before exchange collisions.[174]

A number of studies have used gpHDX to study native proteins with probing conformational states in mind.[175] Multiple uptake distributions during long D_2O exposures have been attributed to multiple gas-phase conformations.[176] The effect of charge state is also seen to differ between the two labelling reagent classes. Denatured proteins have shown a decrease in uptake with increasing charge.[177–179] Whilst ND_3 studies have seen a strong reverse trend,[180–182] this may be due to the requirement that only a single charged site is necessary for exchange and not additional residues for a bridging interaction. A small number of studies have combined IM and gpHDX to study protein conformations,[179,181–184] though none

have used sequential mobility and rapid labelling using ND_3 on native systems, as in this thesis. This will be discussed further in Chapter 6.

References

[1] Richardson, M. J. The Direct Observation of Polymer Molecules and Determination of Their Molecular Weight. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **1964**, *279*, 50–61.

[2] Yamashita, M.; Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *Journal of Physical Chemistry* **1984**, *88*, 4451–4459.

[3] Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64–71.

[4] Meng, C. K.; Mann, M.; Fenn, J. B. Of protons or proteins. *Zeitschrift fur Physik D Atoms, Molecules and Clusters* **1988**, *10*, 361–368.

[5] Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. New Developments in Biochemical Mass Spectrometry: Electrospray Ionization. *Analytical Chemistry* **1990**, *62*, 882–899.

[6] Katta, V.; Chait, B. T. Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *Journal of the American Chemical Society* **1991**, *113*, 8534–8535.

[7] Hilton, G. R.; Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of The Royal Society Interface* **2012**, *9*, 801–816.

[8] Bernstein, S. L.; Bowers, M. T. *Comprehensive Analytical Chemistry*; Comprehensive Analytical Chemistry; Elsevier, 2008; Vol. 52; pp 63–82.

[9] Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 1827–1835.

[10] Benesch, J. L. P.; Sobott, F.; Robinson, C. V. Thermal Dissociation of Multimeric Protein Complexes by Using Nanoelectrospray Mass Spectrometry. *Analytical Chemistry* **2003**, *75*, 2208–2214.

[11] Laszlo, K. J.; Buckner, J. H.; Munger, E. B.; Bush, M. F. Native-Like and Denatured Cytochrome c Ions Yield Cation-to-Anion Proton Transfer Reaction Products with Similar Collision Cross-Sections. *Journal of The American Society for Mass Spectrometry* **2017**, *28*, 1382–1391.

[12] Gabelica, V.; Pauw, E. D. Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrometry Reviews* **2005**, *24*, 566–587.

[13] Benkestock, K.; Sundqvist, G.; Edlund, P. O.; Roeraade, J. Influence of droplet size, capillary-cone distance and selected instrumental parameters for the analysis

of noncovalent protein-ligand complexes by nano-electrospray ionization mass spectrometry. *Journal of Mass Spectrometry* **2004**, *39*, 1059–1067.

[14] Xia, Z.; Williams, E. R. Effect of droplet lifetime on where ions are formed in electrospray ionization. *Analyst* **2019**, *144*, 237–248.

[15] Susa, A. C.; Xia, Z.; Williams, E. R. Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment. *Analytical Chemistry* **2017**, *89*, 3116–3122.

[16] Xia, Z.; Williams, E. R. Protein-Glass Surface Interactions and Ion Desalting in Electrospray Ionization with Submicron Emitters. *Journal of the American Society for Mass Spectrometry* **2018**, *29*, 194–202.

[17] Hernández, H.; Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols* **2007**, *2*, 715–26.

[18] Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the mechanism of electrospray ionization. *Analytical Chemistry* **2013**, *85*, 2–9.

[19] Iribarne, J. V.; Thomson, B. A. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics* **1976**, *64*, 2287–2294.

[20] Dole, M.; Mack, L. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular Beams of Macroions. *The Journal of Chemical Physics* **1968**, *49*, 2240–2249.

[21] De La Mora, J. F. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica Chimica Acta* **2000**, *406*, 93–104.

[22] Rayleigh, L. On the equilibrium of liquid conducting masses charged with electricity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1882**, *14*, 184–186.

[23] Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. Droplet evaporation and discharge dynamics in electrospray ionization. *Journal of Physical Chemistry A* **2002**, *106*, 9957–9967.

[24] McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. Release of Native-like Gaseous Proteins from Electrospray Droplets via the Charged Residue Mechanism: Insights from Molecular Dynamics Simulations. *Journal of the American Chemical Society* **2015**, *137*, 12667–12676.

[25] Tolic, L. P.; Anderson, G. A.; Smith, R. D.; Brothers II, H. M.; Spindler, R.; Tomalia, D. A. Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometric characterization of high molecular mass Starburst dendrimers. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165/166*, 405–418.

[26] Testa, L.; Brocca, S.; Grandori, R. Charge-surface correlation in electrospray ionization of folded and unfolded proteins. *Analytical Chemistry* **2011**, *83*, 6459–6463.

[27] Stengel, F.; Baldwin, A. J.; Bush, M. F.; Hilton, G. R.; Lioe, H.; Basha, E.; Jaya, N.; Vierling, E.; Benesch, J. L. P. Dissecting Heterogeneous Molecular Chaperone Complexes Using a Mass Spectrum Deconvolution Approach. *Chemistry & Biology* **2012**, *19*, 599–607.

[28] Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. R. Studying 18 MDa virus assemblies with native mass spectrometry. *Angewandte Chemie - International Edition* **2013**, *52*, 4020–4023.

[29] Kaltashov, I. A.; Mohimen, A. Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Analytical Chemistry* **2005**, *77*, 5370–5379.

[30] Winger, B. E.; Light-Wahl, K. J.; Ogorzalek Loo, R. R.; Udseth, H. R.; Smith, R. D. Observation and implications of high mass-to-charge ratio ions from electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry* **1993**, *4*, 536–545.

[31] Hartman, R. P.; Brunner, D. J.; Camelot, D. M.; Marijnissen, J. C.; Scarlett, B. Jet break-up in electrohydrodynamic atomization in the cone-jet mode. *Journal of Aerosol Science* **2000**, *31*, 65–95.

[32] Liu, J.; Konermann, L. Protein-protein binding affinities in solution determined by electrospray mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2011**, *22*, 408–17.

[33] Natalello, A.; Santambrogio, C.; Grandori, R. Are Charge-State Distributions a Reliable Tool Describing Molecular Ensembles of Intrinsically Disordered Proteins by Native MS? *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 21–28.

[34] Hall, Z.; Robinson, C. V. Do charge state signatures guarantee protein conformations? *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 1161–1168.

[35] Benesch, J. L. P.; Robinson, C. V. Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Current opinion in structural biology* **2006**, *16*, 245–51.

[36] Metwally, H.; Konermann, L. Crown Ether Effects on the Location of Charge Carriers in Electrospray Droplets: Implications for the Mechanism of Protein Charging and Supercharging. *Analytical Chemistry* **2018**, *90*, 4126–4134.

[37] Marchese, R.; Grandori, R.; Carloni, P.; Raugei, S. A computational model for protein ionization by electrospray based on gas-phase basicity. *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 1903–1910.

[38] Susa, A. C.; Xia, Z.; Tang, H. Y. H.; Tainer, J. A.; Williams, E. R. Charging of Proteins in Native Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2016**, 20–27.

[39] Li, J.; Santambrogio, C.; Brocca, S.; Rossetti, G.; Carloni, P.; Grandori, R. Conformational effects in protein electrospray-ionization mass spectrometry. *Mass Spectrometry Reviews* **2016**, *35*, 111–122.

[40] Beveridge, R.; Phillips, A. S.; Denbigh, L.; Saleem, H. M.; MacPhee, C. E.; Barran, P. E. Relating gas phase to solution conformations: Lessons from disordered proteins. *Proteomics* **2015**, *15*, 2872–2883.

[41] Reading, E.; Liko, I.; Allison, T. M.; Benesch, J. L.; Laganowsky, A.; Robinson, C. V. The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase. *Angewandte Chemie - International Edition* **2015**, *54*, 4577–4581.

[42] Konermann, L.; Douglas, D. J. Unfolding of proteins monitored by electrospray ionization mass spectrometry: A comparison of positive and negative ion modes. *Journal of the American Society for Mass Spectrometry* **1998**, *9*, 1248–1254.

[43] Mehmood, S.; Marcoux, J.; Hopper, J. T.; Allison, T. M.; Liko, I.; Borysik, A. J.; Robinson, C. V. Charge reduction stabilizes intact membrane protein complexes for mass spectrometry. *Journal of the American Chemical Society* **2014**, *136*, 17010–17012.

[44] Loo, J. A.; Udseth, H. R.; Smith, R. D.; Futrell, J. H. Collisional effects on the charge distribution of ions from large molecules, formed by electrospray-ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **1988**, *2*, 207–210.

[45] Zhong, Y.; Han, L.; Ruotolo, B. T. Collisional and Coulombic unfolding of gas-phase proteins: high correlation to their domain structures in solution. *Angewandte Chemie (International ed. in English)* **2014**, *53*, 9209–12.

[46] Catalina, M. I.; Van Den Heuvel, R. H.; Van Duijn, E.; Heck, A. J. Decharging of globular proteins and protein complexes in electrospray. *Chemistry - A European Journal* **2005**, *11*, 960–968.

[47] Pagel, K.; Hyung, S.-J.; Ruotolo, B. T.; Robinson, C. V. Alternate Dissociation Pathways Identified in Charge-Reduced Protein Complex Ions. *Anal. Chem.* **2010**, *82*, 5363–5372.

[48] Zhuang, X.; Gavriilidou, A. F.; Zenobi, R. Influence of Alkylammonium Acetate Buffers on Protein–Ligand Noncovalent Interactions Using Native Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 341–346.

[49] McLuckey, S. A.; Reid, G. E.; Wells, J. M. Ion parking during ion/ion reactions in electrodynamic ion traps. *Analytical Chemistry* **2002**, *74*, 336–346.

[50] Laszlo, K. J.; Munger, E. B.; Bush, M. F. Folding of Protein Ions in the Gas Phase after Cation-to-Anion Proton-Transfer Reactions. *Journal of the American Chemical Society* **2016**, *138*, 9581–9588.

[51] Hopper, J. T. S.; Sokratous, K.; Oldham, N. J. Charge state and adduct reduction in electrospray ionization-mass spectrometry using solvent vapor exposure. *Analytical biochemistry* **2012**, *421*, 788–90.

[52] Scalf, M.; Westphall, M. S.; Krause, J.; Kaufman, S. L.; Smith, L. M. Controlling Charge States of Large Ions. *Science* **1999**, *283*, 194–197.

[53] Chernushevich, I. V.; Thomson, B. a. Collisional Cooling of Large Ions in Electrospray Mass Spectrometry. *Analytical Chemistry* **2004**, *76*, 1754–1760.

[54] Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. a.; Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chemical Reviews* **2007**, *107*, 3544–3567.

[55] Douglas, D. J.; French, J. B. Collisional Focusing Effects in Radio-Frequency Quadrupoles. *J Am Soc Mass Spectrom* **1992**, *3*, 398–408.

[56] Krutchinsky, A. N.; Chernushevich, I. V.; Spicer, V. L.; Ens, W.; Standing, K. G. Collisional damping interface for an electrospray ionization time-of-flight mass spectrometer. *Journal of the American Society for Mass Spectrometry* **1998**, *9*, 569–579.

[57] Sobott, F.; Hernández, F. H.; McCammon, M. G.; Tito, M. a.; Robinson, C. V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Analytical Chemistry* **2002**, *74*, 1402–1407.

[58] de Hoffmann, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications*; Wiley, 2007.

[59] Collings, B. A.; Douglas, D. J. An extended mass range quadrupole for electrospray mass spectrometry. *International Journal of Mass Spectrometry* **1997**, *162*, 121–127.

[60] Wiley, W. C.; McLaren, I. H. Time-of-flight mass spectrometer with improved resolution. *Review of Scientific Instruments* **1955**, *26*, 1150–1157.

[61] Guilhaus, M.; Selby, D.; Mlynski, V. Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrometry Reviews* **2000**, *19*, 65–107.

[62] Dawson, J. H. J.; Guilhaus, M. Orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Communications in Mass Spectrometry* **1989**, *3*, 155–159.

[63] Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Communications in Mass Spectrometry* **1996**, *10*, 889–896.

[64] Chernushevich, I. V.; Verentchikov, A. N.; Ens, W.; Standing, K. G. Effect of ion-molecule collisions in the vacuum chamber of an electrospray time-of-flight mass spectrometer on mass spectra of proteins. *Journal of the American Society for Mass Spectrometry* **1996**, *7*, 342–349.

[65] Kaltashov, I. A.; Eyles, S. J. *Mass spectrometry in structural biology and biophysics: Architecture, dynamics, and interaction of biomolecules*, 2nd ed.; Wiley, 2012.

[66] Lössl, P.; Snijder, J.; Heck, A. J. R. Boundaries of Mass Resolution in Native Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 906–917.

[67] Coles, J. N.; Guilhaus, M. Resolution limitations from detector pulse width and jitter in a linear orthogonal-acceleration time-of-flight mass spectrometer. *Journal of the American Society for Mass Spectrometry* **1994**, *5*, 772–778.

[68] Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. An introduction to quadrupole-time-of-flight mass spectrometry. *Journal of Mass Spectrometry* **2001**, *36*, 849–865.

[69] Lewin, M.; Guilhaus, M.; Wildgoose, J.; Hoyes, J.; Bateman, B. Ion dispersion near parallel wire grids in orthogonal acceleration time-of-flight mass spectrometry: predicting the effect of the approach angle on resolution. *Rapid Communications in Mass Spectrometry* **2002**, *16*, 609–615.

[70] Verenchikov, A.; Yavor, M.; Hasin, Y.; Gavrik, M. Multireflection planar mass analyzer, The analyzer for parallel tandem spectrometer. *Technical Physics* **2005**, *50*, 74–77.

[71] Verenchikov, A.; Yavor, M.; Hasin, Y.; Gavrik, M. Multireflection Planar Time-of-Flight Mass Analyzer. II: The High-Resolution Mode. *Technical Physics* **2005**, *50*, 82–86.

[72] Richardson, K.; Hoyes, J. A novel multipass oa-TOF mass spectrometer. *International Journal of Mass Spectrometry* **2015**, *377*, 309–315.

[73] Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nature Methods* **2012**, *9*, 1084–1086.

[74] Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry* **2000**, *72*, 1156–1162.

[75] Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Analytical Chemistry* **2006**, *78*, 2113–2120.

[76] Fort, K. L.; Van De Waterbeemd, M.; Boll, D.; Reinhardt-Szyba, M.; Belov, M. E.; Sasaki, E.; Zschoche, R.; Hilvert, D.; Makarov, A. A.; Heck, A. J. Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. *Analyst* **2018**, *143*, 100–105.

[77] Maple, H. J.; Scheibner, O.; Baumert, M.; Allen, M.; Taylor, R. J.; Garlish, R. A.; Bromirski, M.; Burnley, R. J. Application of the Exactive Plus EMR for automated protein-ligand screening by non-covalent mass spectrometry. *Rapid Communications in Mass Spectrometry* **2014**, *28*, 1561–1568.

[78] Campuzano, I. D. G.; Li, H.; Bagal, D.; Lippens, J. L.; Svitel, J.; Kurzeja, R. J. M.; Xu, H.; Schnier, P. D.; Loo, J. A. Native MS Analysis of Bacteriorhodopsin and an Empty Nanodisc by Orthogonal Acceleration Time-of-Flight, Orbitrap and Ion Cyclotron Resonance. *Analytical Chemistry* **2016**, *88*, 12427–12436.

[79] Gabelica, V.; Marklund, E. Fundamentals of ion mobility spectrometry. *Current Opinion in Chemical Biology* **2018**, *42*, 51–59.

[80] Mason, E. A.; McDaniel, E. W. *Transport properties of ions in gases*; Wiley, 1988.

[81] Ujma, J.; Giles, K.; Morris, M.; Barran, P. E. New High Resolution Ion Mobility Mass Spectrometer Capable of Measurements of Collision Cross Sections from 150 to 520 K. *Analytical Chemistry* **2016**, *88*, 9469–9478.

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

[83] Giles, K.; Wildgoose, J. L.; Langridge, D. J.; Campuzano, I. A method for direct measurement of ion mobilities using a travelling wave ion guide. *International Journal of Mass Spectrometry* **2010**, *298*, 10–16.

[84] Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid communications in mass spectrometry : RCM* **2004**, *18*, 2401–14.

[85] Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in travelling wave ion mobility resolution. *Rapid Communications in Mass Spectrometry* **2011**, *25*, 1559–1566.

[86] Morsa, D.; Gabelica, V.; De Pauw, E. Fragmentation and Isomerization Due to Field Heating in Traveling Wave Ion Mobility Spectrometry. *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 1384–1393.

[87] Merenbloom, S. I.; Flick, T. G.; Williams, E. R. How hot are your ions in TWAVE ion mobility spectrometry? *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 553–562.

[88] Mortensen, D. N.; Susa, A. C.; Williams, E. R. Collisional Cross-Sections with T-Wave Ion Mobility Spectrometry without Experimental Calibration. *Journal of The American Society for Mass Spectrometry* **2017**, 37–39.

[89] Richardson, K.; Langridge, D.; Giles, K. International Journal of Mass Spectrometry Fundamentals of travelling wave ion mobility revisited : I . Smoothly moving waves. *International Journal of Mass Spectrometry* **2018**, *428*, 71–80.

[90] Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Analytical chemistry* **2010**, *82*, 9557–65.

[91] Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry* **2008**, *80*, 9689–9699.

[92] Baldwin, A. J.; Lioe, H.; Hilton, G. R.; Baker, L. A.; Rubinstein, J. L.; Kay, L. E.; Benesch, J. L. P. Article The Polydispersity of a B-Crystallin Is Rationalized by an Interconverting Polyhedral Architecture. *Structure/Folding and Design* **2011**, *19*, 1855–1863.

[93] Politis, A.; Stengel, F.; Hall, Z.; Hernández, H.; Leitner, A.; Walzthoeni, T.; Robinson, C. V.; Aebersold, R. A mass spectrometry–based hybrid method for structural modeling of protein complexes. *Nature Methods* **2014**, *11*, 403–406.

[94] Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature Chemistry* **2014**, *6*, 281–294.

[95] Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. Collision Cross Sections for Structural Proteomics. *Structure* **2015**, *23*, 791–799.

[96] Benesch, J. L.; Ruotolo, B. T. Mass spectrometry: come of age for structural and dynamical biology. *Current Opinion in Structural Biology* **2011**, *21*, 641–649.

[97] Marklund, E. G.; Ekeberg, T.; Moog, M.; Benesch, J. L.; Caleman, C. Controlling Protein Orientation in Vacuum Using Electric Fields. *Journal of Physical Chemistry Letters* **2017**, *8*, 4540–4544.

[98] May, J. C.; McLean, J. A. Ion Mobility-Mass Spectrometry: Time-Dispersive Instrumentation. *Analytical chemistry* **2014**,

[99] Michelmann, K.; Silveira, J. A.; Ridgeway, M. E.; Park, M. A. Fundamentals of trapped ion mobility spectrometry. *Journal of the American Society for Mass Spectrometry* **2015**, *26*, 14–24.

[100] Fernandez-Lima, F. Trapped Ion Mobility Spectrometry: past, present and future trends. *International Journal for Ion Mobility Spectrometry* **2016**, *19*, 65–67.

[101] Hernandez, D. R.; DeBord, J. D.; Ridgeway, M. E.; Kaplan, D. A.; Park, M. A.; Fernandez-Lima, F. Ion dynamics in a trapped ion mobility spectrometer. *Analyst* **2014**, *139*, 1913–1921.

[102] Silveira, J. A.; Ridgeway, M. E.; Park, M. A. High resolution trapped ion mobility spectrometery of peptides. *Analytical Chemistry* **2014**, *86*, 5624–5627.

[103] Liu, F. C.; Ridgeway, M. E.; Park, M. A.; Bleiholder, C. Tandem trapped ion mobility spectrometry. *Analyst* **2018**, *143*, 2249–2258.

[104] Koeniger, S. L.; Merenbloom, S. I.; Valentine, S. J.; Jarrold, M. F.; Udseth, H. R.; Smith, R. D.; Clemmer, D. E. An IMS - IMS Analogue of MS - MS. *Analytical chemistry* **2006**, *78*, 4161–4174.

[105] Merenbloom, S. I.; Glaskin, R. S.; Henson, Z. B.; Clemmer, D. E. High-Resolution Ion Cyclotron Mobility Spectrometry. *Analytical Chemistry* **2009**, *81*, 1482–1487.

[106] Giles, K.; Wildgoose, J.; Pringle, S.; Garside, J.; Carney, P.; Nixon, P.; Langridge, D. Design and utility of a multi-pass cyclic ion mobility separator.

[107] Eldrid, C. F. S.; Ujma, J.; Kalfas, S.; Tomczyk, N.; Giles, K.; Morris, M. R.; Thalassinos, K. Gas Phase Stability of Protein Ions in a Cyclic Ion Mobility Spectrometry Travelling Wave Device. *Analytical Chemistry Just Accepted Manuscript*.

[108] Hamid, A. M.; Garimella, S. V. B.; Ibrahim, Y. M.; Deng, L.; Zheng, X.; Webb, I. K.; Anderson, G. A.; Prost, S. A.; Norheim, R. V.; Tolmachev, A. V.; Baker, E. S.; Smith, R. D. Achieving High Resolution Ion Mobility Separations Using Traveling Waves in Compact Multiturn Structures for Lossless Ion Manipulations. *Analytical Chemistry* **2016**, *88*, 8949–8956.

[109] Sanders, J. D.; Grinfeld, D.; Aizikov, K.; Makarov, A.; Holden, D. D.; Brodbelt, J. S. Determination of Collision Cross-Sections of Protein Ions in an Orbitrap Mass Analyzer. *Analytical Chemistry* **2018**, *90*, 5896–5902.

[110] Poltash, M. L.; McCabe, J. W.; Shirzadeh, M.; Laganowsky, A.; Clowers, B. H.; Russell, D. H. Fourier Transform-Ion Mobility-Orbitrap Mass Spectrometer: A Next-Generation Instrument for Native Mass Spectrometry. *Analytical Chemistry* **2018**, *90*, 10472–10478.

[111] Mcluckey, S. A.; Goeringer, D. E. Slow Heating Methods in Tandem Mass Spectrometry. *Journal of Mass Spectrometry* **1997**, *32*, 461–474.

[112] Shukla, A. K.; Futrell, J. H. Tandem mass spectrometry: Dissociation of ions by collisional activation. *Journal of Mass Spectrometry* **2000**, *35*, 1069–1090.

[113] Gill, A. C.; Jennings, K. R.; Wyttenbach, T.; Bowers, M. T. Conformations of biopolymers in the gas phase: A new mass spectrometric method. *International Journal of Mass Spectrometry* **2000**, *195-196*, 685–697.

[114] Mikhailov, V. A.; Liko, I.; Mize, T. H.; Bush, M. F.; Benesch, J. L.; Robinson, C. V. Infrared Laser Activation of Soluble and Membrane Protein Assemblies in the Gas Phase. *Analytical Chemistry* **2016**, *88*, 7060–7067.

[115] Lermyte, F.; Konijnenberg, A.; Williams, J. P.; Brown, J. M.; Valkenborg, D.; Sobott, F. ETD allows for native surface mapping of a 150 kDa noncovalent complex on a commercial Q-TWIMS-TOF instrument. *Journal of the American Society for Mass Spectrometry* **2014**, *25*, 343–50.

[116] Jones, C. M.; Beardsley, R. L.; Galhena, A. S.; Dagan, S.; Cheng, G.; Wysocki, V. H. Symmetrical gas-phase dissociation of noncovalent protein complexes via surface collisions. *Journal of the American Chemical Society* **2006**, *128*, 15044–15045.

[117] Benesch, J. L. P.; Robinson, C. V. Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Current opinion in structural biology* **2006**, *16*, 245–51.

[118] Sharon, M.; Mao, H.; Boeri Erba, E.; Stephens, E.; Zheng, N.; Robinson, C. V. Symmetrical Modularity of the COP9 Signalosome Complex Suggests its Multifunctionality. *Structure* **2009**, *17*, 31–40.

[119] Zhou, M.; Morgner, N.; Barrera, N. P.; Politis, A.; Isaacson, S. C.; Matak-Vinkovic, D.; Murata, T.; Bernal, R. A.; Stock, D.; Robinson, C. V. Mass Spectrometry of Intact V-Type ATPases Reveals Bound Lipids and the Effects of Nucleotide Binding. *Science* **2011**, *334*, 380–385.

[120] Pliotas, C. et al. Adenosine Monophosphate Binding Stabilizes the KTN Domain of the Shewanella denitrificans Kef Potassium Efflux System. *Biochemistry* **2017**, *56*, 4219–4234.

[121] Gerken, P. A. et al. Discovery of a Highly Selective Cell-Active Inhibitor of the Histone Lysine Demethylases KDM2/7. *Angewandte Chemie - International Edition* **2017**, *56*, 15555–15559.

[122] Mishra, S.; Chandler, S. A.; Williams, D.; Claxton, D. P.; Koteiche, H. A.; Stewart, P. L.; Benesch, J. L.; Mchaourab, H. S. Engineering of a Polydisperse Small Heat-Shock Protein Reveals Conserved Motifs of Oligomer Plasticity. *Structure* **2018**, *26*, 1116–1126.e4.

[123] Jurchen, J. C.; Williams, E. R. Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers. *Journal of the American Chemical Society* **2003**, *125*, 2817–2826.

[124] Sinelnikov, I.; Kitova, E. N.; Klassen, J. S. Influence of Coulombic Repulsion on the Dissociation Pathways and Energetics of Multiprotein Complexes in the Gas Phase. *Journal of the American Society for Mass Spectrometry* **2007**, *18*, 617–631.

[125] Sciuto, S. V.; Liu, J.; Konermann, L. An electrostatic charge partitioning model for the dissociation of protein complexes in the gas phase. *Journal of the American Society for Mass Spectrometry* **2011**, *22*, 1679–1689.

[126] Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chemistry & biology* **2006**, *13*, 597–605.

[127] Fegan, S. K.; Thachuk, M. Controlling Dissociation Channels of Gas-Phase Protein Complexes Using Charge Manipulation. *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 722–728.

[128] Popa, V.; Trecroce, D. A.; McAllister, R. G.; Konermann, L. Collision-Induced Dissociation of Electrosprayed Protein Complexes: An All-Atom Molecular Dynamics Model with Mobile Protons. *Journal of Physical Chemistry B* **2016**, *120*, 5114–5124.

[129] Schwartz, B. L.; Bruce, J. E.; Anderson, G. A.; Hofstadler, S. A.; Rockwood, A. L.; Smith, R. D.; Chilkoti, A.; Stayton, P. S. Dissociation of tetrameric ions of noncovalent streptavidin complexes formed by electrospray ionization. *Journal of the American Society for Mass Spectrometry* **1995**, *6*, 459–465.

[130] Loo, R. R.; Loo, J. A. Salt Bridge Rearrangement (SaBRe) Explains the Dissociation Behavior of Noncovalent Complexes. *Journal of the American Society for Mass Spectrometry* **2016**, *27*, 975–990.

[131] Englander, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical

Perspective. *Journal of the American Society for Mass Spectrometry* **2006**, *17*, 1481–1489.

[132] Hvidt, A.; Linderstrøm-Lang, K. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochimica et Biophysica Acta* **1954**, *14*, 574–575.

[133] Zhang, Z.; Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein science : a publication of the Protein Society* **1993**, *2*, 522–31.

[134] Engen, J. R. Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Analytical Chemistry* **2009**, *81*, 7870–7875.

[135] Katta, V.; Chait, B. T. Hydrogen/Deuterium Exchange Electrospray Ionization Mass Spectrometry: A Method for Probing Protein Conformational Changes in Solution. *Journal of the American Chemical Society* **1993**, *115*, 6317–6321.

[136] Miranker, A.; Robinson, C.; Radford, S.; Aplin, R.; Dobson, C. Detection of transient protein folding populations by mass spectrometry. *Science* **1993**, *262*, 896–900.

[137] Molday, R. S.; Englander, S. W.; Kallen, R. G. Primary structure effects on peptide group hydrogen exchange. *Biochemistry* **1972**, *11*, 150–158.

[138] Bai, Y.; Milne, J.; Mayne, L.; Englander, W. Primary Structure Effects on Petptide Group Hydrogen Exchange. *Proteins* **1993**, *17*, 75–86.

[139] Wales, T. E.; Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **2006**, *25*, 158–170.

[140] Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **2011**, *40*, 1224–1234.

[141] Oganesyan, I.; Lento, C.; Wilson, D. J. Contemporary hydrogen deuterium exchange mass spectrometry. *Methods* **2018**, *144*, 27–42.

[142] Slysz, G. W.; Percy, A. J.; Schriemer, D. C. Restraining expansion of the peak envelope in H/D exchange-MS and its application in detecting perturbations of protein structure/dynamics. *Analytical Chemistry* **2008**, *80*, 7004–7011.

[143] Trabjerg, E.; Nazari, Z. E.; Rand, K. D. Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): Challenges and emerging solutions. *TrAC - Trends in Analytical Chemistry* **2018**, *106*, 125–138.

[144] Pan, J.; Zhang, S.; Parker, C. E.; Borchers, C. H. Subzero temperature chromatography and top-down mass spectrometry for protein higher-order structure characterization: method validation and application to therapeutic antibodies. *Journal of the American Chemical Society* **2014**, *136*, 13065–71.

[145] Sheff, J. G.; Rey, M.; Schriemer, D. C. Peptide-column interactions and their influence on back exchange rates in hydrogen/deuterium exchange-MS. *Journal of the American Society for Mass Spectrometry* **2013**, *24*, 1006–1015.

[146] Rey, M.; Yang, M.; Burns, K. M.; Yu, Y.; Lees-Miller, S. P.; Schriemer, D. C. Nepenthesin from Monkey Cups for Hydrogen/Deuterium Exchange Mass Spectrometry. *Molecular & Cellular Proteomics* **2013**, *12*, 464–472.

[147] Burns, K. M.; Rey, M.; Baker, C. A. H.; Schriemer, D. C. Platform Dependencies in Bottom-up Hydrogen/Deuterium Exchange Mass Spectrometry. *Molecular & Cellular Proteomics* **2013**, *12*, 539–548.

[148] Maity, H.; Lim, W. K.; Rumbley, J. N.; Englander, S. W. Protein hydrogen exchange mechanism: Local fluctuations. *Protein Science* **2003**, *12*, 153–160.

[149] Persson, F.; Halle, B. How amide hydrogens exchange in native proteins. *Proceedings of the National Academy of Sciences* **2015**, *112*, 10383–10388.

[150] Miller, D. W.; Dill, K. A. A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Science* **1995**, *4*, 1860–1873.

[151] Konermann, L.; Simmons, D. A. Protein-folding kinetics and mechanisms studied by pulse-labeling and mass spectrometry. *Mass Spectrometry Reviews* **2003**, *22*, 1–26.

[152] 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. *Analytical Chemistry* **2012**, *84*, 3771–3779.

[153] 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. *Journal of the American Chemical Society* **2009**, *131*, 12801–12808.

[154] Rand, K. D.; Zehl, M.; Jørgensen, T. J. Measuring the hydrogen/deuterium exchange of proteins at high spatial resolution by mass spectrometry: Overcoming gas-phase hydrogen/deuterium scrambling. *Accounts of Chemical Research* **2014**, *47*, 3018–3027.

[155] Brodie, N. I.; Huguet, R.; Zhang, T.; Viner, R.; Zabrouskov, V.; Pan, J.; Petrotchenko, E. V.; Borchers, C. H. Top-Down Hydrogen-Deuterium Exchange Analysis of Protein Structures Using Ultraviolet Photodissociation. *Analytical Chemistry* **2018**, *90*, 3079–3082.

[156] Mistarz, U. H.; Bellina, B.; Jensen, P. F.; Brown, J. M.; Barran, P. E.; Rand, K. D. UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Analytical Chemistry* **2018**, *90*, 1077–1080.

[157] Green, M. K.; Lebrilla, C. B. Ion-molecule reactions as probes of gas-phase structures of peptides and proteins. *Mass spectrometry reviews* **1997**, *16*, 53–71.

[158] Winger, B. E.; Light-Wahl, K. J.; Rockwood, A. L.; Smith, R. D. Probing Qualitative Conformation Differences of Multiply Protonated Gas-Phase Proteins via H/D Isotopic Exchange with D2O. *Journal of the American Chemical Society* **1992**, *114*, 5897–5898.

[159] Cheng, X.; Fenselau, C. Hydrogen/deuterium exchange of mass-selected peptide ions with ND3 in a tandem sector mass spectrometer. *International Journal of Mass Spectrometry and Ion Processes* **1992**, *122*, 109–119.

[160] Pan, J.; Heath, B. L.; Jockusch, R. A.; Konermann, L. Structural interrogation of electrosprayed peptide ions by gas-phase H/D exchange and electron capture dissociation mass spectrometry. *Analytical Chemistry* **2012**, *84*, 373–378.

[161] Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. Structural and Energetic Constraints on Gas Phase Hydrogen/Deuterium Exchange Reactions of Protonated Peptides with D_2O, CD_3OD, CD_3CO_2D, and ND_3. *Journal of the American Chemical Society* **1994**, *116*, 9765–9766.

[162] Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. Deuterium Exchange Reactions as a Probe of Biomolecule Structure. Fundamental Studies of Gas Phase H/D Exchange Reactions of Protonated Glycine Oligomers with D_2O, CD_3OD, CD_3CO_2D, and ND_3. *Journal of the American Chemical Society* **1995**, *117*, 12840–12854.

[163] Wyttenbach, T.; Bowers, M. T. Gas phase conformations of biological molecules: the hydrogen/deuterium exchange mechanism. *Journal of the American Society for Mass Spectrometry* **1999**, *10*, 9–14.

[164] Evans, S. E.; Lueck, N.; Marzluff, E. M. Gas phase hydrogen/deuterium exchange of proteins in an ion trap mass spectrometer. *International Journal of Mass Spectrometry* **2003**, *222*, 175–187.

[165] Khakinejad, M.; Kondalaji, S. G.; Tafreshian, A.; Valentine, S. J. Gas-Phase Hydrogen-Deuterium Exchange Labeling of Select Peptide Ion Conformer Types: a Per-Residue Kinetics Analysis. *Journal of The American Society for Mass Spectrometry* **2015**, 1115–1127.

[166] Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. Gas-phase ion and neutral thermochemistry. *Journal of Physical and Chemical Reference Data* **1988**, *17*.

[167] Marchese, R.; Grandori, R.; Carloni, P.; Raugei, S. On the zwitterionic nature of gas-phase peptides and protein ions. *PLoS computational biology* **2010**, *6*, e1000775.

[168] Dookeran, N. N.; Harrison, A. G. Gas-phase H-D exchange reactions of protonated amino acids and peptides with ND_3. *Journal of Mass Spectrometry* **1995**, *30*, 666–674.

[169] Zhu, C.; Balta, B.; Aviyente, V.; Lifshitz, C. The Interaction of Protonated Diglycine with Ammonia: A Density Functional Theory Model Study. *The Journal of Physical Chemistry A* **2000**, *104*, 7061–7067.

[170] Lee, S. W.; Lee, H. N.; Kim, H. S.; Beauchamp, J. L. Selective binding of crown ethers to protonated peptides can be used to probe mechanisms of H/D exchange and collision-induced dissociation reactions in the gas phase. *Journal of the American Chemical Society* **1998**, *120*, 5800–5805.

[171] Mistarz, U. H.; Brown, J. M.; Haselmann, K. F.; Rand, K. D. Probing the Binding Interfaces of Protein Complexes Using Gas-Phase H/D Exchange Mass Spectrometry. *Structure* **2016**, *24*, 310–318.

[172] Geller, O.; Lifshitz, C. Applying a new algorithm to H/D exchange of multiply protonated cytochrome c. *Int. J. Mass Spectrom.* **2004**, *233*, 125–129.

[173] Ranasinghe, A.; Cooks, R. G.; Sethi, S. K. Selective isotopic exchange of polyfimctional ions in tandem mass spectrometry: Methodology, applications and mechanism. *Organic Mass Spectrometry* **1992**, *27*, 77–88.

[174] Gur, E. H.; Koning, L. J.; Nibbering, N. M. M. The bimolecular hydrogen-deuterium exchange behavior of protonated alkyl dipeptides in the gas phase. *Journal of the American Society for Mass Spectrometry* **1995**, *6*, 466–477.

[175] Jarrold, M. F. Peptides and proteins in the vapor phase. *Annual Review of Physical Chemistry* **2000**, *51*, 179–207.

[176] Wood, T. D.; Chorush, R. a.; Wampler, F. M.; Little, D.; O'Connor, P. B.; McLafferty, F. W. Gas-phase folding and unfolding of cytochrome c cations. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 2451–2454.

[177] Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. Gas-phase bovine ubiquitin cation conformations resolved by gas-phase hydrogen/deuterium exchange rate and extent. *International Journal of Mass Spectrometry* **1999**, *185-187*, 565–575.

[178] McLafferty, F. W.; Guan, Z. Q.; Haupts, U.; Wood, T. D.; Kelleher, N. L. Gaseous conformational structures of cytochrome c. *J. Am. Chem. Soc.* **1998**, *120*, 4732–4740.

[179] Robinson, E. W.; Williams, E. R. Multidimensional separations of ubiquitin conformers in the gas phase: Relating ion cross sections to H/D exchange measurements. *Journal of the American Society for Mass Spectrometry* **2005**, *16*, 1427–1437.

[180] Geller, O.; Lifshitz, C. A Fast Flow Tube Study of Gas Phase H/D Exchange of Multiply Protonated Ubiquitin. *The Journal of Physical Chemistry A* **2005**, *109*, 2217–2222.

[181] Rand, K. D.; Pringle, S. D.; Murphy, J. P.; Fadgen, K. E.; Brown, J.; Engen, J. R. Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Analytical chemistry* **2009**, *81*, 10019–28.

[182] Beeston, H. S.; Ault, J. R.; Pringle, S. D.; Brown, J. M.; Ashcroft, A. E. Changes in protein structure monitored by use of gas-phase hydrogen/deuterium exchange. *Proteomics* **2015**, 2842–2850.

[183] Valentine, S. J.; Clemmer, D. E. H/D Exchange Levels of Shape-Resolved Cytochrome c Conformers in the Gas Phase. *Journal of the American Chemical Society* **1997**, *119*, 3558–3566.

[184] Zhu, S.; Campbell, J. L.; Chernushevich, I.; Le Blanc, J. C.; Wilson, D. J. Differential Mobility Spectrometry-Hydrogen Deuterium Exchange (DMS-HDX) as a Probe of Protein Conformation in Solution. *Journal of the American Society for Mass Spectrometry* **2016**, *27*, 991–999.

3

Protein structure and solution dynamics through HDX-MS

3.1 Abstract

H/D exchange (HDX) of labile hydrogens associated with the amide backbone is highly sensitive to the local structural environment. Monitoring the rate of deuterium incorporation using mass spectrometry (MS) therefore provides a powerful means of reporting on the solvent accessibility and hydrogen bonding network for each residue of the protein. In this chapter, we present the utility of HDX-MS for probing the in solution dynamics and assembly of two protein systems. Firstly, we use solution HDX to probe the light-induced structural change of the FAD-binding protein *Dm*-cryptochrome. Quantitative analysis of the exchange kinetics revealed a region at the C-terminus that undergoes a long-lived conformational change with high specificity to blue light. This suggests the potential role the cryptochrome C-terminus may have in down-stream interactions and signalling associated with circadian processes.

Secondly, we use HDX-MS to probe the interface of a homo-dimeric ancestor of extant haemoglobin. Modern-day haemoglobin forms a tetramer, or dimer of dimers, with two discrete interfaces between the α and β subunits. To understand how the interfaces evolved in modern-day haemoglobin, we studied the assembly of the phylogenetically engineered last common ancestor (Anc) of the α and β subunits. Anc was shown to form homo-dimers and differential HDX-MS was used to reveal the protected peptide regions at the interface. Our results suggest that the ancestral interface is consistent with the *packing* interface of the α/β subunits in modern-day haemoglobin.

3.2 Introduction

The H/D exchange (HDX) rate of the labile backbone amide hydrogens is extremely sensitive to their accessibility to a deuterated solvent, and restrictions imparted by the local hydrogen bonding network. These features therefore make monitoring the rate of exchange a powerful means of understanding protein structure and corresponding dynamic motions.[1–4] The deuterium incorporation can be readily measured using nuclear magnetic resonance (NMR) at a residue level or with mass spectrometry, which is capable of either a global measurement or localised to near-residue level with prior enzymatic digestion. Whilst HDX strategies cannot provide a high-resolution 3-dimensional structure as with X-ray crystallography, electron microscopy (EM) and NMR it is particularly capable of measuring local differences between two structural states, with significantly fewer experimental limitations than the aforementioned methods. Understanding differences in HDX can be significantly aided when related high resolution structures are available through homology models, or if a single "static" state has been resolved. HDX has a particularly important place, alongside NMR, in reporting on the dynamics not always resolved or discovered through diffraction techniques. In this Chapter we present two systems in which HDX-MS, with the aid of native MS, is used to monitor and quantify structural changes as a result of a conformational change and the formation of an interface, in light of existing static or homology structures.

3.3 Part I : Structural dynamics of a light-induced conformational change

3.3.1 Biological context

Cryptochromes (Cry) are a group of photoreceptors sensitive to blue light that are found across all kingdoms of life. A host of light-induced functional responses have been shown, from the entrainment of circadian rhythms in animals to the growth and development of plants and most intriguingly as a candidate for avian magnetoreception.[5–8] Upon exposure to blue light, crytpochromes from *Drosophila melanogaster*, *Dm*-Cry, are thought to undergo a structural change and have been shown to interact in a light dependent manner with a number of partners associated with biological clocks, including but not limited to, TIMELESS (TIM),[9] PERIOD (PER)[10] and jetlag (JET).[11] In doing so they are thought to disrupt downstream complex formation, preventing a negative feedback loop from operating whilst stimulated by light conditions. Little is known of the biophysics of the conformational or structural change responsible for this signalling cascade and the cause of downstream interactions.

Cryptochromes have a conserved photolyase homology domain, PHR, 514 residues long in which a chromophore flavin adenine dinucleotide, FAD, is bound,[12,13] shown in yellow in Figure 3.1A & B. Photoactivation of the oxidised form of FAD is well understood and is known to result in the formation of a radical pair.[7,8,14–16] A light activated state of FAD is produced, and following the sequential electron transfer from a neighbouring tryptophan tetrad[i] chain a flavosemiquinone radical is created, $FAD^{\bullet -}$, at the core of the protein and a radical tryptophan, $TrpH^{\bullet +}$, near the exterior and close to the C terminal tail (CTT), Fig. 3.1B. It is the existence of this radical pair that is thought to be crucial to the potential origins of magnetoreception seen in insects and avian species. Figure 3.1C illustrates the simplified photochemical scheme present in cryptochromes. The initial radical pair is generated as a singlet state, this may then quickly recombine to give the ground

[i]Evidence exists of the involvement of a fourth tryptophan though this was long reported as a functional triad.[17,18]

state or it may undergo deprotonation to give a second long-lived radical pair. The singlet state of the first radical pair may also undergo singlet-triplet interconversion; the triplet state is only capable of conversion to radical pair 2, owing to spin selection rules, and cannot recombine directly to the ground-state. The relative abundance of the triplet state generated will influence the amount of radical pair 2 formation; it is this long-lived state that is thought to be the signalling state. An external magnetic field is capable of perturbing this interconversion and as such cryptochrome has a biochemical means of populating states not only influenced by light but also a magnetic field. For more detail I would refer the reader to the review by Hore & Mouritsen[7] which gives a detailed overview of the radical-pair mechanism in magnetoreception.

A number of experiments implicate the CTT in being responsible for downstream regulation through a conformational change. The CTT is a region that differentiates cryptochrome from photolyases and differs significantly between species. The most terminal helix, α23, forms intramolecular contacts close to the bound FAD, shown in red (Fig. 3.1A). CTT deficient studies have been used *in vivo* to suggest a regulatory role of the CTT[19–21] with a small number utilising a limited proteolysis approach to note increased susceptibility to cleavage on exposure to light.[16,22–24] Structural elements surrounding the CTT that are thought to play an allosteric role and/or are involved in complex formation include the C-terminal lid (residues 420–446), the protrusion motif (residues 288–306), the phosphate-binding loop (residues 249–263), and the CTT base loop (residues 154–160), all shown in Figure 3.1A.

In this section we utilise HDX-MS to probe the local light-induced conformations of *Dm*-Cry. Limited proteolysis has previously provided insight into the light induced conformational change but the technique is inherently low resolution, whose sensitivity to change is limited by the mechanism of cleavage. In contrast, HDX-MS provides a non-invasive labelling method which in principle reports on the local structural changes for each backbone amide. HDX-MS is particularly suited to studying such a moderately sized dynamic system and can be coupled to an external

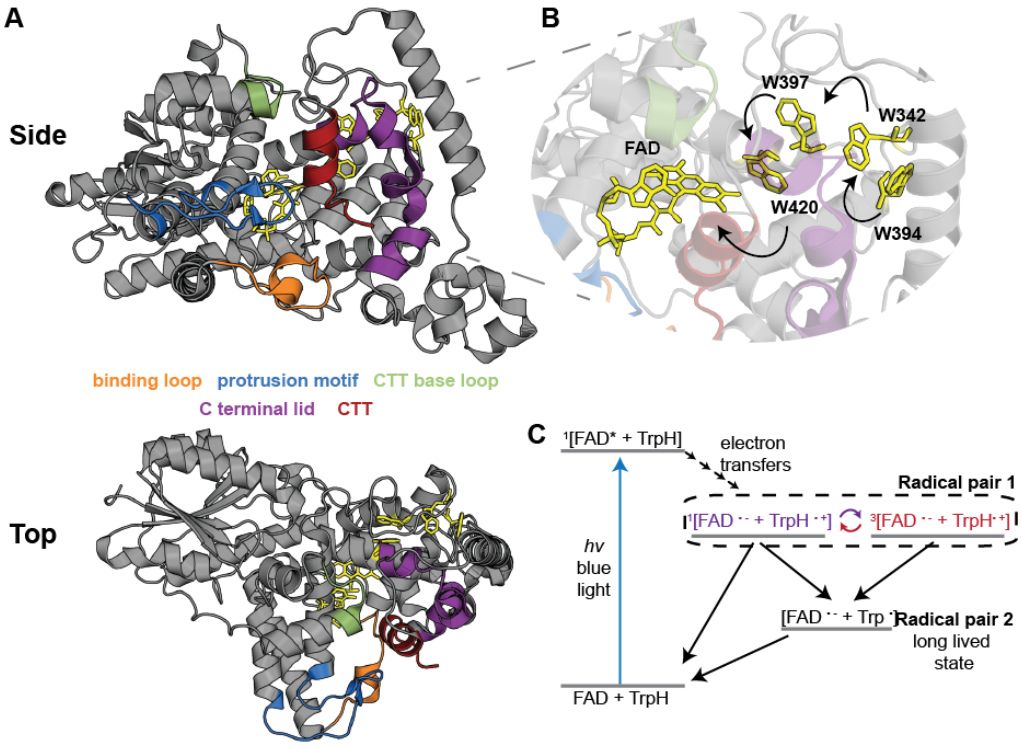


Figure 3.1. Photo-activation of *Drosphila melanogaster* Cryptochrome A *Dm*-Cry X-ray crystal structure (PDB: 4GU5[12]) in the dark state conformation, shown from the side and top. The bound cofactor, FAD, and corresponding tryptophan resides shown in yellow result in the formation of a radical pair upon blue light activation. The C-terminal tail (CTT) α23 helix, a region of significant structural interest, is shown in red. Motifs localised to the C-terminal tail are also highlighted for reference. **B** A zoom of the electron transfer and radical pair formation involving FAD and the Trp tetrad after blue light activation. The light activated state of FAD forms a radical after the electron transfer from W420. The resulting electron 'hole' then migrates to W397, W342 and finally W394 through further electron transfers. W394 is close to the exterior of the protein near the α22 helix. The radical pair is separated by ~2nm between FAD and W394. **C** A simplified photochemical scheme for the photo-activation and radical pair formation in cryptochromes. The ground state FAD becomes photoactivated and through electron transfers from the Trp tetrad forms a singlet state, radical pair 1. This can recombine directly to the ground state or produce a long lived deprotonated radical pair 2 state. It may also interconvert to a triplet state which may also convert to radical pair 2, but cannot directly recombine due to spin selection rules. The interconversion can be influenced by an external magnetic field thus affecting the population of radical pair 2. This forms the basis of the proposed biophysical magneto-sensor.

light source with relative ease.[25,26] NMR studies are often highly complementary to, and in many cases at higher resolution than HDX-MS, assuming size limitations can be circumvented. Understanding the cryptochrome's role in magnetoreception is of considerable interest and as NMR relies on a large external magnetic field to be present, this makes HDX-MS an ever more appropriate technique to sensitively probe the conformational states present. Here we report how HDX-MS is able to report on the dark state dynamics and importantly localise regions of light induced change. Some regions described here displayed a form of HDX kinetics not commonly observed for a native system[27] and allowed us to monitor absolute conformeric populations as a function of time and light.

3.3.2 Results and Discussion

Dark-state dynamics of *Dm*-Cryptochrome

In order to understand the basal dynamics of *Dm*-Cry under dark conditions we performed a local HDX-MS experiment to determine the relative rates of exchange as a function of time. Protein was kept in complete darkness for the entire duration of the experiment with prior manual handling performed exclusively under low-level red lights. The deuterium incorporation was determined for each peptide product obtained through solution enzymolysis using the experimental methodology briefly described in the Introduction and detailed in the Methods. The dark-state sequence coverage amounted to 80.5 % after data processing, with most regions lacking digestion rich in proline residues, typical when proline is local to the cleavage site in the P2-P3' positions[ii].[28] The large number of prolines limits the practical coverage losses due to the inability to exchange, because of the lack of a backbone amide. There was a high level of redundancy closer to the N and C termini with greater peptide overlap.

The resulting uptake is shown for each time point, 0.25 - 60 min, as a heat-map over the entire sequence, Figure 3.2A. Uptake values were normalised to the theoretical maximum, though no correction was applied for residual H_2O buffer,

[ii]For a peptide P3-P2-P1-P1'-P2'-P3' where peptic cleavage occurs between P1 and P1'.

linearised and plotted for each exchangeable residue. The resulting relative uptake ranged from 0 - 70 % (blue - red) with many regions displaying very low uptake, around 44 % of the covered sequence has a relative uptake <5 %. This only decreases to ∼36 % after 1 hour of continuous exposure suggesting a significant amount of the structure remains rigid and undergoes little conformational flexibility. Included in these regions are three of the tryptophan residues, W397, W420, W394[iii], in the radical pair Trp-tetrad of cryptochrome, with the fourth, W342, not observed in the peptidic products.

Some regions however do display rapid exchange (residues: 187-192, 210-214, 289-303, 333-336), >50% is indicated by a yellow-orange colour, even at the shortest time points with all regions involved in a loop or lacking secondary structure. Owing to residual H_2O after dilution, the relative uptake maximum should be considered 92 % and with back-exchange estimated to be between 20 - 30 % (§3.6.3), values of 62 - 72 % can be deemed close to saturation. Only three regions appear to reach the estimated saturation however, 289-293, 439-444 and 502-509. Of these the exposed 439-444 residues reach this limit after only 15 seconds, this region therefore is highly unlikely to display an increase in uptake when compared to a light state. Correlation between relative uptake and the degree of protectivity should not be over-interpreted as the intrinsic exchange rates for given peptides can vary significantly, depending on the local chemistry of nearby side-chains, see p. 50.

The uptake heat-maps for the dark-state at the shortest, 15 s, and longest, 1 hr, time points were mapped on to the X-ray crystal structure (PDB: 4GU5), Fig. 3.2B. The structural features observed in the crystal structure tend to correspond well with most buried regions maintaining slow exchange whilst loops and exposed regions more likely possess rapid exchange. The α22 helix displays a particularly large increase in uptake over an hour, 10.5 - 63.4 %, and is highlighted.

[iii]The fourth tryptophan W394 is not directly covered as it is the leading N-terminal residue in each observed peptide and as such its exchange cannot be directly measured. Flanking regions either side of W394, however, both display little exchange over all time points and from this we can infer little change at the Trp site also.

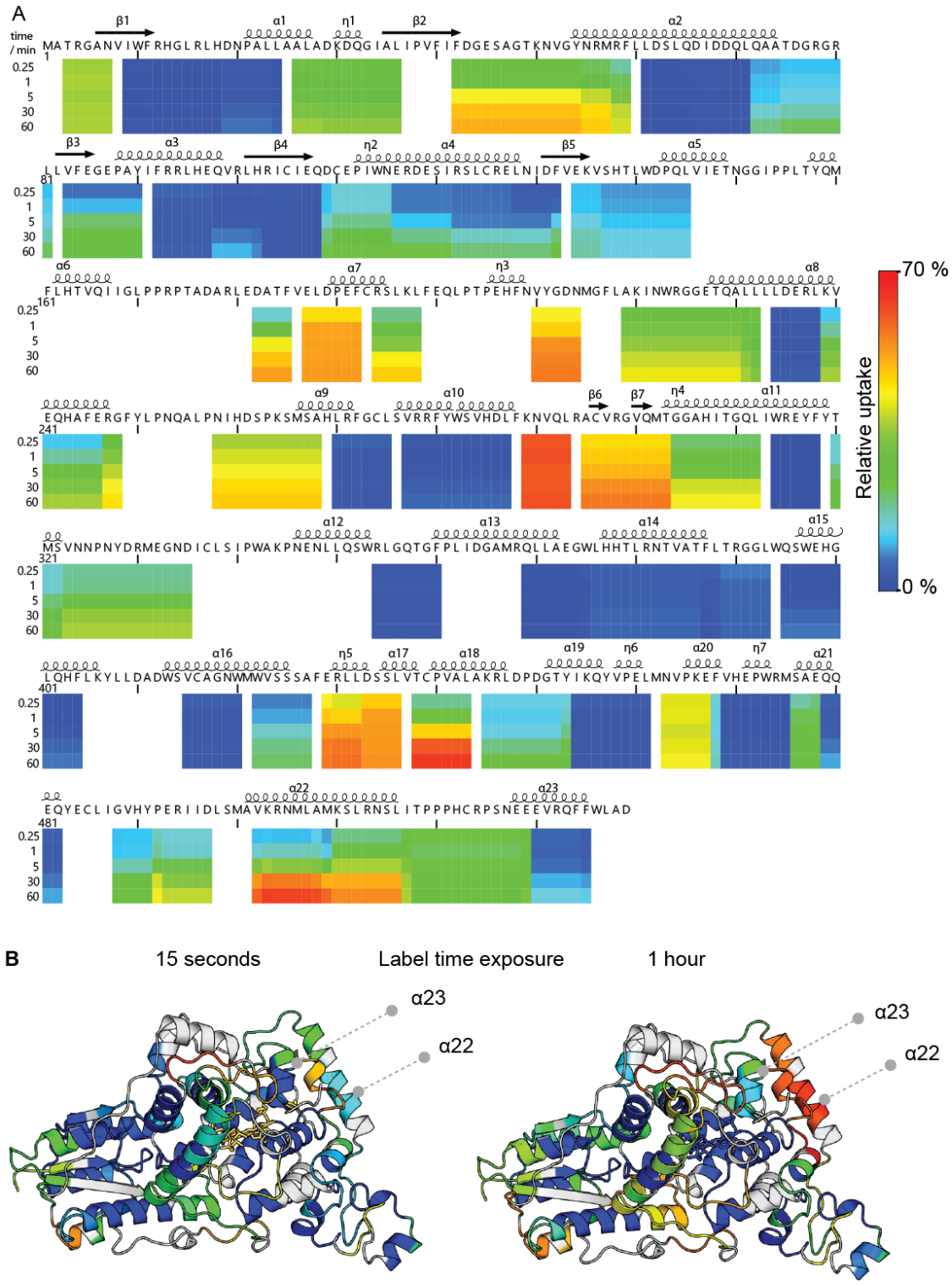


Figure 3.2. Dark state dynamics of *Dm*-Cry. A The deuterium uptake heat-map for *Dm*-Cry under dark conditions. HDX-MS was performed on *Dm*-Cry after exposure to a deuterated buffer solution for 0.25, 1, 5, 30 and 60 minute incubation periods. Estimates of the uptake at each residue are shown having been normalised to the theoretical maximum and correspond to values between 0 % (blue) and 70 % (red) as indicated in the colour bar. Secondary structure elements, α-helices, β-sheets, 3_{10}-helices (η), are depicted above the sequence. **B** Uptake heat-maps are mapped on to the crystal structure (PDB: 4GU5) for the time points 15 seconds and 1 hour. C-terminal helices are marked for future reference. A coverage of 80.5 % was obtained after data processing with regions of no coverage shown in white.

S. A. Chandler

Light-induced conformational change

A differential HDX-MS experiment was then performed to determine which regions became exposed to a greater or lesser extent. Cryptochrome underwent an identical experiment as described previously with the addition of a fixed wavelength blue light, 450 ±10 nm, with a controlled intensity. Cryptochrome was exposed to blue light for the duration of labelling and a prior equilibration step using a purpose built LED "cage" surrounding the samples on all sides at a fixed distance. The labelling chamber was temperature controlled but an external temperature probe was introduced and found no temperature fluctuations were introduced due to the lighting rig.

Uptake levels were then determined for the dark and light states from data acquired within the same 12 hour period. Taking the difference between these two states then provides a means of observing relative degrees of protection when exposed to different external light conditions. The relative increase in *Dm*-Cry uptake after blue light exposure relative to the dark state is shown as a difference heat-map in figure 3.3A. Regions in bright red correspond to those regions where blue light induces a change that results in greater exchange. The first thing to note, however, is just how much of the structure does not display a significant difference, shown in dark blue. Close to 80 % of the covered structure did not change upon exposure to blue light after 15 seconds, defined as an uptake change of < 4.3 % which corresponds to 3 standard deviations. This decreased slightly to 72 % after 1 hour, for an uptake change of < 5.8 % also corresponding to 3 standard deviations. Some localised regions however see a large increase in uptake propensity on exposure to blue light, as much as 45 %. All of these regions are clustered close to the CTT, with the largest increases seen for short exposure times at the α22 and α23 helices. The differential HDX results were mapped on to the crystal structure for the 15 second time point, Fig. 3.3B. Example spectra for peptides from these regions are also shown when in darkness and blue light across the time course, 0.25 - 60 min. An undeuterated time zero, t_0, spectrum is shown in each case for

reference. The peptide spanning residues 530-535, α23 region, maintains very low uptake when in the dark over the entire duration of the experiment, while, when exposed to blue light rapid exchange is seen after just 15 s. The peptide 501-509 from a neighbouring helix, α22, does not display such behaviour as the dark-state uptake is seen to increase over time until converging with the uptake obtained in the light state throughout. These spectra explain the decrease in difference over time seen for the α22 peptides whilst the large differences for α23 is maintained. A minimum of three overlapping peptides (each in triplicate) for both states exist for each of the regions providing a degree of certainty of the behaviour which is discussed further on page 86.

A small number of localised regions do display notable increases after 1 hour between 25 and 34 % and correspond to peptides spanning residues 110-121, 389-393 and 422-428. The latter has a relatively large initial difference of 20 % rising to 28 % after 1 hour. This region corresponds to the N-terminal portion of the C terminal lid, Fig. 3.1A, that is also situated close to W420 of the Trp-tetrad. Residues 389-393 exist as a short loop prior to the most terminal Trp-tetrad residue, W394, which is buried by its neighbour α22. That α22 and α23 become less structured may allow for the 389-393 and 422-428 regions to become less protected manifesting as an increase in uptake. It cannot be ruled out that these regions also become disordered, though given they both support the photo-sensitive Trp-tetrad this may be unlikely. All of the observed increases are located close to the CTT in space and highly localised to the photo-sensitive core, with the exception of residues 110-121 which is distal to all of these conformeric motions.

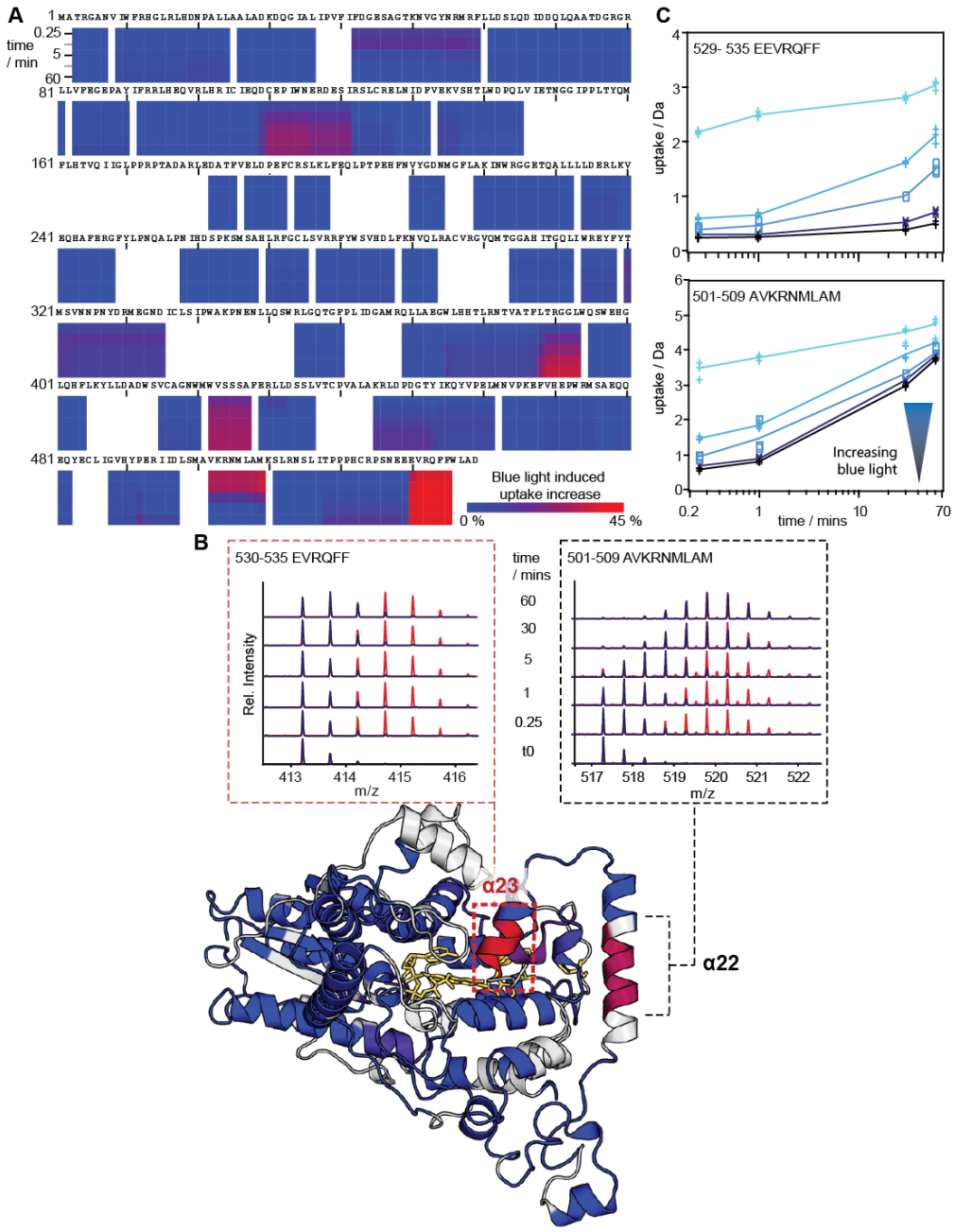


Figure 3.3. Light-induced conformational change in *Dm*-Cryptochrome Differential HDX-MS reveals a blue light-induced conformational change localised to the C-terminal tail. **A** A differential uptake heat-map for *Dm*-Cry where the positive increase corresponds to greater uptake for the light environment. Regions which become significantly less protected are shown in red, corresponding to a relative change of 45 %. Each residue has five associated time points: 0.25, 1, 5, 30, 60 minutes. **B** The difference heatmap is mapped on to the *Dm*-Cry crystal stucture (PDB: 4GU5) after 15 s of labelling. Helices α22 and α23 are highlighted with uptake spectra corresponding to peptides spanning these regions shown. Dark spectra are shown in blue with those exposed to light in red over the time course 0.25 - 60 min. Undeuterated time zero, t_0, spectra are shown on the bottom for reference. The proximity to the bound FAD and the Trp-tetrad (W342, W394, W397 and W420) is clear, shown in yellow. Regions with no coverage are shown in white. **C** Plots showing the blue light dependence on the amount of uptake for two peptides, 501-509 and 529-535, from the α22 and α23 regions identified to have a large induced change. Uptake is plotted under each light intensity between 0.25 - 60 mins where the lighter blue colour corresponds to greater light intensity.

S. A. Chandler

By adjusting the blue light intensity we explored the sensitivity of *Dm*-Cry to photo-activation rates, Fig 3.3C. The absolute uptake is plotted against time as a function of increasing light intensity for a peptide in each of the α22 (501-509) and α23 (529-535) regions. We observe that lower light intensities yield lower uptake across both regions, showing that the photo-activation does not yield a binary response but that the rate of the conformational change is sensitive to the external stimulus. The α22 peptides are seen to converge faster towards the maximum of ∼4.7 Da in the presence of a more intense light whilst the α23 peptides show little sign of reaching the upper limit of 3 Da. The terminal tip of the CTT appears most sensitive to the intensity of light.

The bound FAD and tryptophan residues associated with the tryptophan photochemical tetrad, are highlighted in yellow and extend from the core outwards towards the C-terminal tail α22 and α23 helices. Blue light exposure causes little deprotection of the regions flanking the observable photochemical tryptophans. Earlier it was shown that these regions were well protected over 2 orders of magnitude of time and now, after exposure to blue light, they still exchange slowly. These data suggest that the core has few structural perturbations despite the photo-activation triggering a conformational change localised at the C-terminus. It is conceivable that the emergence of charged species as a result of photo-activation could result in an increase in intrinsic uptake rates without a conformational change. However, as the photochemical core, those residues harbouring the induced charge and most prone to the electronic influence on uptake, remains protected one can likely rule out the direct effect of electrostatics on the uptake differences observed. The crystal structure shows the C-terminal regions to be helical with a bridging loop between α22 and α23. The decrease in protection would imply there is a conformational change in which the CTT is released with increased flexibility and/or the helices begin to locally unfold. That the dark-state peptides of α22 were capable of converging to the same observed uptake as for those in blue light suggests light promotes the conformational change at these sites. The peptides associated with the final helix, α23, have two distinct uptake populations corresponding to the dark

and light states, where even after an hour the dark uptake does not reach that seen after light exposure. The kinetics of these conformational changes are discussed later but show the sensitivity HDX-MS has to localising a conformational change and also resolving the structural kinetics.

Conformational specificity towards blue light

FAD absorption of light wavelengths greater than 500 nm is very low and so red-light-induced photochemistry was thought to be unlikely. To probe the difference in light activation between blue and red wavelengths, we performed the same differential experiment under a comparable intensity of red wavelength light. The difference between the dark and red light exposure was plotted on a per peptide basis across the sequence for each time point, Fig. 3.4A. A positive value corresponded to an increase in uptake due to exposure to red light. Very small deviations do exist for some individual peptides but within a very narrow range –5.4 - +3.8 %, with the vast majority of peptides limited to ±2 %. These values are an order of magnitude lower than the +40.9 % observed for the α23 peptides in blue light. In addition, very little systematic uptake behaviour is observed between sequential time-points. All peptides have a relative difference error of less than ±6.5 %,whether between dark-red or dark-blue, with the error distributions shown in Figure 3.4C. The large majority of peptides have an error as low as ±1 % with similar distributions for those exposed to red or blue light, but blue light does have a greater proportion of peptides with larger errors than in red light. The magnitude of change in uptake for the red light exposure is of the same order as the median error ∼1 %, whilst the magnitude of change under blue light is significantly higher.

Two peptides shown previously, from α22 and α23, are highlighted that showed a large increase in uptake when exposed to blue light, Fig. 3.4B. The absolute uptake is plotted for *Dm*-Cry peptides after exposure to the dark, red light and blue light between 0.25 - 60 minutes. It is clear that red light exposure very closely resembles the dark state suggesting it is the blue wavelength that is crucial to trigger the conformational change. These data support the idea that blue light activation

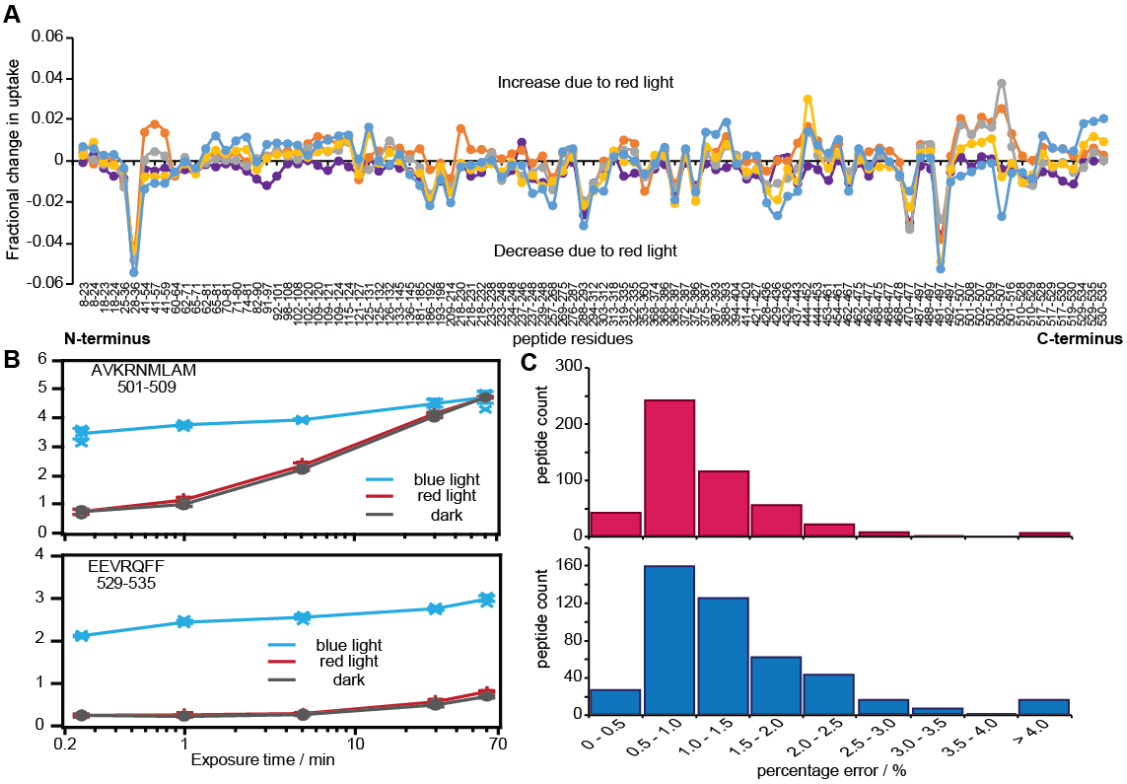


Figure 3.4. Conformational specificity towards blue light Performing the same differential local HDX experiment between a red light stimulus and darkness shows no significant change. **A** A plot of the change in uptake upon red light exposure for each identified peptide. Peptides are loosely arranged from the N-terminus to the C-terminus from left to right, with the residues indicated. Each peptide is plotted for the following exposure times and coloured accordingly: 0.25 (purple), 1 (orange), 5 (grey), 30 (yellow) and 60 minutes (blue). **B** Absolute uptake of peptides from the α22, 501-509, and α23, 529-535, regions under dark, red light and blue light conditions, coloured black, red and blue respectively. **C** Error distributions of all peptides and time points for the red-dark comparison (top) and the blue-dark comparison (lower). The histogram has a bin size of 0.5 % corresponding to the relative difference error. A total of 505 peptide time points are present for the red exposure and 465 for the blue exposure. Changes observed in **A** are on the order of the red light errors whilst the changes seen in figure 3.3 are significantly higher than the blue light errors.

causes a localised conformational change and not a global unfolding event. Peptides in other regions of the protein act as internal calibrants as they do not display a change of the magnitude seen for the C-terminal tail. The red light confirms this and shows that the same C-terminal tail, and indeed all other regions of the protein, are not perturbed by a red light stimulus.

Extracting EX1 kinetics

A bimodal uptake envelope is indicative of two distinct, interconverting conformations. If a slow exchanging, or exchange incompetent, conformation interconverts

with a more "open" fast exchanging population on a time-scale much slower than the open HDX exchange rate, two uptake profiles will be observed for the corresponding peptide region. This is termed EX1 type kinetics in HDX-MS where, $k_{\text{close}} << k_{\text{ex}}$, and provides a means of obtaining the relative populations of the closed a open states in addition to k_{open}. This is described further in the Chapter 2 on page 52. By treating the isotope distributions of the two populations as two distinct binomial distributions with different probabilities of exchange one can fit the isotopic peaks and thus extract the relative intensity and probability of exchange. This procedure was performed using the HX express 2 macro developed by the Engen group.[29,30]

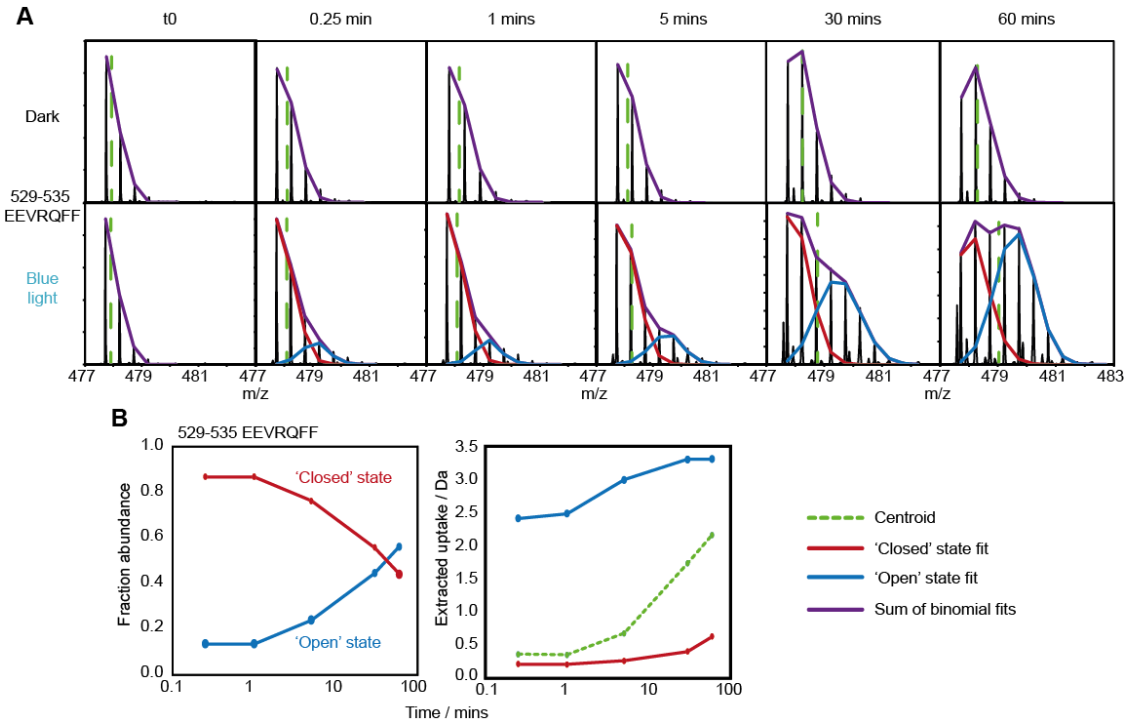


Figure 3.5. EX1 fitting of C-terminal peptide spectra. A Binomial fits of a C-terminal peptide, 529-535, under dark (**top**) and blue light (**lower**) conditions after 0.25 to 60 min of labelling. A slow exchanging "closed" distribution (red line) and a fast exchanging "open" distribution (blue line) are required to adequately fit the blue light spectra. The sums of the two binomial distributions are shown in purple and provide a good fit, a reduced chi-squared per degree of freedom (χ^2_ν) of 3 for all data and $\chi^2_\nu = 0.9$ for labelling times above 1 min, further details provided in the Methods section. Weighted centroids, reported in previous sections, are shown in a dashed green line. HX Express 2[29,30] was used for spectral fitting as described in the text. **B** Extracted relative abundances of the open (blue line) and closed (red line) states for each time point, **left**. The centroid of the extracted open and closed state was plotted for each exposure time, **right**. The centroid shift is shown in green and is a result of the increase in the fast exchanging open state abundance.

Spectra were fitted for the dark and blue light states of *Dm*-Cry for all peptides

in the α22 and α23 regions, interestingly the only peptides that displayed significant bimodal behaviour, Fig. 3.5. The dark state peptides are very well described by a single binomial distribution as demonstrated for the peptide 529-535. The corresponding weighted centroid is shown as a dashed green line. However, two binomial distributions corresponding to an open and closed state are necessary to fit blue light data. The open distribution is shown in blue and the closed in red with the sum of the two fits shown in purple. The sum can be seen to accurately represent the spectra for all time points. It can be seen from these fits that the blue light closed state distribution corresponds well to the single dark state distribution over time. The relative populations of each state were extracted from these fits for the open and closed states, Fig. 3.5B (left). The closed state is seen to decrease from an initial population of ∼0.85 to ∼0.45 after 1 hour with little difference between 15 s and 1 min. The centroid from each distribution in the light-exposed peptides can also be plotted as a function of time. The closed state is seen to have a minor increase in uptake from 0.2 Da to 0.6 Da over an hour. The open state sees a greater, but still small, increase from 2.4 Da to 3.3 Da. As the two distributions display small increases the peptides must also be coupled to low level EX2 behaviour. The total centroid is also plotted as a dashed green line for reference. The large centroid shift observed for these peptides in blue light is due to the population increase of the fast-exchanging open state, rather than the increase in uptake for each of the two states. These fits provide a sensitive measure of the abundance of the local open population at each time point.

This analysis was performed for all peptides, in triplicate, in the α22 (501-509) and α23 (529-535) regions and the relative amount of the "open" state was plotted as a function of time, Fig. 3.6B (left). All peptides are seen to increase exponentially (Figure 3.6 plots are on a logarithmic x-axis) from an initial abundance of ∼0.3 for α22 and ∼0.2 for α23. A simple exponential rising from zero to unity did not result in a good fit for any of the peptides. As all peptides appear to start with an initial open population, an initial y-intercept value was allowed resulting in a very good fit (R^2 = 0.9991, 0.9947, 0.9952 for 501- 507, 508 and 509 respectively)

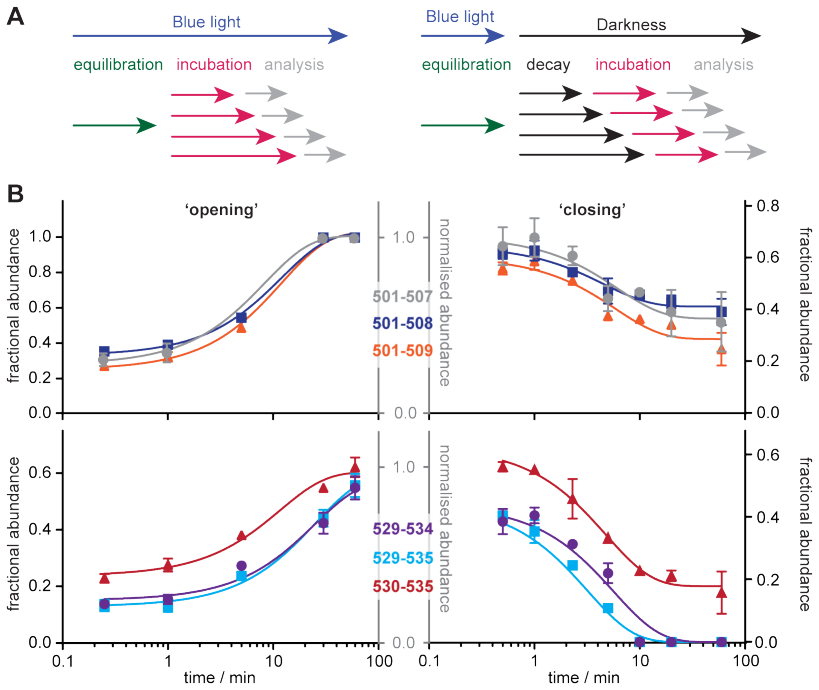


Figure 3.6. Kinetics of the light induced conformational change. EX1 kinetics of the C-terminal conformational change measured by local HDX-MS. **A** Experimental HDX-MS time sequence for measuring the "opening" (**left**) and "closing" (**right**) kinetics of *Dm*-Cry. Opening experiments involve a fixed equilibration in blue light after which sample is incubated under blue light in a deuterated buffer for a variable time prior to analysis, sample remains in blue light throughout. The "closing" experiment has a fixed equilibration step under blue light of 20 minutes after which the sample is kept in darkness for a variable time. The sample is then incubated under darkness in a deuterated buffer for a fixed 15 seconds and analysed in the dark. **B** The amount of *Dm*-Cry open state, determined as described in Figure 3.5, after an opening or closing experiment is plotted as a function of time for the α22 (501-507, 501-508 and 501-509) and α23 (529-534, 529-535 and 530-535) peptides. Each plot has the absolute fractional abundance shown but is plotted having been normalised to the highest abundance in each plot, the normalised axis is shown in the centre coloured in grey. Data were fitted to a single exponential with a variable intercept and plateau to account for the lack of saturation seen in most peptides. Quality of fits and the extracted rate constants for opening and closing are reported in the text.

for the 501-509 peptides with the entire population reaching an open state after 30 minutes. The α23 population never became fully "open" and reached a maximum for all peptides of ∼0.6, the exponential fit for this region required an additional "plateau" parameter to account for not reaching saturation (R^2 = 0.9733, 0.9914, 0.986 for 529-534, 529-353 and 530-535 respectively). These data indicate that a small proportion of *Dm*-Cry protein exists in an open state even in darkness, though this was not detectable from the acquired dark state spectra and is only based on the observed fits. In addition, the α22 helix is able to become fully open whilst the α23 region reaches a maximum close to 60 %.

The rate of opening can be extracted from these fits with the α22 having an average rate constant of opening of 0.096 ± 0.02 min^{-1}. Whilst the C-terminal tip has a slower rate of opening, with the 529-534 and 529-535 peptides having rate constants of 0.044 ± 0.02 min^{-1} and 0.040 ± 0.01 min^{-1}, respectively. The 530-535 residue peptide has a comparable rate constant to the other α22 peptides, 0.084 ± 0.03 min^{-1}. That two of the peptides have a lower rate constant may reflect the limitations in the EX1 fitting performed but it is also encouraging that we were able to extract local kinetics comfortably to within an order magnitude. It could equally be that residue 529 is located in a region which remains more "closed".

Return to the dark state

Exploitation of the C-terminal EX1 kinetic behaviour also allowed us to monitor the relaxation of the open state back to the closed state. If rather than maintaining a constant exposure to blue light and varying the label incubation time, Fig. 3.6A (left), we equilibrate *Dm*-Cry under blue light and turn the light off after different time intervals we can watch the decay of the open state population, Fig. 3.6A (right). This was performed by equilibrating the sample for 20 minutes in blue light and then turning the light off in a dark environment before performing a 15 second labelling experiment. The abundance of the open state is shown in Figure 3.6B (right) for the same two C-terminal regions. Plots have been normalised, central axis shown in grey, to the highest abundance observed in each of the regions for comparison with the "opening" experiment. The absolute fractional abundance is also shown on the right axis. The abundance of the open state is seen to decrease, also exponentially, but generally not returning to zero. The 501-509 peptides plateau to a value between 0.25 - 0.4 which is similar to the levels of the initial open state in the forward "opening" experiments. Work by others also notes this finding, reporting ∼37 % of the light state still remaining with the decay back to the dark state following a bi-exponential function.[13] The 529-535 region also displayed similar behaviour when closing as with opening. The 530-535 region displayed more open character than the other two peptides, this is likely the time left for equilibration as the

starting values in the relaxation experiment are comparable to those obtained in the opening experiment after 20 minutes. The 530-535 peptide also decays to a value of ∼0.2 which mirrors the initial abundance seen for this peptide in the opening experiments. Unusually, the 529-534 and 529-535 peptides show entirely closed conformations after 10 - 15 minutes, whereas, we observed a fraction of around 0.15 in the forward experiment. It should be noted that the fit quality is lower for these peptides with large errors (R^2 = 0.9433 and 0.9277 for 529-534 and 529-353 respectively), though 530-535 remained high, R^2 = 0.9925.

Calculation of the closing rate constant gives values of 0.16 ± 0.06 min^{-1}, 0.22 ± 0.05 min^{-1} and 0.17 ± 0.06 min^{-1} for the α22 peptides, 501-507, 501-508 and 501-509 respectively, and 0.12 ± 0.26 min^{-1}, 0.30 ± 0.12 min^{-1} and 0.20 ± 0.02 min^{-1} for the α23 peptides, 529-534, 529-535 and 530-535 respectively. Despite the high errors seen for two of the peptides the rate constant for closing is larger than that of opening in all cases, reflecting approximate half-lives for opening of 5-10 min and closing <5 min. Both regions demonstrate that after photo-activation *Dm*-Cry is able to maintain its open conformation on the order of minutes. These experiments quantify the long-lived nature and reversibility of the C-terminal conformational change. This data is supported by some previous studies into *Dm*-Cry in which the authors showed a half-life of the signalling state *in vivo* to be around 15 mins and the half-life of FAD reoxidation to be comparable.[22]

3.3.3 Conclusion

HDX-MS has conclusively determined those regions which display a blue light dependent conformational change. We have built upon low resolution experiments presented[22,23] previously offering a more localised quantitative approach. We confirm the role of the CTT in a light-induced conformational change and demonstrate it is localised to α22 and α23 after short exposure times and encompassing surrounding regions proximal to both the CTT and the Trp-tetrad. These results are highly consistent with the limited proteolysis approach presented by Ozturk *et al.*[23] These same regions were corroborated by Vaidya and co-workers,[22] though they also

discovered regions to be more prone to proteolysis that Ozturk *et al.*, and the work presented here, do not support. They report a number of regions distant from the CTT with the inference being a global structural change has occurred. In the work presented here only one small localised region, 110-121, away from the CTT is observed suggesting the core PHR remains conformationally constrained.

We have additionally quantified the relative population of the "open" state for each of the α22 and α23 regions using EX1 kinetic extractions. In doing so we were able to calculate the opening rate constant at a local level and note it has a half-life of between 5-10 minutes depending on the peptide observed. Decay experiments were performed to monitor how quickly the light state returns to the dark. Rate constants for the opening appeared to be faster than that of closing and importantly demonstrates that the process is largely reversible. Other work suggests that Cry adopts a sacrificial role in its signal transduction in circadian regulation and suggesst an irreversible change.[31] We, and others, note that light exposure introduces instabilities and a small degree of precipitation is observed.[16] In our work, however, we demonstrate that the remaining solution population is capable of reversibility, recovering >60-80 % of the dark state.

In this work we have demonstrated that the CTT release is light sensitive and produces a long lived state capable of a degree of reversibility, yet little is known exactly how the CTT is ejected from its binding pocket close to the FAD cofactor. Recent molecular dynamics simulations suggest that upon photo-activation histidine-378 which is sandwiched between FAD and α23 becomes protonated owing to an increased pKa caused by the light induced reduction of the neighbouring FAD.[32] The charge site is thought to mediate the release of the CTT allowing down-stream interactions to form. This mechanism does not however appear universal across cryptochrome species with molecular dynamics simulations in European Robin Cry 4 displaying little difference at the corresponding histidine site.[33]

The HDX-MS experiments presented here demonstrate the power of the technique to sensitively probe the light induced conformer states of the elusive

cryptochromes. This work paves the way for further experiments in exploring the role of the Trp-tetrad and the susceptibility of cryptochromes to external magnetic fields.

3.4 Part II : Protein-protein interface determination

3.4.1 Biological context

Many classes of protein assemble into multimeric complexes, and whilst many cellular functions are known to rely on such structures, the origins of oligomerisation and the evolutionary pathway often remains unclear. Haemoglobin is an archetypal beneficiary of complex formation aiding function through subunit allostery. Haemoglobin is an essential oxygen transporter to almost all modern-day vertebrates and assembles to form a tetrameric complex.[34,35] The highly helical globin fold, consisting of 7-8 α-helices, of each subunit is common to all extant haemoglobins and the related cousins, myoglobin. The conserved fold provides a binding pocket for an iron-heme complex, crucial for oxygen binding, at its core. In mature humans, the complex exists as a dimer of dimers, $\alpha_2\beta_2$, consisting of two sets of closely related α and β chains. Each α subunit has a strong interaction, a *packing* interface, with a β subunit to form the heterodimers, termed $\alpha1\beta1$ and $\alpha2\beta2$. These two symmetric heterodimers interact via a *sliding* interface between unlike α and β chains, $\alpha1$ and $\beta2$ (also $\alpha2$ and $\beta1$), Fig. 3.7A.

The tetrameric complex is well known to have allosteric properties to promote oxygen binding in neighbouring subunits.[36–38] The cooperativity is brought about through its existence as a hetero-oligomer yet little is known of how extant haemoglobin evolved from monomeric globins and how the two distinct interfaces came to be. It is known that the α and β forms diverged from a common ancestor, Anc, around 450 million year ago.[39,40] Our collaborators from the Thornton Laboratory, University of Chicago, used ancestral protein reconstruction (APR) a computational approach using phylogenetic techniques to reconstruct statistical approximations of the ancestral haemoglobin proteins.[41–44] Following this analysis they expressed, purified and structurally examined Anc. A multiple sequence alignment of the α and β haemoglobins and their last common ancestor, Anc, is shown in Figure 3.7B.

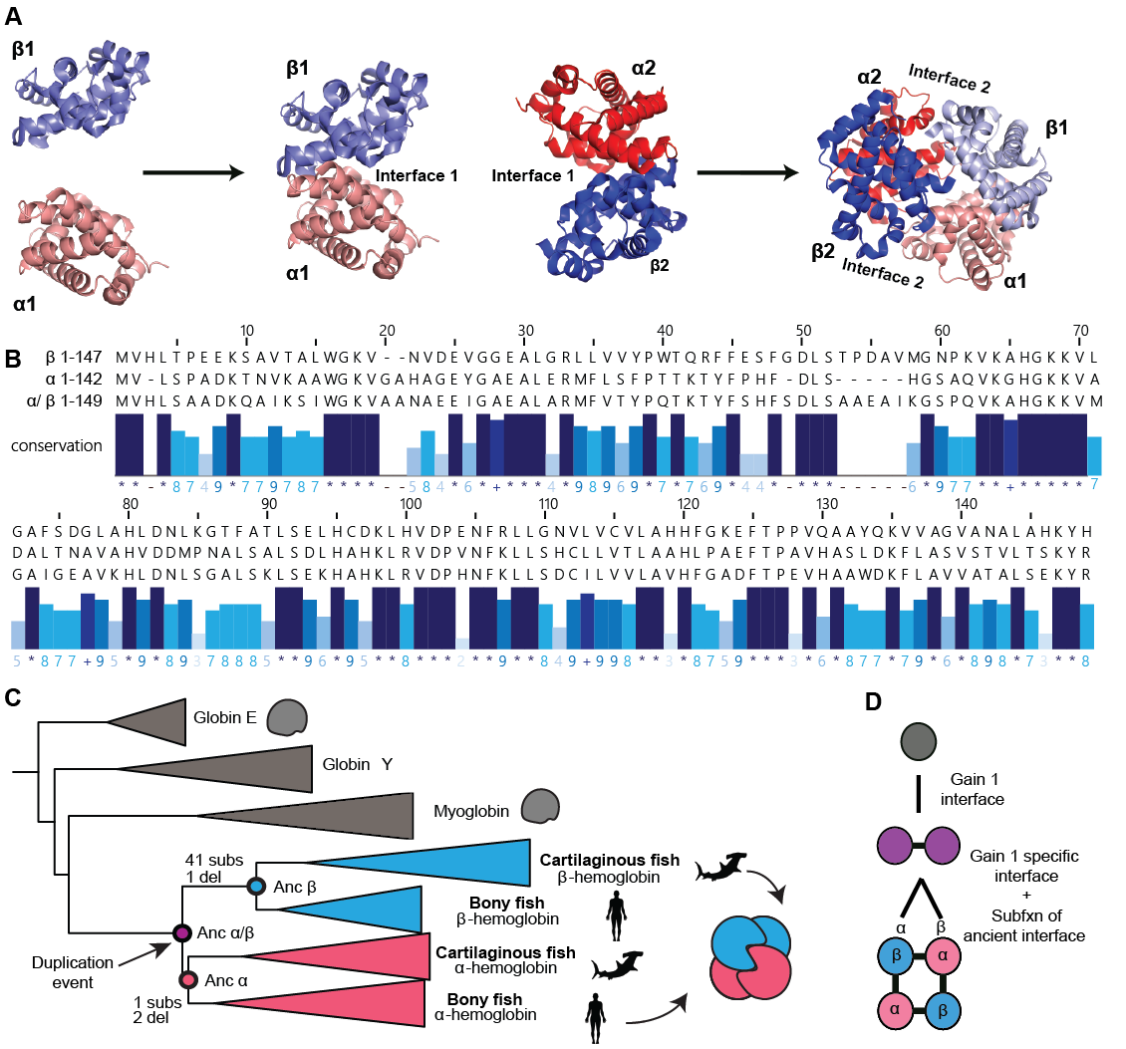


Figure 3.7. Oligomeric origins of extant haemoglobin. A Quaternary structure of extant human haemoglobin. A dimer *packing* interface, interface 1, between α and β subunits is first indicated. Each dimer interacts via a secondary *sliding* interface, interface 2, with an unlike chain to form a tetramer, $\alpha_1\beta_2$. **B** Sequence alignment of the α/β, α- and β- haemoglobin subunits from humans. The degree of conservation is indicated below with an * denoting fully conserved residues and + residues where a residue differs but its physio-chemical properties remain conserved. The darker the blue the more conserved the systems are. **C** A phylogenetic tree of globin protein classes determined from APR. The last common ancestor of the α/β chains is indicted in purple. The ancestral forms, that ultimately diverge to give extant haemoglobin, of each α and β are shown in blue and pink respectively. **D** The hypothesised evolution from a monomeric globin to the tetrameric extant haemoglobin via a dimeric common ancestor of α and β. All phylogenetic analyses and studies were performed by the Thornton lab, who are identified in the acknowledgements, with all hypotheses provided entirely by them. The sequence alignment was performed by ourselves. Arvind Pillai provided panels A,C & D of this figure with only aesthetic changes made.

The degree of conservation between the species is shown with a darker blue corresponding to higher conservation. Their initial analytical size exclusion chromatography (SEC) data suggested Anc existed predominantly as a dimer. We

verified these findings using native MS to find both monomeric and dimeric populations, discussed in greater detail in § 3.4.2. Since native mass spectrometry (MS) measurements were performed in the low μM range and typical haemoglobin concentrations in human erythrocytes reach near 20 mM, this would comprise a predominantly dimeric system.[45] That Anc exists primarily as a dimer suggests it gained a single interface from evolution of the monomeric ancestor. If the homodimeric Anc subunits then diverged to form discrete α and β units that tetramerise, it then follows that the existing dimeric interface subfunctionalised whilst elsewhere gaining a new specific interface. The last common ancestor of α-haemoglobin of cartillaginous and bony fish, Anc α, underwent a single substitution and 2 deletions from Anc. Whilst the last common ancestor of β-haemoglobin of cartillaginous and bony fish, Anc β, underwent a single deletion and 41 substitutions, Fig 3.7C.

The Thornton lab engineered the ancestral proteins and showed the progression from a monomeric globin to dimeric Anc, which after a duplication event gave rise to the α and β subunits capable of forming tetramers. This progression led to the hypothesis that Anc dimers are formed by a precursor to interface 1, in the extant haemoglobin tetramer. The emergence of the weaker interface 2 then came about after the gene duplication of Anc. A schematic of this hypothesis is presented in Figure 3.7D.

In this section we use differential H/D exchange (HDX) MS and native MS to determine the likely interface formed in the dimerisation of the last common ancestor of α- and β-haemoglobin, in order to validate the hypothesis of interface evolution.

3.4.2 Results and Discussion

Native mass spectrometry of ancestral haemoglobin

Mutant systems were engineered to selectively disrupt the proposed interface 1 in Anc dimers: firstly, a single point mutation of P127R, and, secondly, 7 mutations selected to revert the protein back to its last common ancestor with its monomeric

myglobin cousins, a reversion mutant (Rev). We acquired native mass spectra of these systems to ascertain the oligomeric assembly. Spectra are shown in Figure 3.8A for Anc and the two mutants at a concentration of 10 μM. The ancestor shows an equilibrium of monomer and dimeric states centred around the charge states 8+ and 11+ respectively. Almost half of the species observed are dimeric with approximately 64 % of subunits existing as a dimer at this concentration. Spectra also show that from an average of 3 repeats around 6 % of monomer heme sites are unoccupied and 2 % for the dimer. Spectra were acquired using a quadrupole filter to prevent transmission of low m/z species to monitor any gas-phase heme dissociation, shown in the inset. Very low ion counts corresponding to heme B, 617.27 m/z, are recorded as a result of gas-phase loss at these low energies, please note the axis is not normalised but reflects the relative intensity at a maximum of 0.8 %. Both P127R and the reversion mutants, however, show an exclusively monomeric distribution also centred around the 8+ charge state. That the charge states are comparable to the folded ancestor suggests the mutations have not induced any appreciable unstructuring. The reversion mutant saw a negligible amount of *apo* state but P127R saw as much as 15 % of the monomeric state without a bound heme. Once again, a quadrupole selection shows little gas-phase dissociation is at play, a maximum of 1.5 %, rather suggesting a lack of binding in solution. Appreciable gas-phase loss was not observed for any of the Anc or mutant species until between 15 - 30 V was applied.

Native MS was also used to help characterise the equilibrium behaviour of Anc as a function of concentration. The concentration was increased over a narrow range amenable to native MS, here 10 - 30 μM, Fig. 3.8B. Upon increasing the concentration a small shift towards the dimeric population is observed. The solution heme loss becomes clearer at higher concentrations revealing a partially occupied dimer and a small population of completely unoccupied sites. The percentage of apo sites for the monomer and dimer were: 6 % and 2 %, 12 % and 5 %, 15 % and 15 %, at 10 μM, 20 μM and 30 μM respectively. The trend worth noting for future HDX-MS studies, albeit for these three concentrations, is that as the Anc

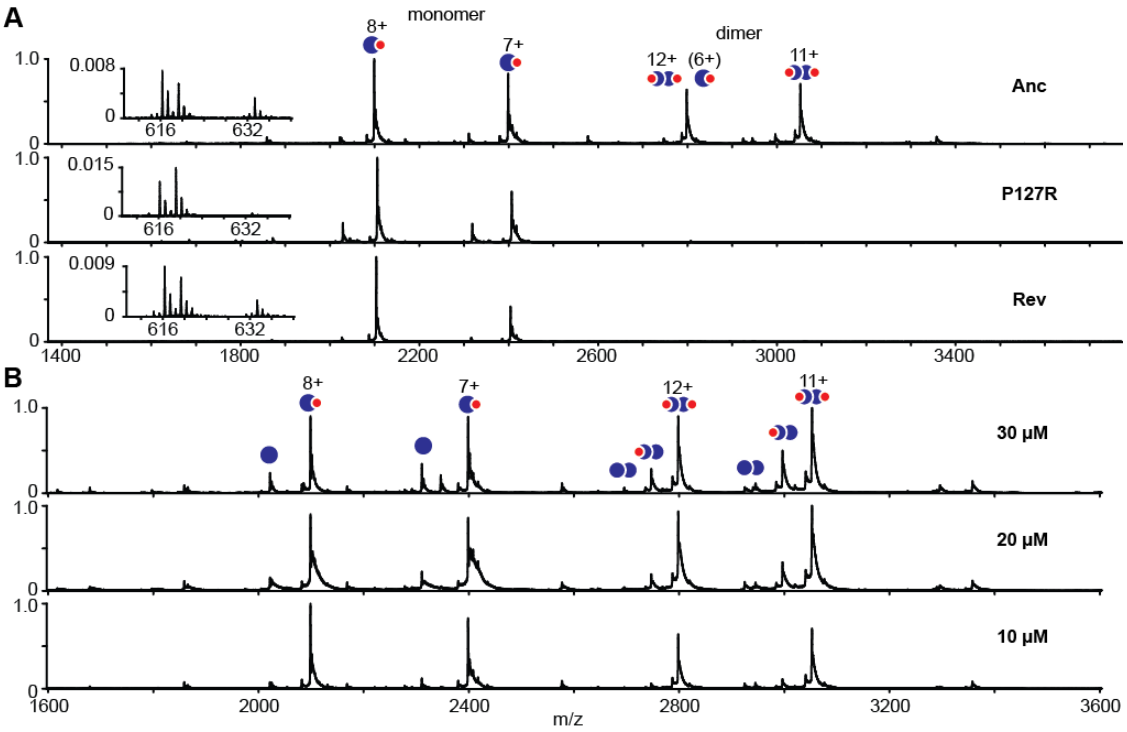


Figure 3.8. Native mass spectra of ancestral haemoglobin and mutants. A Native mass spectra of Anc (top), and, P127R (middle) and Rev (lower) mutants at 10 µM. Anc exists as a monomer and dimer with the heme sites fully occupied. Spectra were acquired with a quadrupole selection and the resulting low level gas-phase dissociation of heme is shown as an inset, Note the intensity axis is a maximum of 1.5 %. Monomeric subunits are indicated in blue with bound heme shown in red. No dimer was observed for either mutant. **B** Native mass spectra of Anc were acquired over the indicated concentration range, 10 - 30 µM. At higher concentrations more apo heme sites are observed, indicated by the schematics.

concentration is increased, ligand binding becomes less favourable and leaves the binding pocket less shielded to HDX. This would mean that when we compare the "dimer" to the "monomer" systems, the dimer may have increased exchange in the binding pocket region relative to the monomer. Finally, an approximate dissociation constant for the monomer-dimer, K_d, was calculated to be ∼15 µM from native MS experiments as described in the Methods section.

Native MS experiments have validated the dimerisation of Anc subunits and verified the expected disruptive nature of the mutants. The observed charge states are consistent with a folded system[46] and are consistent between the wild type and mutants. Additionally, native MS has shown that there is a small concentration dependence on heme binding. These results indicate that the ancestral haemoglobin created, is natively folded and that mutants pertaining to the modern day interface

1 disrupt its oligomerisation. Further validation using a technique more sensitive to tertiary and quaternary three-dimensional structure, HDX-MS, is described in the following section to further localise the observed interface.

Mapping uptake differences for monomeric and dimeric ancestral α/β haemoglobin

To localise the interfacial residues in the dimeric ancestral haemoglobin we performed a differential HDX-MS experiment comparing the solvent accessibility of the dimer and the free monomer. Regions found at the interface are likely to have reduced uptake due to greater steric protection and an increased number of amide interactions offering protection for backbone hydrogens. The two mutants, P127R and reversion, provide a means of obtaining exclusively free monomers but have the disadvantage that the intrinsic rate of exchange may vary with a change in residue but also structural perturbations may be present as a result of the mutants. Careful control of the ancestral wild type concentration provided a means of accessing predominantly the monomeric or dimeric states by shifting the equilibrium. Operating at 0.67 μM would result in almost entirely monomer (92 %) whilst working at 75 μM means we can obtain 74 % dimer. This provides a means of obtaining an uptake difference with identical primary structures eliminating any artificial structural distortion due to mutations. Uptake differences are likely to be smaller however as the monomer and dimer are in dynamic equilibrium meaning both states are never exclusive, which is not the case for the mutant systems. Together, however, through comparison of the uptake of high concentration wild-type to low concentration WT and the mutants systems we have a means of comparing the uptake at a local level.

An on-line peptic digestion resulted in a good sequence coverage, comparable to that previously reported for modern day haemoglobin tetramers,[47] consisting of 47 unique peptides giving 91.3 % with an overlap redundancy of 5.3, Fig. 3.9. A number of long peptides were obtained which limit our ability to localise uptake differences and make linearising the sequence troublesome with existing methods. Peptides were therefore limited to 25 amino acids, still a relatively high threshold,

but short peptides no longer than 10 residues account for 60 % of the sequence. Point mutations are highlighted for the reversion mutant, red, and P127R, purple. Considerations of the impact of these residues on the resulting peptide uptake are discussed on page 103.

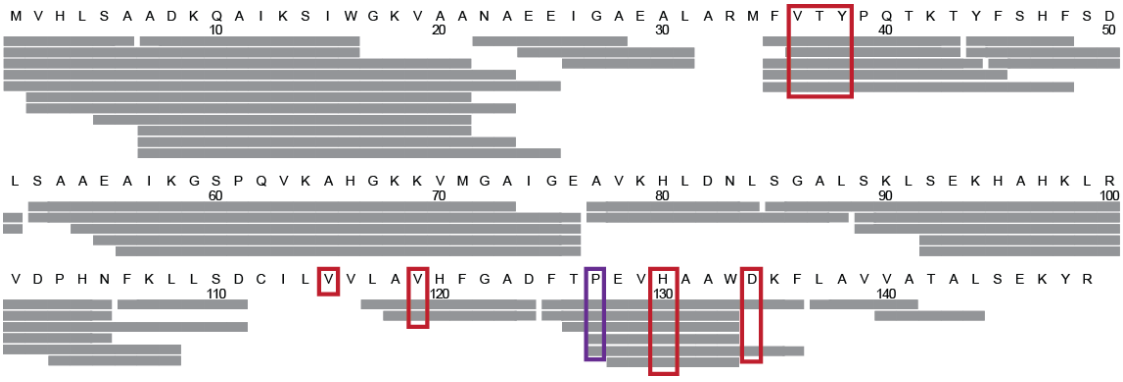


Figure 3.9. Pepsin digest of ancestral haemoglobin. An on-line pepsin digestion, at pH 2.4 and 20 °C, was performed on ancestral haemoglobin with resulting peptides indicated below the wild type sequence. A coverage of 91.3 % was obtained consisting of 47 unique peptides and a redundancy of 5.3. Residues that underwent mutation in the P127R and Rev mutants are indicated by purple and red boxes, respectively.

Initially, a local HDX-MS experiment was performed, as described previously and in the methods, to obtain the deuterium incorporation for dimeric Anc haemoglobin and the corresponding monomeric mutant systems, P127R and Rev. Anc haemoglobin was diluted into a deuterated buffer to give a final concentration of 75 μM, resulting overwhelmingly in dimer. Uptake spectra were acquired between 0.25 - 60 minutes and spectra are shown for three peptides in figure 3.10. Spectra are shown for Anc dimer (blue) alongside P127R (red) and/or Rev (purple) haemoglobin; N.B. the m/z axis has been shifted for effective comparison of isotopic envelopes. The peptide spanning residues 1-16 shows very little difference in uptake distribution between the dimer and mutant species and is typical of most peptides. Whereas, for the 35-45 peptide differences do begin to emerge between the Anc dimer and P127R, Rev is not shown as this peptide features a mutant in this region. Spectra obtained were of high signal-to-noise and high reproducibility, so whilst these differences only correspond to a 0.45 ± 0.04 Da shift after 15 s and 0.74 ± 0.08 Da after 1 hour, errors remain small. In addition as the absolute uptake of the 35-45 peptide had an

absolute uptake of 3.0 ± 0.03 Da and 4.7 ± 0.06 Da after exposures of 15 s and 1 hr respectively, these correspond to percentage increases of ∼15 % and ∼16 %.

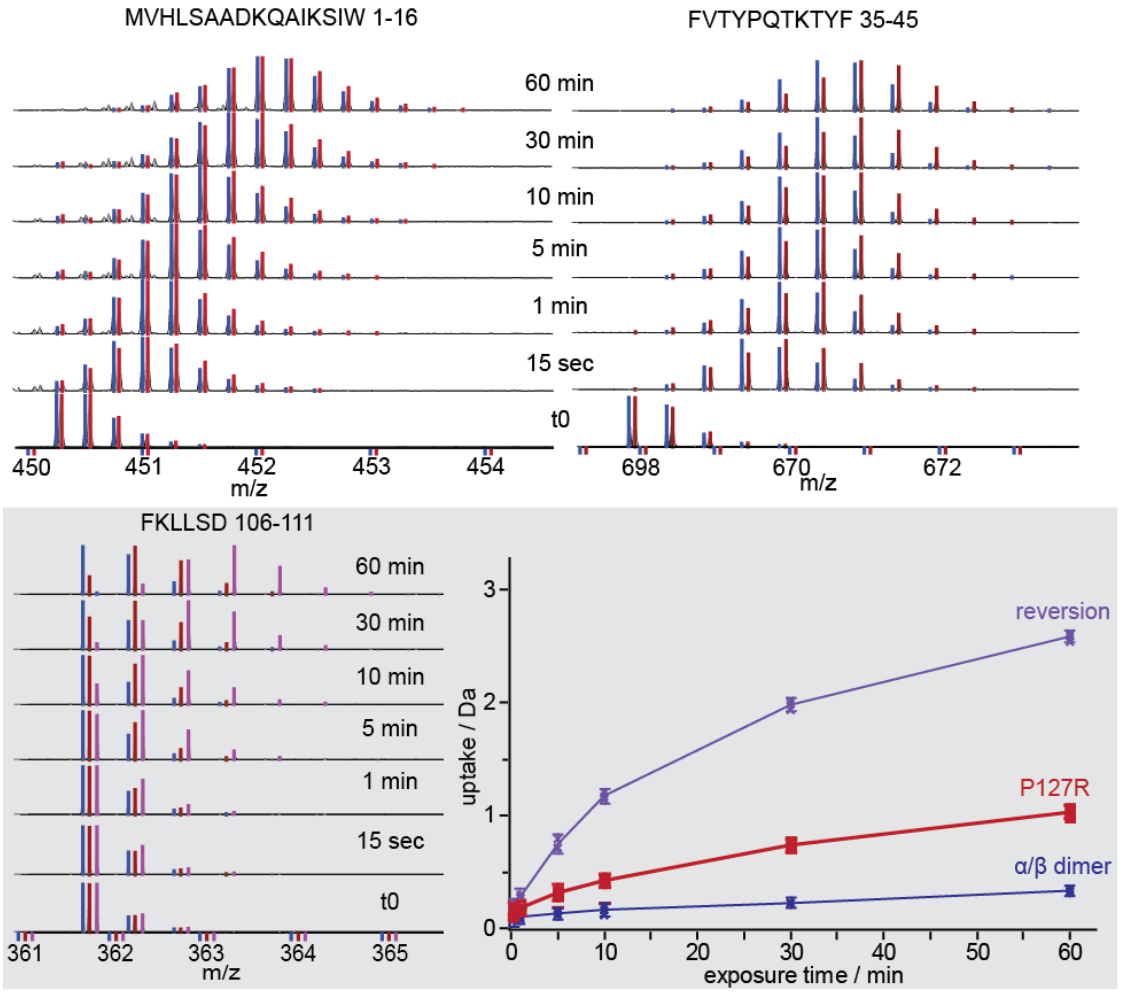


Figure 3.10. Uptake spectra for dimeric Anc haemoglobin and monomeric mutants. Mass spectra of Anc haemoglobin (blue) and mutants P127R (red) and Rev (purple) showing the spectral increase due to deuterium incorporation as a function of time. Highlighted peptides are shown with no significant difference, 4+ 1-16, a small difference between Anc and P127R, 2+ 35-45, and a large difference between Anc haemoglobin and the P127R or Rev mutants, 2+ 106-111. A corresponding uptake plot, highlighted in grey, over time is shown for the peptide 106-111 for Anc, P127R and Rev haemoglobin. None of the mutations are associated with, or immediately flank, this peptide.

Other regions see further increases still, such as the highlighted 106 - 111 region. Particularly large differences were observed in uptake between Anc and the Rev mutant, with modest increases also seen for the P127R mutant. The extracted uptake centroid was plotted as a function of time for the three systems demonstrated the increased ability for the two mutants to exchange over all time points. After 1

hour the reversion mutant had increased uptake of 2.25 ± 0.07 Da corresponding to a 7.5 fold increase in uptake from the dimeric state. The P127R showed less of an increase at 3.3 fold but this still represents a significant amount of deprotection between the dimeric and monomeric forms. Given the short peptide length of 5 residues, the N-terminal residue does not contribute to uptake, this is a highly localised change.

To map these changes across all peptides an external script was written to allow the comparison of the Anc dimer with the mutants. Existing proprietary software is not able to automatically compute the differences between peptides in mutants regions and as such a Python code was developed to use the output from DynamX software and automatically produce uptake plots representing each peptide, known since in other work as Woods plots.[48] The difference in uptake per exchangeable residue is a crude measure of determining how important an uptake change is as it provides a length independent measure and can crucially help account for mutations which introduce or lose an unexchangeable proline residue, such as P127R. Figure 3.11 is a plot of the relative uptake change for each peptide (grey bar) having been normalised to the maximum relative uptake observed for each peptide at that time point. This acts to emphasise those regions which have a large uptake change as a proportion of the observed exchange. The normalised uptake change is plotted for each peptide shown as a bar spanning the residues the peptide covers. The plots show the change in uptake of the monomer species relative to the Anc dimer, as such a positive uptake difference corresponds to greater uptake for the monomer over the dimer.

The Woods plots were produced for P127R (top) and Rev (lower), relative to Anc haemoglobin, across all time points. Neither mutant observes a significant number of peptides where a decrease in uptake is observed which is consistent with deprotection of an interface on forming free monomer. A number of regions display large increases in uptake most notably those around the 106-111 peptide (highlighted in blue) discussed previously in Figure 3.10. The normalised difference is seen to increase for both mutants over time starting initially at a value of 0.48

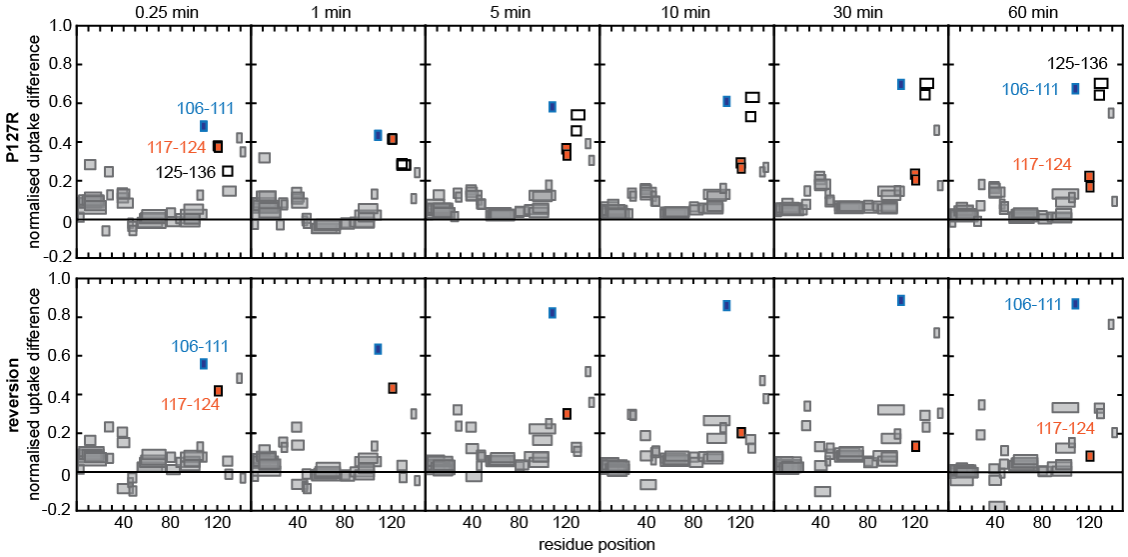


Figure 3.11. Differential HDX-MS reveals interfacial residues. The difference in uptake is plotted for each peptide for monomeric mutants relative to dimeric Anc haemoglobin. Uptake is displayed as a grey bar spanning the residues in corresponding peptide. The relative uptake change is first calculated and scaled by the number of residues. To highlight significant differences, the relative uptake change is then plotted having been normalised to the maximum relative uptake for the given peptide at that exposure time. P127R (**top**) and reversion (**lower**) are plotted for each time point, 0.25 - 60 minutes exposure. Residues 106-111 (blue) and 117-124 (orange) are highlighted for both systems. Peptides including the point mutation in P127R are shown as a white boxes.

and 0.56 for P127R and reversion respectively, this increased after an hour to 0.68 and 0.87. The P127R mutant sees a cluster of peptides from residue 106 onwards displaying larger increases. At short time points, 0.25 - 10 mins, the peptides 117-124 and 118-124 (orange bar) show a modest increase, peaking at 0.42. The reversion mutant also displays the same increase though the increased difference is only observed for exposure times less than 10 minutes. For both mutants there are a cluster of peptides around residue 40 with a small increase, 0.15 - 0.25, and have greater uptake over those immediately flanking. This is particularly evident for the P127R mutant after 30 - 60 mins.

For P127R, two peptides, 125-133 and 125-136, (white bars) also display large increases comparable to those seen for the 106-111 peptide particularly after 5 minutes. Differences for these peptides should be treated with particular caution as they incorporate the single point mutation seen in P127R. Whilst calculating the relative uptake per residue helped compensate for the impact of this mutation

it is difficult to determine the absolute contribution to uptake of each residue and so scaling in this manner may distort the significance of the peptide's differences. However, the absolute uptake differences for the 125-133 and 125-136 peptides are 1.8 ± 0.2 Da and 1.4 ± 0.1 Da, respectively. The maximum possible increase brought about solely by the mutation is 1 Da and so it certainly appears that at least some of increase is due to the lack of an interface, though it is hard to decouple the exact contributions.

To test the observed uptake behaviour of the monomeric mutants a series of differential HDX experiments were performed on a concentration series of Anc haemoglobin. In diluting Anc down to concentrations of 15, 2, and 0.67 μM a sliding scale of monomeric abundance can be obtained, at ∼55 %, ∼85 % and ∼94 %, respectively, using Eq. 3.1 in the Methods section. An uptake comparison at low concentration against the high concentration provides a means of obtaining an interfacial uptake difference to compare with the mutant data.

A Woods plot of the uptake difference for the lowest concentration, 0.67 μM, relative to the high concentration "dimer" is shown in Figure 3.12A. The majority of peptides display little difference, <15%, with most peptides resembling the mutant data across all time points. Regions 106-111, 117-124 and 125-136 are highlighted as previously. These three regions consistently account for the majority of elevated uptake differences seen over the exposure range. That these regions display similar increased levels of deprotection across all monomeric systems supports the idea that these are involved at an interface. The 106-111 peptide in particular has consistently displayed a large difference. Indeed, peptides spanning residues 85-111 which overlap with the small but significant peptide see a small increase also. The increase in uptake corresponding to these peptides however is almost entirely localised at the 106-111 residues, for example, a comparison of 0.67 μM against 75 μM after 1 hour shows an uptake difference of 0.5 ± 0.1 Da for 106-111 whereas 0.9 ± 0.2 Da for 89-111. In this example, neglecting back-exchange contributions, ∼55 % of exchange occurs with 25 % of exchange competent residues, demonstrating just how localised the uptake is to this region.

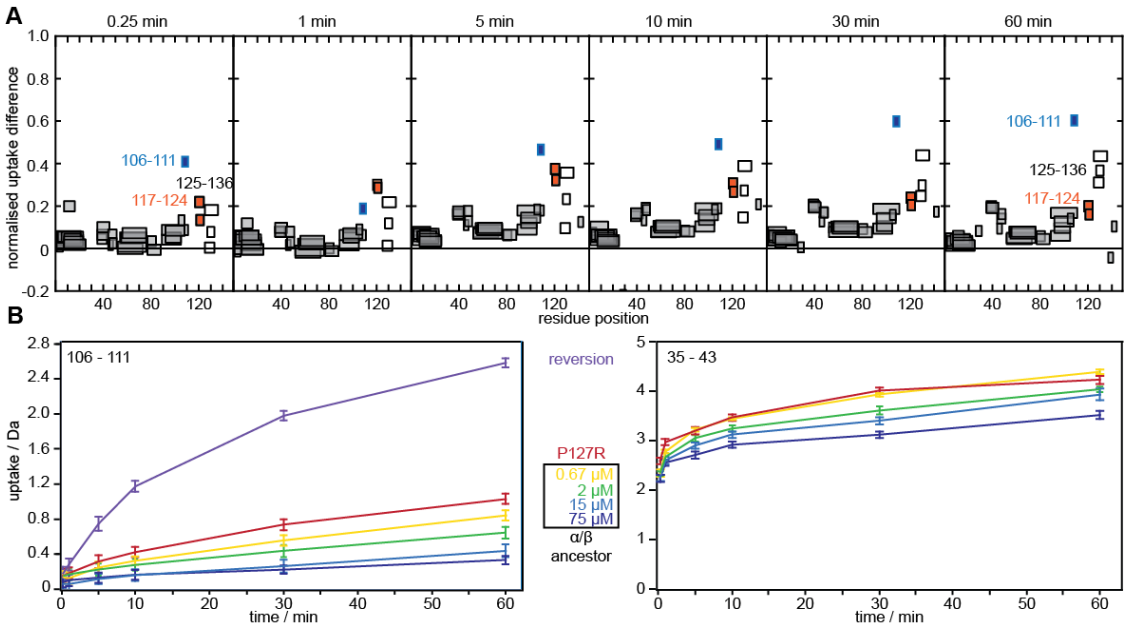


Figure 3.12. Titration of ancestral α/β haemoglobin supports interfacial regions determined by HDX-MS. Ancestral α/β haemoglobin was diluted from 75 μM to 0.67 μM to shift the equilibrium in favour of the monomer. A differential HDX-MS experiment was performed at each concentration relative to the dimeric 75 μM. **A** The resulting increase in peptide uptake, relative to the highest uptake observed at the given exposure time and peptide, was plotted for each time point, 0.25 - 60 mins. An increase corresponds to less protection for the monomeric state. Each peptide is shown as a grey bar spanning the residues in the peptide. Peptide 106-111 has a particularly high uptake difference and is highlighted in blue. Peptides from other notable regions, 117-124 and 125-136, are coloured orange and white respectively. **B** Plots of absolute uptake against exposure time are shown for peptides 106-111 and 35-43, over a concentration series. These peptides show a clear concentration dependence on uptake supporting the involvement in a dimeric interface. Decreasing the concentration, more monomeric, during labelling results in greater uptake. The P127R mutant uptake is shown for reference and the reversion data is shown for 106-111, not for peptide 35-43 owing to mutations in this region.

To further confirm these trends in uptake with an increased monomer population, plots of absolute uptake as a function of time and Anc concentration are shown in Figure 3.12B. The uptake difference between low and high concentration Anc increases over time. Two peptides are shown where the concentration dependence is particularly clear in the absolute uptake data. The first being the 106-111 peptide where increasing the about of free monomer results in a steadily increasing uptake. The uptake levels remain below that of the P127R mutant and significantly below the reversion system. The lowest concentration begins to approach the uptake levels of the P127R mutant though it is not surprising that even the lowest concentration Anc is lower in uptake than the mutants, as a proportion of the Anc exists in exchange

with a dimer. Peptide 35-43 displays similar difference behaviour; the reversion is not shown owing to mutations in this region. Here the lowest concentration of Anc has a comparable level of uptake to the monomeric P127R mutant throughout. Peptides around this region do not display uptake differences as large as for 106-111 but are consistently elevated relative to neighbouring peptides in all monomer systems. That a concentration dependence exists supports the mutant data as the WT equilibrium is shifted towards the monomer the uptake behaviour more closely mirrors the mutants.

Uptake differences indicate the locality of the dimeric interface

The motivation behind these studies was to determine how the interfaces involved in modern day haemoglobin tetramers came to be. By studying the dimeric interface of the last common ancestor of the α and β chains of tetrameric haemoglobin we can begin to understand the evolution from monomeric globins to modern day haemoglobins and the order in which interfaces were formed. Two candidate interfaces for Anc haemoglobin exist based on precursors to the two modern-day interfaces. These interfaces were determined by members of the Thornton group, University of Chicago, through phylogenetic analysis of residue substitutions and deletions, in addition to homology model interface docking.

The proposed interfaces, termed 1 or 2, are mapped onto the uptake difference plots for each of the monomeric systems, Fig. 3.13A. Only peptides above an arbitrary threshold of 15 % are shown for the exposure times, 1, 10 and 60 minutes. Residues from interface 1 are coloured blue and interface 2 in red. It becomes clear that the vast majority of peptides that display elevated increases in uptake exist at either interface 1 or 2. There is only one peptide in Anc that does not incorporate any residues from interface 1 or 2, peptide 140-144, though it remains close in both sequence and space. Both mutants have additional peptides, though number less than 5, which do not incorporate any of the proposed interface residues. However, all reside close to proposed interfaces with the exception of peptide 7-16, though

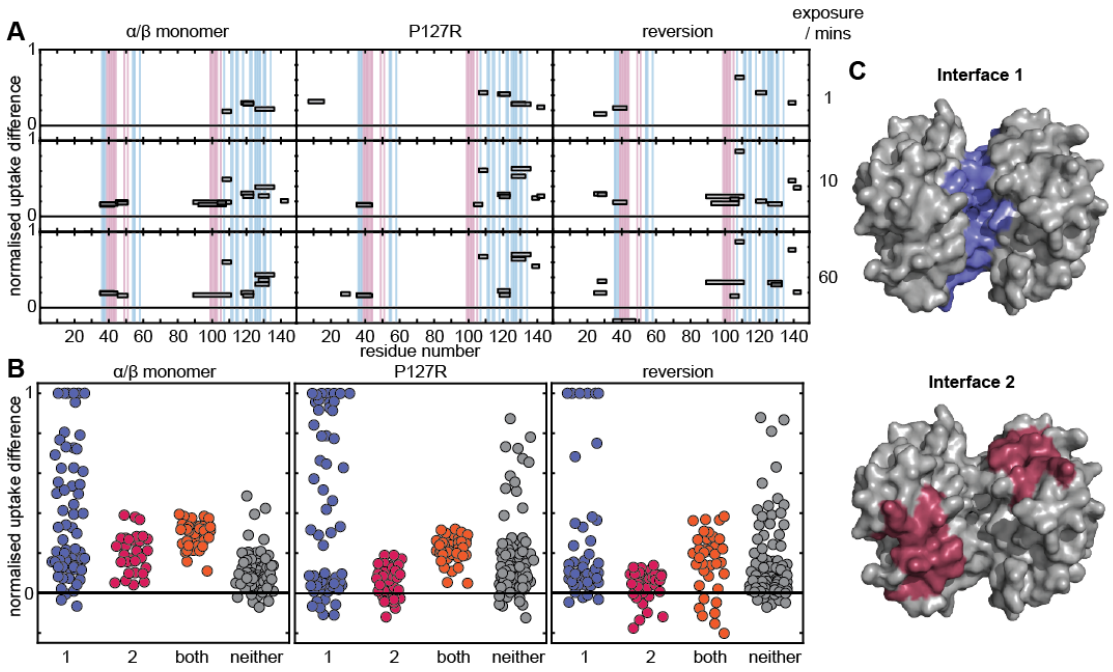


Figure 3.13. HDX-MS data suggest residues buried at the ancestral haemoglobin dimer interface. A Uptake difference for the monomer systems relative to Anc dimer plotted for each peptide. Data is shown for all peptides displaying a relative change of >15 %. Interfacial residues proposed by phylogenetic analysis are indicated in blue, interface 1, and red, interface 2. **B** The normalised uptake is plotted for each peptide, across all time points, having been grouped by whether the peptide incorporates proposed interface 1 (blue), 2 (red), both (orange) regions or neither region (grey). Uptake differences, those seen in Figures 3.11 & 3.12, are normalised to the maximum across all time points for each system. **C** Homology models for dimeric Anc based on modern haemoglobin are shown. The model interface corresponds to that of interface 1, coloured blue. Proposed interface 2 is highlighted in maroon relative to interface 1.

this is only observed for the 1 min time point and has just a 0.3 ± 0.1 Da absolute difference.

What is also important to note is that all peptides that do overlap with proposed interfaces always contain at least one residue from interface 1, with the vast majority only containing interface 1. This suggests that of the two proposed regions interface 1 is more likely to be the dimeric interface owing to a greater number of peptides displaying increased exposure. This is made clearer if the difference in uptake for all peptides are plotted and grouped according to which proposed interface they belong, 1, 2 or both, or do not belong at all. Such a plot is shown in Figure 3.13B for all peptides across all time points for each monomeric comparison. Uptake differences are normalised to the largest difference observed across all peptides for each time point, in most cases this means normalising to the

106-111 peptide. If we first consider the Anc system in which no mutants exist, we see that the largest uptake differences belong to interface 1 (blue). Peptides containing solely interface 2 (red) are slightly elevated with respect to those at neither peptides at neither proposed interface (grey). Additionally, those peptides which contain both interface 1 and 2 residues (orange) has a difference distribution resembling both. Mutant data also supports these observations with the same trend between interface 1, 2 and both. As the number of mutations increases, peptides that aren't predicted to contain interfacial residues (grey) are also seen increase though not quite to the levels of interface 1. These differences are likely to be as a result of minor structural differences in the mutant systems that aren't observed Anc. Allosteric structural changes could also cause such increases away from the interface, though Anc doesn't show appreciable allostery in O_2 binding assays, data not shown. These data presented suggest the proposed interface 1 is the more likely candidate for the dimeric interface in Anc haemoglobin. This interface is shown mapped on to the minimised homology model structure with interface residues coloured blue, figure 3.13C. Whilst this data doesn't support interface 2 these residues are shown on the same structure in red relative to the docked interface 1.

Proline-127 features at interface 1 and the reversion mutants exist primarily at interface 1 with some also featuring at interface 2. Peptides spanning residues 106-111, 117-124 and 125-136, which have been previously highlighted, all feature exclusively in proposed interface 1.

Whilst these data cannot definitively eliminate the possibility that some of the residues from interface 2 are involved in the dimeric interaction, our data strongly suggest the dominant role of interface 1 residues. Increases in uptake elsewhere may not simply be due to deprotection at the interface but could be a coupled, local allosteric, change in structure upon binding. The consistency across all three monomer systems, P127R, reversion and low concentration Anc, demonstrate the importance of the common peptides which almost all correspond to the hypothesised interface 1.

3.4.3 Conclusion

The comparative HDX-MS experiments presented here, supported by native mass spectrometry, provide compelling evidence for the residues involved in the interface of the ancestral dimer of α/β haemoglobin. Our data supports the phylogenetic hypothesis that the interface of the last common ancestor evolved to be interface 1 in extant haemoglobin. A combination of mutant systems, which disrupted the dimeric interface, and a serial dilution of the WT dimer provided means of probing the monomer surface relative to the dimer. We localise a region of high difference between the monomeric and dimeric forms to the residues 106-111. These differences are evident in all the systems probed and a titration series provided further validation. Such differences cannot be accounted for by a loss of heme from the binding pocket as heme binding for the monomeric systems is comparable to the dimer. These data provide the first biophysical structural data on the ancestral form of the 450 million year old haemoglobin and aid our understanding of the evolution to the modern-day tetramer.

3.5 Concluding remarks

In this Chapter we have demonstrated the versatility of HDX-MS for probing the unknown structural states in two dynamic systems. Firstly, we presented conclusive evidence in support of a light induced conformational change at the CTT of *Dm*-Cry. We showed the quantitative nature of the technique through the extraction of EX1 kinetic behaviour from such regions. This provides a unique read-out of the relative populations of the dark and light states of the protein for each associated region. Much is still to be gained from HDX studies on cryptochrome systems and these methods provide the foundation for doing so. Secondly, we use HDX-MS in combination with native MS to map the interface of dimeric ancestral hameoglobin. Phylogenetic hypotheses in addition to selective mutations pointed towards a precursor to an interface seen in modern-day haemoglobin, though a second binding mode could not be simply eliminated. Our HDX-MS data provided

strong evidence for those residues involved at the interface from both a titration series and disruptive mutants. These two studies demonstrate the utility of HDX-MS as a biophysical tool to compliment high resolution techniques and modelling for quantifying conformational changes and determining binding interactions.

3.6 Methods

3.6.1 Materials

All chemicals and reagents were purchased from Sigma Aldrich (Gilligham, UK). Haemoglobin was expressed and purified from *E. coli* by Arvind Pillai and Georg Hochberg of the Thornton Laboratory (University of Chicago, USA). *Dm*-Cry was expressed and purified from *E. coli* by members of the Weber Laboratory (University of Freiburg). Native equilibration buffer: 5 mM potassium phosphate (H_2O), pH 7.4. Labelling buffer: 5 mM potassium phosphate (D_2O), pD 7.4. Quench buffer: 50 mM potassium phosphate (H_2O), pH 1.9.

3.6.2 HDX-MS

Labelling experiments are performed by diluting 5 µL of protein sample into 55 µL of a deuterated buffer of the same composition and corresponding pD. This results in a labelling solution ∼92 % D_2O. Samples were incubated between 15 s and 1 hour before quenching with an ice-cold H_2O buffer (pH 1.9) of equal volume. The quenched solution pH was around 2.5 at 0 °C. The haemoglobin samples required an additional 1 M guanadinium chloride to be added to aid denaturation before digestion. This was quickly injected into an on-line HDX manager (Waters, Milford, MA, USA) which is described in Chapter 2 on p. 52. Briefly the sample was injected on to a 50 µL sample loop at 0 °C before passing over an immobilised pepsin column, Enzymate Pepsin (Waters), at 20 °C using an isocratic H_2O (0.1 % v/v) formic acid solution (200 µL/min). Peptide products were collected on a trapping column, BEH C_{18} (Waters), held at 0 °C. After 2 minutes of collection, and de-salting, peptides were eluted from the trap column on to an analytical column for separation using a reverse-phase gradient with a flow rate of 40 µL/min. The elution profile

using a H_2O/MeCN (+0.1% formic acid v/v) gradient was as follows: 1-7 minutes 97 % water to 65 % water, 7-8 minutes 65 % water to 5 % water, 8-10 minutes hold at 5 % water. The analytical flow rate was 40 μL/min and was electrosprayed directly into a Synapt G2Si (Waters, Wilmslow, UK) Q-ToF instrument for mass analysis. Sample handling was semi-automated using a robotic liquid handling HDX system (LEAP technologies, Ringwood, Australia) to ensure reproducibility in timings. A blank and cleaning injection cycle was performed in between each labelling experiment.

Mass spectrometry conditions were optimised to avoid H/D exchange in the step-wave source region with water vapour leading to spurious EX1 kinetics.[49] Mass spectrometry conditions were as follows: capillary 2.8 kV, sample cone 30 V, source offset 30 V, trap activation 4 V, transfer activation 2 V. The source temperature was set to 80 °C and cone gas flow 80 L/hr, the desolvation temperature was 150 °C and the desolvation gas flow of 250 L/hr. LeuEnk was used as an internal calibrant and acquired every 30 s.

Dm-Cry

To monitor the conformational rate of opening, continuous HDX labelling was performed in the presence and absence of blue light in a temperature controlled environment. Experiments were performed on 15 μM protein samples. All experiments were performed in a dedicated dark-room to control light-intensity throughout. Blue light exposure was enabled by fixed wavelength LEDs (450± 10 nm) at a light intensity of ∼0.1mW/cm^2. This was increased to ∼100mW/cm^2 when monitoring the conformational dependency on blue light intensity. Red light LEDs 0.1-10mW/cm^2 were used to as a negative control to determine the specificity towards blue light. Fully deuterated controls were not possible due to sample instabilities after around 2 - 3 hours.

The closing rates were determined using pulsed HDX-MS. Samples were equilibrated in a H_2O based buffer in the presence of blue-light for 20 minutes before variable periods of complete darkness. After a time in darkness samples were

labelled using the same continuous labelling strategy described above for a fixed 15 s.

Ancestral haemoglobin

Sample concentration was varied to control the relative populations of monomeric and dimeric species of Anc. After dilution into the labelling buffer Anc haemoglobin concentrations were 0.67, 2, 15, and 75μM; to avoid significant sample over-loading of the column when using high concentrations of Anc, samples were diluted during quenching to give an injection quantity of ∼15 pmol. To ensure back-exchange occurred equally across all diluted samples the final ratio of H_2O:D_2O after quenching was kept constant at 54:46. This allows for all concentrations to be compared without correcting for back-exchange.

3.6.3 Back-exchange controls of the system

To verify our HDX-MS workflow was capable of maintaining high levels of deuterium incorporation the back-exchange was calculated for the peptide angiotensin II and peptides from cytochrome C. Lyophilised samples were dissolved in D_2O (pH 4.0) and left for 48 hours. A reduced pH deuterated solution is required for long exposure times due to the slow exchange of the C_2 hydrogen in histidine.[50,51] Manual liquid handling of angiotensin II was seen to have a back-exchange of 23 ±3 %. The robotic sample handling was more reproducible and displayed back exchange of 29.1 ±0.1 %, 31.8 ±0.2 %, and 34.0 ±0.1 % for 1, 2 and 3 minute peptide trapping times, respectively. Intact labelling of cytochrome C followed by a 3 minute digestion was 30 ±7 % averaged across all peptides.

3.6.4 Native MS

Native mass spectrometry of the haemoglobin variants was performed on a modified Q-ToF platform (Synapt G1 HDMS, Wilmslow, UK). Samples were buffer exchanged and diluted into 200 mM ammonium acetate at the concentration indicated and nano-electrosprayed in positive polarity using borosilicate capillaries coated in gold.[52]

Low-energy spectra were acquired in ToF mode using the following conditions: capillary 1.4 kV, sample cone 10 V, extraction cone 3 V, trap activation 10 V, transfer activation 8 V, backing pressure 3.2 mbar, trap Ar pressure 12μbar. When checking for gas-phase ligand loss a fixed quadrupole was applied at 2000 *m/z* to allow transmission greater than this range.

3.6.5 Data analysis

Peptides were identified, in the absence of labelling, by data-independent MS/MS analysis (MS^E) of the eluted peptides and subsequent database searching in the Protein Lynx Global server 3.0 software (Waters). Peptide identifications were filtered according to fragmentation quality (minimum fragmentation products per amino acid: 0.2), mass accuracy (maximum $[MH]^+$ error: 5 ppm), and reproducibility (peptides identified in all MS^E repeats) before their integration into HDX analysis. HDX-MS data were processed in DynamX 3.0 software (Waters), and all automated peptide assignments were manually verified, with noisy and overlapping spectra discarded. Spectra displaying bimodal distributions according to EX1 type kinetics were quantified using HX-Express 2.0 as described in the given references.[29,30] Dark state data were fit to a single binomial distribution (2 fitting parameters) and light state data to a bimodal binomial distribution (4 fitting parameters). A reduced chi-squared per degree of freedom (χ^2_ν) was calculated from the difference between the experimental intensity and that determined from the binomial, or sum of binomials, in the fitting procedure. The average χ^2_ν across all time points was equal to 3, with time points greater than 1 minute displaying a better fit, $\chi^2_\nu = 0.9$. Standard deviations used to calculate χ^2_ν were calculated from 3 experimental repeats. External python scripts were written to generate and analyse the Woods plots from data outputs of DynamX. Processing of mass spectra was carried out using MassLynx V4.1 software (Waters, Wilmslow, UK).

Secondary stucture assignments were performed using ESPript 3.0 (SBG grid consortium). PyMOL software (Schrödinger, NY, USA) was used for 3-D protein structure visualisation. Sequence alignment of the haemoglobin mutants was

performed using MUSCLE software (drive5). In a monomer-dimer equilibrium the monomer concentration, $[M]$, may be calculated for a given dissociation constant, K_d, and known sample concentration, $[M]_{Tot}$, using Eq. 3.1.

$$[M] = \frac{\sqrt{{K_d}^2 + 8K_d[M]_{Tot}} - K_d}{4} \tag{3.1}$$

Where, $[M]_{Tot}$ is the total sample concentration of protein with respect to monomer, $[M]$ is the monomeric concentration, $[D]$ is the dimeric concentration and $[M]_{Tot} = [M] + 2[D]$. The ratio of monomer to dimer may therefore be calculated for a given K_d. The difference between the experimentally observed monomer-dimer ratio and the equation derived value was calculated. The least squares error was then minimised for all data sets using the SOLVER function in Excel software (Microsoft, Redmond, USA). The monomer and dimer intensities were determined from native MS spectra using peak heights from all charge states acquired from MassLynx V4.1 software (Waters, Wilmslow, UK).

3.7 Acknowledgments

First and foremost I would like to thank our collaborators who approached us with such interesting biological questions and made both pieces of work possible. Specifically to Katie Buckton, Stuart Mackenzie, Peter Hore and Chris Timmel for the work on exploring *Dm*-Cry. I thank Katie for introducing and advising me on the cryptochrome systems and with whom it was a pleasure to work. Whilst all experimental work presented here were my own efforts towards this project, much of our recent efforts are credited to Katie. I would also like to thank Joe Thornton, Arvind Pillai and Georg Hochberg for presenting us with such a fascinating system to study, in the form of ancestral haemoglobins. Aside from collaborators, a number of people advised me in establishing and improving our HDX-MS protocols, not least, Shahid Mehmood and Jon Fox. I thank Reid Alderson and Peter Hore for their critical review of this chapter. Finally, to Justin Benesch who allowed me to assume group responsibility for the instrumentation and provided critical advice and discussion throughout.

References

[1] Mandell, J. G.; Falick, A. M.; Komives, E. A. Identification of protein-protein interfaces by decreased amide proton solvent accessibility. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 14705–14710.

[2] Liu, Y. H.; Belcheva, A.; Konermann, L.; Golemi-Kotra, D. Phosphorylation-Induced Activation of the Response Regulator VraR from Staphylococcus aureus: Insights from Hydrogen Exchange Mass Spectrometry. *Journal of Molecular Biology* **2009**, *391*, 149–163.

[3] Ehring, H. Hydrogen exchange/electrospray ionization mass spectrometry studies of structural features of proteins and protein/protein interactions. *Analytical Biochemistry* **1999**, *267*, 252–259.

[4] Engen, J. R. Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Analytical Chemistry* **2009**, *81*, 7870–7875.

[5] Chaves, I.; Pokorny, R.; Byrdin, M.; Hoang, N.; Ritz, T.; Brettel, K.; Essen, L.-O.; van der Horst, G. T. J.; Batschauer, A.; Ahmad, M. The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. *Annual Review of Plant Biology* **2011**, *62*, 335–364.

[6] Ahmad, M. Photocycle and signaling mechanisms of plant cryptochromes. *Current Opinion in Plant Biology* **2016**, *33*, 108–115.

[7] Hore, P. J.; Mouritsen, H. The Radical-Pair Mechanism of Magnetoreception. *Annual Review of Biophysics* **2016**, *45*, 299–344.

[8] Ritz, T.; Adem, S.; Schulten, K. A model for photoreceptor-based magnetoreception in birds. *Biophysical Journal* **2000**, *78*, 707–718.

[9] Ceriani, M. F.; Darlington, T. K.; Staknis, D.; Mas, P.; Petti, A.; Weitz, C. J.; Kay, S. A. Light-Dependent Sequestration of TIMELESS by CRYPTOCHROME. *Science* **1999**, *285*, 553–556.

[10] Busza, A.; Emery-le, M.; Rosbash, M.; Emery, P. Roles of the Two Drosophila CRYPTOCHROME Structural Domains in Circadian Photoreception. *Science* **2004**, *304*, 1503–1506.

[11] Peschel, N.; Chen, K. F.; Szabo, G.; Stanewsky, R. Light-Dependent Interactions between the Drosophila Circadian Clock Factors Cryptochrome, Jetlag, and Timeless. *Current Biology* **2009**, *19*, 241–247.

[12] Levy, C.; Zoltowski, B. D.; Jones, A. R.; Vaidya, A. T.; Top, D.; Widom, J.; Young, M. W.; Scrutton, N. S.; Crane, B. R.; Leys, D. Updated structure of Drosophila cryptochrome. *Nature* **2013**, *495*, E3–E4.

[13] Berndt, A.; Kottke, T.; Breitkreuz, H.; Dvorsky, R.; Hennig, S.; Alexander, M.; Wolf, E. A Novel Photoreaction Mechanism for the Circadian Blue Light Photoreceptor Drosophila Cryptochrome. *The Journal of biological chemistry* **2007**, *282*, 13011–13021.

[14] Schulten, K.; Swenberg, C. E.; Weller, A. A Biomagnetic Sensory Mechanism Based on Magnetic Field Modulated Coherent Electron Spin Motion. *Zeitschrift für Physikalische Chemie* **1978**, *111*, 1–5.

[15] Liu, Z.; Tan, C.; Guo, X.; Li, J.; Wang, L.; Sancar, A.; Zhong, D. Determining complete electron flow in the cofactor photoreduction of oxidized photolyase. *Proceedings of the National Academy of Sciences* **2013**, *110*, 12966–12971.

[16] Lin, C.; Top, D.; Manahan, C. C.; Young, M. W.; Crane, B. R. Circadian clock activity of cryptochrome relies on tryptophan-mediated photoreduction. *Proceedings of the National Academy of Sciences* **2018**, 201719376.

[17] Nohr, D.; Franz, S.; Rodriguez, R.; Paulus, B.; Essen, L. O.; Weber, S.; Schleicher, E. Extended Electron-Transfer in Animal Cryptochromes Mediated by a Tetrad of Aromatic Amino Acids. *Biophysical Journal* **2016**, *111*, 301–311.

[18] Cailliez, F.; Müller, P.; Firmino, T.; Pernot, P.; De La Lande, A. Energetics of Photoinduced Charge Migration within the Tryptophan Tetrad of an Animal (6-4) Photolyase. *Journal of the American Chemical Society* **2016**, *138*, 1904–1915.

[19] Hemsley, M. J.; Mazzotta, G. M.; Mason, M.; Dissel, S.; Toppo, S.; Pagano, M. A.; Sandrelli, F.; Meggio, F.; Rosato, E.; Costa, R.; Tosatto, S. C. E. Linear motifs in the C-terminus of D. melanogaster cryptochrome. *Biochemical and Biophysical Research Communications* **2007**, *355*, 531–537.

[20] Dissel, S.; Codd, V.; Fedic, R.; Garner, K. J.; Costa, R.; Kyriacou, C. P.; Rosato, E. A constitutively active cryptochrome in Drosophila melanogaster. *Nature Neuroscience* **2004**, *7*, 834–840.

[21] Conrad, K. S.; Manahan, C. C.; Crane, B. R. Photochemistry of flavoprotein light sensors. *Nature chemical biology* **2014**, *10*, 801–9.

[22] Vaidya, A. T.; Top, D.; Manahan, C. C.; Tokuda, J. M.; Zhang, S.; Pollack, L.; Young, M. W.; Crane, B. R. Flavin reduction activates Drosophila cryptochrome. *Proceedings of the National Academy of Sciences* **2013**, *110*, 20455–20460.

[23] Ozturk, N.; Selby, C. P.; Annayev, Y.; Zhong, D.; Sancar, A. Reaction mechanism of Drosophila cryptochrome. *Proceedings of the National Academy of Sciences* **2011**, *108*, 516–521.

[24] Partch, C. L.; Clarkson, M. W.; Özgür, S.; Lee, A. L.; Sancar, A. Role of structural plasticity in signal transduction by the cryptochrome blue-light photoreceptor. *Biochemistry* **2005**, *44*, 3795–3805.

[25] Lindner, R.; Heintz, U.; Winkler, A. Applications of hydrogen deuterium exchange (HDX) for the characterization of conformational dynamics in light-activated photoreceptors. *Frontiers in molecular biosciences* **2015**, *2*, 33.

[26] von Horsten, S.; Straß, S.; Gruth, V.; Hellwig, N.; Klasen, R.; Essen, L. O. Mapping light-driven conformational changes within the photosensory module of plant phytochrome B. *Nature Communications* **2016**, 1–8.

[27] Kaltashov, I. A.; Eyles, S. J. *Mass spectrometry in structural biology and biophysics: Architecture, dynamics, and interaction of biomolecules*, 2nd ed.; Wiley, 2012.

[28] Keil, B. *Specificity of Proteolysis*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1992; pp 14–15.

[29] Weis, D. D.; Engen, J. R.; Kass, I. J. Semi-Automated Data Processing of Hydrogen Exchange Mass Spectra Using HX-Express. *Journal of the American Society for Mass Spectrometry* **2006**, *17*, 1700–1703.

[30] Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *Journal of the American Society for Mass Spectrometry* **2013**, *24*, 1906–12.

[31] Kutta, R. J.; Archipowa, N.; Scrutton, N. S. The sacrificial inactivation of the blue-light photosensor cryptochrome from Drosophila melanogaster. *Physical Chemistry Chemical Physics* **2018**, *20*, 28767–28776.

[32] Ganguly, A.; Manahan, C. C.; Top, D.; Yee, E. F.; Lin, C.; Young, M. W.; Thiel, W.; Crane, B. R. Changes in active site histidine hydrogen bonding trigger cryptochrome activation. *Proceedings of the National Academy of Sciences* **2016**, *113*, 10073–10078.

[33] Kattnig, D. R.; Nielsen, C.; Solov'yov, I. A. Molecular dynamics simulations disclose early stages of the photo-activation of cryptochrome 4. *New Journal of Physics* **2018**, *20*, 083018.

[34] Adamson, J. W.; Finch, C. A. Hemoglobin Function, Oxygen Affinity, and Erythropoietin. *Annual Review of Physiology* **1975**, *37*, 351–369.

[35] Schechter, A. N. Hemoglobin research and the origins of molecular medicine. *Blood* **2008**, *112*, 3927–3938.

[36] Yonetani, T.; Laberge, M. Protein dynamics explain the allosteric behaviors of hemoglobin. *Biochimica et Biophysica Acta - Proteins and Proteomics* **2008**, *1784*, 1146–1158.

[37] Eaton, W. A.; Henry, E. R.; Hofrichter, J.; Mozzarelli, A. Is cooperative oxygen binding by hemoglobin really understood? *Rendiconti Lincei* **2006**, *17*, 147–162.

[38] Bellelli, A.; Brunori, M. Hemoglobin allostery: Variations on the theme. *Biochimica et Biophysica Acta - Bioenergetics* **2011**, *1807*, 1262–1272.

[39] Hardison, R. C. Evolution of hemoglobin and its genes. *Cold Spring Harbor Perspectives in Medicine* **2012**, *2*.

[40] Storz, J. F.; Opazo, J. C.; Hoffmann, F. G. Gene duplication, genome duplication, and the functional diversification of vertebrate globins. *Molecular Phylogenetics and Evolution* **2013**, *66*, 469–478.

[41] Pauling, L.; Zuckerkandl, E.; Henriksen, T.; Lövstad, R. Chemical Paleogenetics. Molecular "Restoration Studies" of Extinct Forms of Life. *Acta Chemica Scandinavica* **1963**, *17 supl.*, 9–16.

[42] Thornton, J. W. Resurrecting ancient genes: Experimental analysis of extinct molecules. *Nature Reviews Genetics* **2004**, *5*, 366–375.

[43] Harms, M. J.; Thornton, J. W. Evolutionary biochemistry: Revealing the historical and physical causes of protein properties. *Nature Reviews Genetics* **2013**, *14*, 559–571.

[44] Wheeler, L. C.; Lim, S. A.; Marqusee, S.; Harms, M. J. The thermostability and specificity of ancient proteins. *Current Opinion in Structural Biology* **2016**, *38*, 37–43.

[45] Sarma, P. R.; Wheby, M. S. Red Cell Indices. *Archives of Internal Medicine* **1980**, *140*, 720–723.

[46] Stengel, F.; Baldwin, A. J.; Bush, M. F.; Hilton, G. R.; Lioe, H.; Basha, E.; Jaya, N.; Vierling, E.; Benesch, J. L. P. Dissecting Heterogeneous Molecular Chaperone Complexes Using a Mass Spectrum Deconvolution Approach. *Chemistry & Biology* **2012**, *19*, 599–607.

[47] Sowole, M. A.; Konermann, L. Comparative analysis of oxy-hemoglobin and aquomet-hemoglobin by hydrogen/deuterium exchange mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2013**, *24*, 997–1005.

[48] Lau, A. M. C.; Martens, C.; Ahdash, Z.; Politis, A. Deuteros: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry. **2019**,

[49] Guttman, M.; Wales, T. E.; Whittington, D.; Engen, J. R.; Brown, J. M.; Lee, K. K. Tuning a High Transmission Ion Guide to Prevent Gas-Phase Proton Exchange during H/D Exchange MS Analysis. *Journal of the American Society for Mass Spectrometry* **2016**, *27*, 662–668.

[50] Miyagi, M.; Nakazawa, T. Determination of pKa values of individual histidine residues in proteins using mass spectrometry. *Analytical Chemistry* **2008**, *80*, 6481–6487.

[51] Tran, D. T.; Banerjee, S.; Alayash, A. I.; Crumbliss, A. L.; Fitzgerald, M. C. Slow histidine H/D exchange protocol for thermodynamic analysis of protein folding and stability using mass spectrometry. *Analytical Chemistry* **2012**, *84*, 1653–1660.

[52] Hernández, H.; Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols* **2007**, *2*, 715–26.

4

Implementation and application of gas-phase HDX for probing protein surfaces

4.1 Abstract

Mass spectrometry (MS) based approaches are revolutionising the structural characterisation of proteins and the assemblies they form. These strategies either utilise a solution probe followed by accurate mass analysis, such as hydrogen/deuterium exchange MS (HDX-MS), detailed in Chapter 3, or interrogate the intact structure directly in vacuum through native mass spectrometry. The latter strategy is attractive primarily because it allows exploitation of mass-selective investigation.

Coupling native mass spectrometry to gas-phase labelling would allow for a highly selective probe of each constituent from an ensemble of states, discriminated by size and/or mass. Gas-phase HDX (gpHDX) involves introducing deuterated ND_3 vapour into the mass spectrometer to exchange the isotopic label with labile side-chain hydrogens on the accessible surface of the protein ion. gpHDX is, in principle, sensitive to both the chemistry and structure of the surface akin to solution labelling, however, mechanistic details of the reaction make predicting exchange behaviour difficult.

In this chapter, I describe the implementation, characterisation, and use of gas-phase HDX to interrogate native proteins and biomolecules in the mass spectrometer. Using a model system of green fluorescent proteins (GFPs) with varying surface electrostatics, we uncover sensitivity to surface-chemistry and how the degree of uptake can inform on the mechanistic behaviour.

4.2 Introduction

Native mass spectrometry (MS) provides a powerful means of probing many facets of a protein's tertiary and quaternary structure, Chapter 1.1. Native mass analysis of a protein complex inside the vacuum of the spectrometer lends itself to the accurate characterisation of subunits and their connectivity through tandem MS approaches. Whilst ion mobility (IM) MS has furthered the technique through its ability to understand gas-phase and/or solution-like structure at a course level, little can be gleaned about the protein surface or the interfaces involved in complex formation. To understand these aspects of a protein's structure one does not typically probe the species directly in the gas-phase but utilises the introduction of an label in solution first.

In chapter 3, I demonstrated how solution labelling through isotopic exchange can provide significant insight into a protein's conformational dynamics and interfacial assembly. Whilst a hydrogen/deuterium exchange (HDX) MS experiment has the ability to report on a single state, through data deconvolution, it will ultimately always be a measure of the ensemble of states present. Where, the reported incorporation value will predominantly reflect the most abundant species or that whose reactivity towards the label dominates. Whilst following the dynamics can often, usefully, provide information on all of these states it does limit the experiment in its general ability to selectively report a given state's structural dynamics. If we consider Figure 3.8, p. 98, previously described in chapter 3, ancestral haemoglobin exists in an equilibrium between its monomeric and dimeric forms. An additional level of complexity exists as each subunit also has an *apo* and *holo* form with regards to heme binding. Chapter 3 demonstrated solution HDX to be powerful in providing structural information but it does demonstrate that even for a relatively simple monomer-dimer system the complexity can drastically increase. This would be particularly true if the *holo* state were less abundant and the protein complex was heterogeneous or polydisperse in its make-up. Such complexity makes it difficult to selectively probe species in solution.

If we were to imagine performing a similar exchange reaction inside the spectrometer then we would now be able to measure the uptake directly for each species in isolation based on the snapshot of states present in solution. The mass spectrometer has a unique ability to selectively isolate individual protein species with an incredibly high separation power, especially when compared to solution separative techniques.

To this end, this chapter describes my implementation of gas-phase hydrogen/ deuterium exchange (gpHDX) in an ion mobility (IM) mass spectrometer to selectively probe the surfaces of native protein assemblies. This work builds upon the fundamental gpHDX studies with peptides and small proteins detailed in the Introductory chapter by developing applied methods to inform on larger protein systems and complexes, and how they evolve in the gas phase. We explore the important parameters involved in obtaining high quality gpHDX data whilst maintaining the integrity of the core native MS methodology. After which, we test the sensitivity of gas-phase HDX to differentiate between model surfaces of differing surface chemistry and uncover the behaviour of different exchangeable moieties.

4.3 Results and Discussion

4.3.1 Implementation and modifications

In order to gain the best understanding of ion structure our gpHDX modifications were carried out on a mass spectrometer enhanced for high-mass transmission that is capable of ion mobility separation. Our Waters Synapt G1 instrument was modified in a similar manner to that described by the Rand group[1,2] with additions to ensure finer pressure control and non-deuterated control experiments. This was achieved by leaking a low pressure of deuterated ammonia (ND_3) into the collision cells of the ion mobility spectrometer. This involved bypassing the inbuilt gas-flow controllers typically used for the argon collision partner in the trap and transfer regions and restricting the flow from the high pressure ND_3 source. We aimed to exploit the instrument's architecture by introducing the deuterated gas either before

or after the ion mobility cell allowing for a multitude of sequential experiments. The current, and following, chapters utilise this gpHDX set-up in both of these configurations. Switching valves were installed to provide control and experimental flexibility. Further practical implementation details are provided in the methods of this chapter, p. 143. In order to make gpHDX compatible with existing native MS methodologies, parameters including gas pressures and optics voltages were explored around typical operating values, as detailed in the following sections.

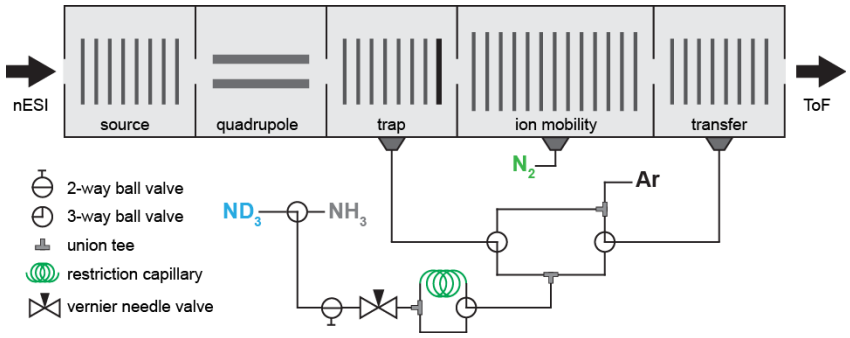


Figure 4.1. Gas-phase hydrogen/deuterium exchange modifications to an ion mobility mass spectrometer. Instrument schematic detailing the modifications around the gas inlets allowing for isotopic labelling in the trap or transfer regions. Figure adapted from Mistarz & Chandler *et al.*,[3] original designed by Ulrik Mistarz.

Protein ions have been shown to maintain aspects of native structure in the gas-phase on the order of milliseconds,[4] Chapter 1.1, and so the ability to label within this time-scale allows gpHDX to inform on gas-phase structure. The trap and transfer labelling cells utilise travelling waves (TW) to help push ions through the cell. Ions can be additionally accelerated into the cell using a direct current (DC) potential. A combination of the acceleration voltage and the travelling wave motion will determine the residence time of the ion in the labelling cell. If we first consider an upper bound by assuming ions are pushed through the cell at the same speed as the travelling waves then we obtain an upper residence limit. For much of these experiments we operated at a velocity around 200 ms^{-1} which corresponds to a transit time of ~0.5 ms. In reality this time could be longer if ions are not effectively trapped in the mobile potential wells, as in a mobility separation, but our labelling gas is only introduced at around 10 µbar which is an order of magnitude lower in pressure than an IM experiment. For the shortest residence time possible,

we could consider the time it would take for an ion to traverse the cell only after an initial injection voltage, impeded only by gas collisions and without any additional electric field. A model protein, GFP, was used throughout these studies and the time taken to pass through a 10 cm cell was estimated. For a 9+ ion, of diameter 55 Å and mass 28 kDa, under an argon pressure of 10 μbar at room temperature, an initial acceleration voltage of 8 V would result in a residence time of 0.2 - 0.3 ms[i]. This corresponds to nearly 1000 collisions and a mean free path of ∼0.1 mm. If the voltage is increased to 80 V the time reduces to 0.06 - 0.1 ms. Clearly these times are faster than the travelling waves and so the time taken will depend on how effectively the ions are trapped in the potential wells. If ions quickly become trapped after the initial acceleration into the ion guide they will exit the cell closer to the 0.5 ms limit, otherwise they will "skip" the wave and only be retarded by the collisions. Without simulations the exact residence time for the ions is hard to determine but it is certainly on the order of 0.1 - 0.5 ms.

4.3.2 Application to biomolecules

Upon exposure to ND_3 vapour in the trap ion-guide, labile hydrogens from a range of moieties can be seen to exchange in biomolecules, ranging from short peptides and sugars to small folded proteins. Mechanistic details are discussed in Section 2.4. Uptake is visible in small systems as a shift in the isotopic distribution towards higher *m/z*. Efficient labelling of two small peptides, 1+ LeuEnk and 2+ GluFib, consisting of 5 and 14 residues respectively, is demonstrated in Figure 4.2A & B. The theoretical maxima for the fast-exchanging labile hydrogens are shown as dashed grey lines. An example of a trisaccharide, melitriose, is shown in Figure 4.2C where limited exchange of 0.9 Da was observed, significantly below the maximum possible (11 Da). Complications in obtaining labelled sugars will be discussed further on page 125.

[i]This was calculated/simulated iteratively for each collision using Equation 2.8 on p. 45 using an Excel spreadsheet created by Justin Benesch.

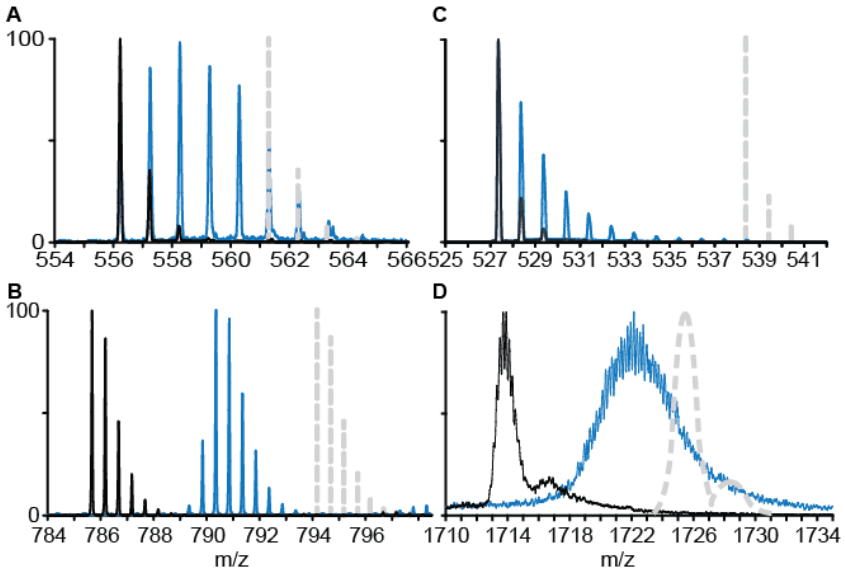


Figure 4.2. Gas-phase hydrogen/deuterium exchange of biomolecules. MS spectra of two peptides **A** LeuEnk 1+ and **B** GluFib 2+, a trisaccharide **C** melitriose(+ Na) 1+ and a small protein **D** ubiquitin 5+. Isotopic distributions obtained with Ar as a collision partner are shown in black. Gas-phase exchanged spectra, after infusion of ND_3 vapour, are shown in blue where a shift in m/z can be seen. Dashed grey lines represent the theoretical maximum isotopic envelope achievable should all fast exchanging sites react.

When labelling is performed on proteins, which have significantly greater secondary/tertiary structure, we observe equally efficient exchange as for small peptide systems. In these cases a quadrupole isolation was used to achieve exclusively $[M + H]^+$ ions and prevent transmission of other cationated ions. Salt adducts commonly found in native MS will complicate the resulting uptake spectra and so stringent quadrupole isolations are enforced throughout all gpHDX studies in this thesis. Native gpHDX of 5+ ubiquitin, ∼8.6 kDa, is shown in Figure 4.2D where the protein surface exchanged ∼43 D to give a broad Gaussian distribution. The broadening of the isotopic envelope is typical, and expected, as it represents the probabilities for exchange across all exchangeable sites. The isotopic pattern was still discernible for a protein of this molecular weight and charge state on this ToF instrument[ii] Any further increase in mass or charge resulted in an unresolved isotope distribution and as such, systems described later often appear as a single peak whose centroid was used as a measure of its uptake.

[ii]The instrument resolution required to isotopically resolve a protein is equivalent to the molecular weight of the ion, i.e. on a ToF with a resolution of ∼7000 a 7 kDa protein can be isotopically resolved at half height, see §2.1.4.

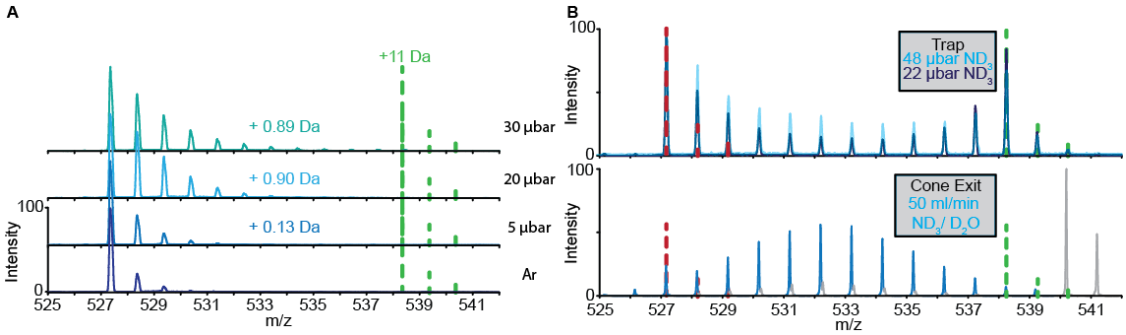


Figure 4.3. Gas-phase HDX of sugars. A MS spectra of sodiated melitriose 1+ ions with an increasing ND_3 pressure with uptake saturation evident at 20 µbar around 0.9 Da. Uptake maxima are shown as dashed green lines. **B** Representative spectra of unusual exchange behaviour resulting from electrospray variations. Melitriose labelled in the trap collision cell under 22 µbar and 48 µbar ND_3 results in an uptake range from no uptake,red, to complete uptake, green, and all values between (top). Broad isotope distribution spanning no uptake to complete uptake, obtained using an alternative labelling method using ND_3/D_2O vapour in the source region.[5] Grey lines are peaks not associated with $[\text{melitriose} + \text{Na}]^+$.

Unlike the peptide and protein systems studied, the uptake behaviour of sugars was highly dependent on the electrospray conditions. The melitriose system studied was found to have limited exchange capability with a centroid shift of less than 1 Da, over an ND_3 pressure range of 5 -30 µbar, Fig. 4.3A. Given no further exchange was observed between 20 and 30 µbar it appeared the system was fully saturated. Whilst the sugars systems studied often behaved in this manner, a number of spectra were obtained in which the uptake was less predictable or reproducible. Examples of this behaviour are shown in Figure 4.3B. A broad uptake distribution is seen for melitriose corresponding to no uptake (red), maximum uptake (green), and an uptake distribution in between. This is indicative of multiple conformations with different exchange competencies. This may reflect the solution conformations or, more likely, the distribution of conformers that are a product of the transferral to the gas-phase. Similar behaviour was also found when performing an alternative method of gpHDX,[5] where a D_2O/ND_3 vapour was introduced into the source region. A broad uptake distribution was found bounded by the theoretical minimum and maximum, Fig. 4.3B (lower). Examples in the literature suggest that efficient exchange can occur with sugars using ND_3 and the difference between our results may reflect the spray conditions or the absence of a bound sodium,[6] which may

interact with the hydroxyl groups. This reflects the highly sensitive nature of gpHDX to gas-phase structure and non-covalent interactions.

4.3.3 Effect of ion manipulation on uptake

Gas-phase labelling efficiency has been shown to be affected by ion residency time, acceleration potential and the number of collisions.[1,7–9] The travelling wave ion guides, described in Chapter 2, in our IM mass spectrometer are capable of such manipulations through control of the T-wave height or velocity. This has the effect of trapping ions in the potential well and pushing ions through the cell with differing speeds. At the pressures employed here, ∼10 µbar, few "roll-over" events, in which the ion is not effectively trapped, will occur but this will likely increase with decreased wave heights.[10] We set out to explore how controlling the wave parameters can influence the levels of uptake for peptides and proteins. Figure 4.4 first demonstrates the impact wave velocities have on the uptake of the peptide, GluFib. Travelling wave velocities were increased under two different wave heights, 6 V and 15 V, in both the trap and transfer ion guides in the presence of ND_3. The 2+ GluFib ion was labelled under an ND_3 pressure of 10 µbar with an initial uptake of 9.2 - 10.5 Da depending on the wave height and labelling region. Uptake levels decreased slightly as the wave velocities were increased to 200 $m\,s^{-1}$. After a wave velocity of ∼500 $m\,s^{-1}$ the uptake began to increase significantly, irrespective of the labelling region or wave height. This increase was more pronounced for higher wave heights of 15 V in both the trap and transfer cells. At these high velocities the energy transferred on collision may help to facilitate the exchange reaction.

Between 750 - 900 $m\,s^{-1}$ ion transmission begins to suffer but for those that do survive, particularly in the transfer region, a further decrease in uptake is observed. The uptake decrease is also accompanied by significant envelope broadening, as can be seen for high wave velocities in Fig. 4.4A. A distribution broadening represents a differing ability to exchange, either as a result of structural differences or the translational energy range of the ions. If the fast waves cause ions not to become

effectively trapped or even induce minor structural changes then uptake broadening could be observed.

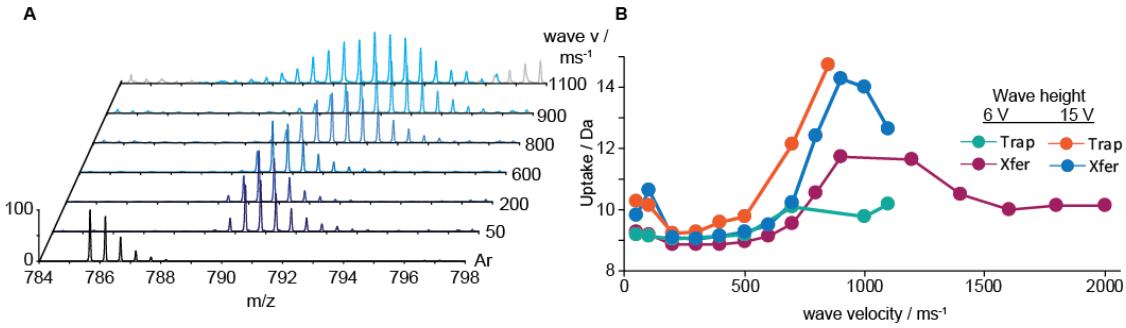


Figure 4.4. Manipulation of peptide deuterium incorporation. **A** Gas-phase labelled spectra of glufibrino peptide 2+ under a constant ND_3 pressure (10 µbar) with variable wave velocities. **B** Uptake determination across transmittable wave velocity range in the trap and transfer (Xfer) TWIGs at wave heights of 6 V and 15 V. Spectra seen in **A** correspond to transfer gpHDX data at 15 V. Grey peaks in the top spectrum are not associated with the peptide, no quadrupole isolation was used.

To determine whether protein uptake levels could be manipulated in the same manner, as for peptides, the 28 kDa monomeric green fluorescent protein (GFP) was exposed to 10 µbar ND_3 inside the same instrument ion guides. Figure 4.5 (bottom) shows the native 9+ ion under non-labelling Ar conditions. At low wave velocities a large m/z shift corresponding to an uptake of 81 Da is observed. Increasing the wave velocity from 25 $m\,s^{-1}$ to 400 $m\,s^{-1}$ we see a gradual decrease in uptake corresponding to ∼52 Da. At higher wave velocities, under a wave height of 6 V, transmission becomes poor with a much lower signal-to-noise. Increases in uptake were not observed with faster wave velocities, in the trap or transfer, when operating with a continuous beam in ToF mode. Differences between peptides and proteins are likely given the large difference in CCS altering the number of roll-over events and also the differing impact of TW electric fields on ion heating and structure.[11]

Faster travelling waves will cause the ions to evacuate the labelling cell faster owing to a greater ion velocity; the decrease in residency time will also impact the number of collisions with the deuterated bath gas non-linearly.[12] The number of collisions is dependent on the mean free path of the ion, which in turn is influenced

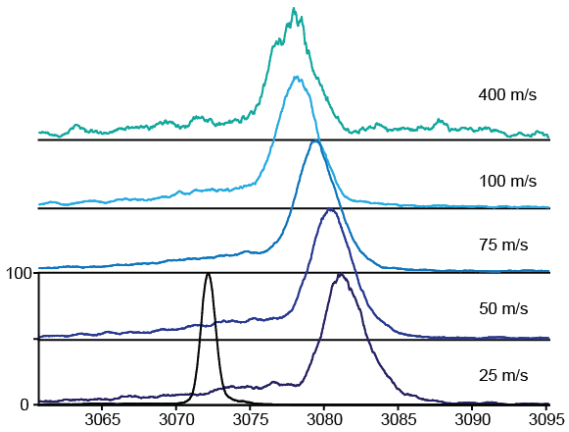


Figure 4.5. Travelling wave parameter control over protein labelling. Stacked native MS spectra of gas-phase labelled GFP 9+. A quadrupole isolation of the $[M + 9H]^{9+}$ under argon is shown in black (bottom). Ions were labelled in the transfer region with variable T-wave velocities at a constant wave height of 6 V and operated in ToF mode. The transfer wave velocity is indicated alongside the corresponding spectra with a decrease observed for increased speeds.Transmission becomes poor at higher velocities as is demonstrated at 400 $m\,s^{-1}$ (top).

by the influenced by the ratio of the ion velocity, ν_{ion}, to the relative velocity, ν_{rel}, of the collision partners.

$$\lambda = \frac{\nu_{ion} k_B T}{\nu_{rel} p \sigma} \tag{4.1}$$

Where, λ is equal to the mean free path of the ion-neutral collision, σ is the collision cross section of the ion-neutral, p is the buffer gas pressure,k_B is the Boltzmann constant and T the absolute temperature.

For a room temperature collision cell of ND_3, the relative velocity will always be greater than the ion velocity, $\nu_{rel} > \nu_{ion}$. At very high ion velocities the relative velocity will converge towards the ion velocity, i.e. ν_{ion}/ν_{rel} will approach 1. However, at low ion speeds the relative velocity tends towards the average thermal velocity of the gas, often much higher than the ion velocity. The value of ν_{ion}/ν_{rel} will increase as the ion velocity increases having the effect of increasing the mean free path, resulting in fewer exchangeable collisions.

Additionally, let us consider long-range interactions between collision partners beyond the summed van der Waals radii. The average velocity, $\bar{\nu}$, of an Ar gas molecule is $\sim 25\sqrt{T}$ $m\,s^{-1}$ at temperature, T, corresponding to $\sim$430 $m\,s^{-1}$ at 293 K. If ions are moving at, or close to, the travelling wave velocity range 50 - 2000 $m\,s^{-1}$ then buffer gas molecules and ions are moving at a comparable speed. It has been

shown previously that when ion-neutral collisions are not treated as hard spheres but instead include a Leonard-Jones potential, there is a temperature dependence on the collisional cross-section.[13,14] This is because the relative velocity is greater at higher temperatures and so the time the collision gas has to interact and perturb the ion's trajectory is reduced, §2.1.6. The capture cross-section of an ion-dipole interaction is related to velocity by:[15]

$$\langle \sigma \rangle \propto \frac{1}{{\nu_{rel}}^2} + c \tag{4.2}$$

Where, $\langle \sigma \rangle$ is the angle averaged capture cross section, ν_{rel} is the relative velocity of the ion to the gas and c is a constant. If now we do not change the gas thermal velocity but increase the ion speed using the T-waves, a decrease in CCS will also be observed. This may, in part, explain the decrease in uptake observed at high velocities for proteins and the initial decrease seen in the peptide case.

The rate of gpHDX will be governed by the number of collisions, the capture cross section associated with the ion-neutral complex and also the kinetic energy of the collision effecting the complex lifetime. All of these factors will act to decrease uptake as the ion velocity increases for the reasons discussed here and in Section 2.4. Understanding the exchange behaviour for a large protein remains challenging but these experiments demonstrate that the ion's motion may be manipulated, allowing an element of control over an ion's uptake.

Many previous gas-phase HDX studies have utilised ion trap spectrometers in which the residence time can be tuned by the trapping duration.[9] In our Synapt G1 IM instrument ions are accumulated and stored in the trap ion guide for the duration of the previous IM cycle. We set out to use the trap cell, in IM mode, as a means of increasing the residency time. In Figure 4.6 we explored the effect of the TW conditions on the uptake whilst ions are stored for 33.6 ms (200 bins × 168 μs).

Travelling wave velocities were varied from 50 - 250 $\mathrm{m\,s^{-1}}$ and wave heights increased from 0.2 - 2 V; where sufficient transmission was possible the uptake

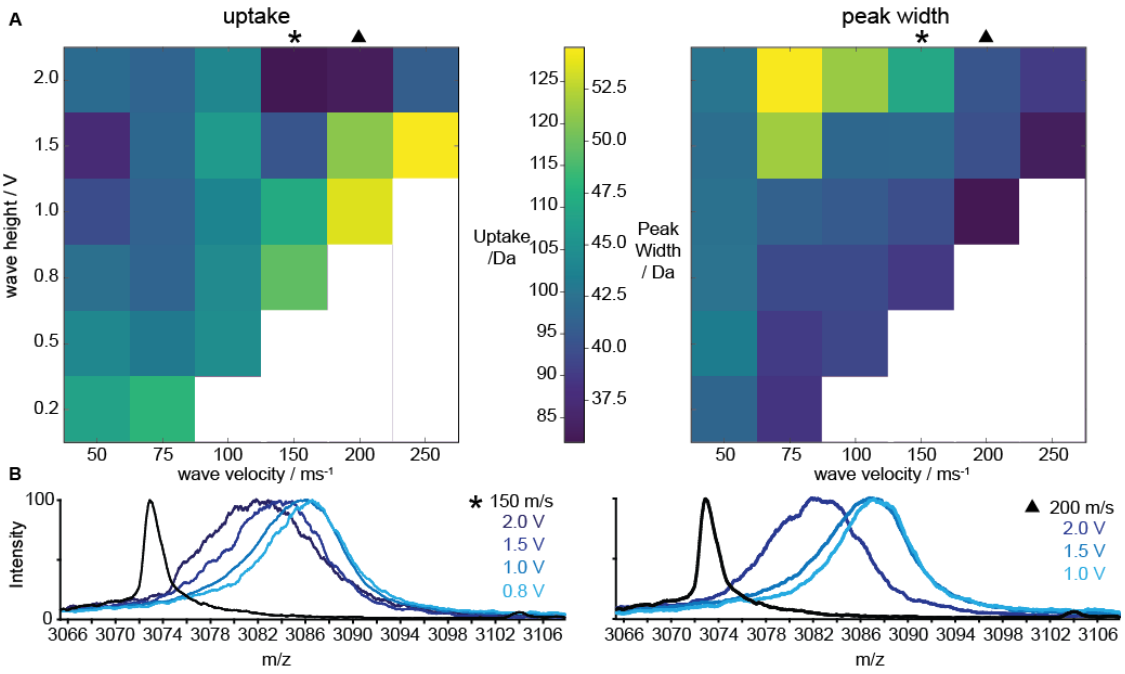


Figure 4.6. Travelling wave effects on the gpHDX labelling efficiency in a trapped ion mode. A Plots to show the impact of T-wave velocities and heights on the uptake efficiency and peak broadening for 9+ GFP in the trap region. Data could not be acquired for those regions in white due to poor signal transmission. **B** Native MS spectra of two selected wave velocities, 150 and 250 $\mathrm{m\,s^{-1}}$, demonstrating the decrease in uptake centroid with increasing wave height. A negative control using argon is shown in black with different wave conditions shown in blue.

centroid and peak width were plotted, Fig. 4.6A. Moving from low to high wave heights, for a given velocity, we see a decrease in uptake. The magnitude of this change is most evident for higher velocities. The spectra in Figure 4.6B demonstrate these observations for the wave velocities 150 and 250 $\mathrm{m\,s^{-1}}$, where blue spectra correspond to the deuterated species and the black peak, around 3074 *m/z*, is the argon reference. The peak width loosely inversely correlates with the uptake, this is not unexpected as the uptake tends towards the maximum. Comparison of these results with the continuous ion beam data presented in Figure 4.5 we note that for a given wave height, the uptake also tends to decrease with increased velocities. In the case of IM mode, ions are trapped and unable to evacuate the labelling cell. The waves must therefore impact the ion motion and/or the structural reactivity,[11,16] a further comment is made in Chapter 6. It is desirable to operate under conditions where high uptake levels are achieved whilst still maintaining narrow peak widths.

4.3.4 Considerations for ion mobility MS

In the knowledge that ion uptake can be tuned using the TW guides we then explored the parameter space in which a good ion mobility experiment can still be achieved. As ions are trapped in the trap cell prior to release it is these parameters on which we focus. The transfer guide during an IM experiment allows ion packets to pass through without hindrance, using high wave heights to maintain the separation achieved in the previous IM cell, and so little can be changed in this region. It was necessary for ions to undergo a narrow ion isolation, excluding adducts, to make the shifts in *m/z* more reliable and a centroid easier to obtain. This also has the added advantage that adducts cannot be lost due to activation and therefore influence the observed mass. When a wide isolation window is applied allowing sole transmission of the 10+ charge state, good ion transmission is possible across a range of trap wave heights 0.2 - 1.5 V at a fixed velocity of 150 ms^{-1}, Fig. 4.7A. However, when the quadrupole window is narrowed to only allow transmission of the $[M + 10\,H]^{10+}$ ion without adducts, transmission is only possible at the higher trap heights. Note, this is not of concern when operating in continuous ion mode (ToF mode). This effect is also demonstrated for the isolated 9+ ion under a stringent quadrupole selection, where fixing the wave velocity at 200 ms^{-1} but decreasing the trap wave height sees a dramatic loss in transmission, Fig. 4.7B. Working at lower wave velocities can alleviate these problems as shown for a low trap wave height, 0.2 V, Fig. 4.7C. The exact reason behind this behaviour is hard to determine from these experiments, though it appears the lack of charge density in the trap has an important role. When species are well isolated fewer ions enter the trap region and so when the IM gate is released the repulsive force pushing ions out is reduced. When space charge effects were reduced a higher wave potential was required to drive the ions out of the trapping cell towards the higher pressure IM region. This could also be achieved by slowing the waves down, which may more efficiently cause the ions to be trapped.

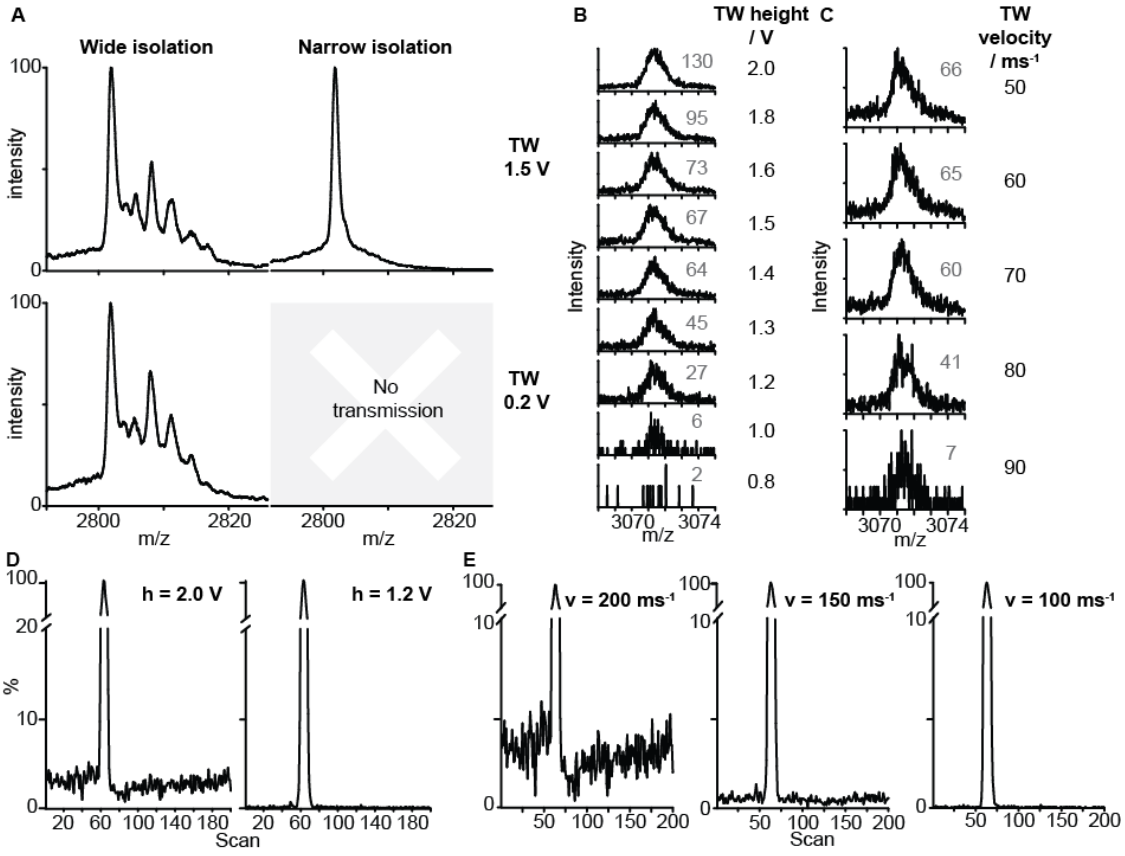


Figure 4.7. Travelling wave impact on transmission and subsequent mobility. A GFP basic 10+ native spectra illustrating the transmission efficiency in ion mobility mode whilst under a wide (LM:12, HM:14) or narrow (LM:15, HM:17) quadrupole selection with high or low trap wave heights (h). No transmission was possible for a narrow isolation and low wave heights. **B** Effect of trap wave height on the transmission of native WT GFP 9+. Spectra were acquired for 100 scans under ion mobility mode with a fixed wave velocity (v) of 200 ms^{-1}, total ion counts are recorded in grey next to each spectrum. **C** Decreasing TW velocities improves ion transmission at a wave height of 0.2 V. High wave heights result in a high baseline in the mobilogram, **D**, at a velocity of 200 ms^{-1}, which can be alleviated with decreasing wave velocities, **E**. Mobilograms have a split axis to highlight the high background signal, split at 20% and 10% for panels **D** and **E** respectively.

Having shown that when operating with a narrow quadrupole window either slow travelling waves or high wave heights are required, we demonstrated how these parameters impact the mobilogram. High wave heights are shown to cause an unacceptably high background signal, around 4%, across all 200 bins in the IM spectrum. This can be alleviated by decreasing the wave height as seen in Figure 4.7D. Therefore, if high wave heights are required for transmission and/or labelling efficiency, then decreasing the wave velocity may enable this. Figure 4.7E shows a mobilogram under a trap wave height of 2.0 V and a wave velocity of

200 ms^{-1} which also has a baseline of ~4 %. Reduction of the wave velocity to 100 ms^{-1} reduces the background to near zero. These effects suggest ions are being forced over the gating potential and causing a background of ions throughout all bins.

To summarise, whilst controlling the uptake of a species using the travelling wave parameters is effective one has to also optimise the conditions to aid transmission when ion currents are low, and maintain a good signal-to-noise in mobility space. High wave heights may be used to alleviate transmission difficulties, however to preserve a good mobilogram this must also be coupled to low wave velocities. Tuning the TW parameters is therefore a compromise between transmission and labelling efficiency.

4.3.5 Gas-phase HDX control experiments

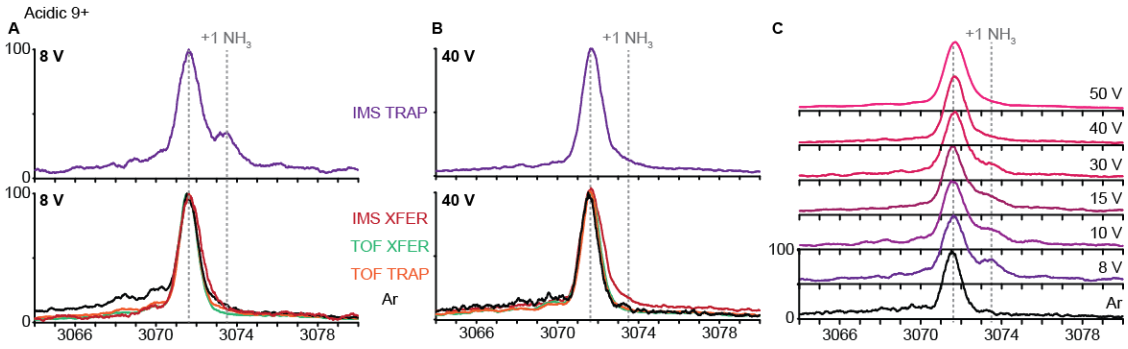


Figure 4.8. Undeuterated control experiments To determine whether ammonia can form gas-phase adducts, GFP 9+ ions were quadrupole selected and passed through either the trap or transfer labelling cells containing a partial pressure of 12 µbar NH_3, **A** & **B**. This was performed in both ToF,trap (orange) and transfer (green), and IMS, trap (Palatinate purple) and transfer (red), mode. An argon control is shown in black and the addition of a single ammonia adduct is marked in grey. Under a low labelling cell collision voltage (8 V), **A**, an adduct is seen when labelling occurs in the trap in IMS mode. At higher energies (40 V), **B**, the adduct is lost. **C** Collision energy ramp, 8 - 50 V in the trap in IM mode shows the decrease in ammonia adduct formation at higher collision voltages.

As uptake is calculated from the increase in m/z relative to the argon control it was important to establish that the measured shift was due to HDX and not adduct formation of the highly basic ammonia. Un-deuterated ammonia, NH_3, was used as a control for the ND_3 labelling partner. Experiments were performed at pressures typically used during a native experiment (~10^{-6} bar) in the two labelling cells, in

both gated IM and continuous beam ToF mode. Spectra in Figure 4.8 show that when ions are continuously passed through the labelling cell, both trap or transfer, no adducts are observed. The argon reference is shown in black and ToF mode spectra overlay well, Fig. 4.8A. In IM mode ions are gated and trapped in the trap cell before release into the IM cell followed by the lower pressure transfer cell to the pusher region. In the transfer region ion packets pass through without any trapping. As such when ions are labelled in the transfer in IM mode the overlay with argon is also very good. When labelling occurs in the trap cell however (purple trace), the presence of an ammonia adduct is observed. This can be mitigated by increasing the energy into the collision/labelling cell. Figure 4.8B demonstrates this point at 40 V, compare to A at 8 V, where labelling in all regions and modes display no adduction. If labelling is to be performed in the trap cell when IM gating is occurring, one has to be mindful of this small adduct contribution to the m/z shift at low energies, Fig. 4.8C. Otherwise, adducts should not be of concern when ion transmission is continuous.

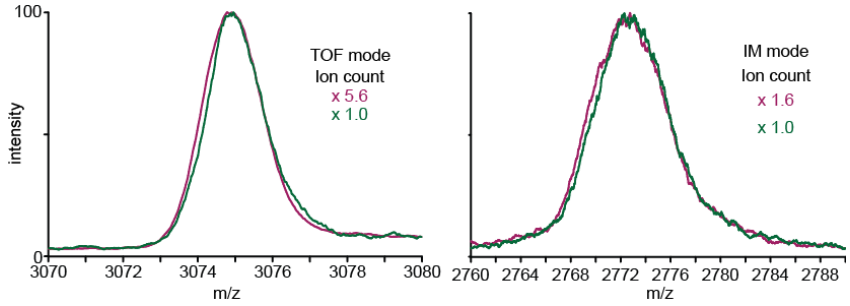


Figure 4.9. Effect of ion current on labelling. Gas-phase labelled GFP 9+ in ToF mode (left) and GFP 10+ in IMS mode (right). Higher ion currents are shown in purple and lower currents in green, corresponding to a 5.6 increase for ToF mode and 1.6 increase for IM mode. No significant difference is noted.

One final consideration on the uptake efficiency is in relation to the space-charge density of the ions. As the capture event is determined by the local electrostatics between the protein ion and neutral labelling partner, if ions are significantly close in space this may influence neighbouring energy potentials. We performed a simple experiment to verify the space-charge densities we work with have little bearing on the uptake ability. By attenuation of the ion-beam prior to the labelling cell we

see no impact on uptake. The ion beam on this generation of instrument can be attenuated by the frequency of which a gating lens has a voltage applied. Figure 4.9 demonstrates this effect for 9+ GFP in ToF mode and 10+ GFP in IMS mode, where signal was attenuated by a factor of 5.6 and 1.6 respectively, with no significant shift in peak centroid. Therefore we needn't be concerned about comparing experiments with different ion transmission, spray efficiency or concentration.

4.3.6 Gas-phase labelling is sensitive to protein surface chemistry

Earlier in this chapter we demonstrated that gas-phase HDX can effectively label, in a tuneable manner, the surface of a protein on a sub-millisecond time-scale. As gpHDX is reliant on the formation of an ion-neutral encounter complex it is highly sensitive to the exchangeable moieties present, the charge localisation and, the sterics and electrostatics of the local structure. Initial experiments were devised to test the sensitivity of gas-phase labelling to differentiate between distinct surface topologies. We used a system of green fluorescent proteins (GFP),[17,18] Fig. 4.10,

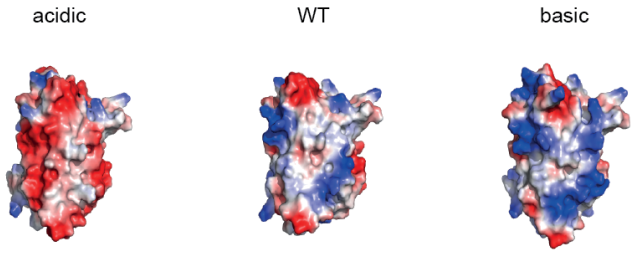


Figure 4.10. Surface electrostatics of model GFP variants. Homology models of the three variants, acidic, wild type and basic, GFP. Structures are coloured according to gas-phase electrostatics. Regions coloured red are residues electronegative in nature and blue are those that are electropositive.

which had been modified to enhance the surface electrostatics, by mutation of a number of Lys/Arg residues to Asp/Glu to result in negative surface in solution, 'Acidic' GFP, and the reverse to give a positive surface, 'Basic' GFP. The exact number of which are detailed in Table 4.1. These systems act as an ideal probe for gpHDX studies as the electrostatic nature and residue chemistry of the surface

are substantially different and yet the resulting gas-phase charge states are highly similar owing to electrospray charge deposition, Fig. 4.11. In addition, the mutants are within 5% of the WT native gas-phase CCS (wild type experimental average: $\sim$2484 $\pm$ 7 Å^2 , projection approximation (PA) theoretical: $\sim$2250 Å^2)[iii] and all possess the strict native fold necessary for maintaining an active fluorophore.[17] These features allow for an effective comparison between ion HDX and surface chemistry, uncoupled to other determinants.

Table 4.1. Exchangeable residues of the GFP variants. The number of each amino acid residue capable of gas-phase exchange in the GFP variants. The associated number of fast and slow exchanging hydrogens are also provided. Residues are grouped by acidic, basic and neutral residues in solution.

Residue	Acidic			Wild type			Basic		
	Number	Labile Hydrogens		Number	Labile Hydrogens		Number	Labile Hydrogens	
		Fast	Slow		Fast	Slow		Fast	Slow
D	25	25	25	17	17	17	12	12	12
E	24	24	24	15	15	15	14	14	14
C term	1	1	1	1	1	1	1	1	1
H	16	16	16	16	16	16	16	16	16
K	13	26	26	19	38	38	25	50	50
R	6	18	24	8	24	32	16	48	64
N term	1	2	2	1	2	2	1	2	2
C	2	0	2	2	0	2	2	0	2
S	10	0	10	10	0	10	10	0	10
T	16	0	16	17	0	17	18	0	18
Y	9	0	9	8	0	8	9	0	9
N	9	0	18	13	0	26	11	0	22
Q	7	0	14	7	0	14	5	0	10
Total	139	112	187	134	113	198	140	143	230

All residues with an associated labile side-chain hydrogen are shown in Table 4.1, with the maximum number capable of exchange provided for each mutant. In addition the number of fast and slow exchanging hydrogens corresponding to these residues are given. Residues are grouped depending on whether they are from a

[iii]The experimental average CCS was determined from the charge states, 8+, 9+ and 10+. The CCS was calculated using a modified Waters synapt with a linear drift cell. Further discussion and calculations are presented in Chapter 6.

nominally acidic, basic or neutral residue. All three mutants have approximately the same number of exchangeable sites but begin to differ significantly on the number of associated hydrogens. The basic mutant has by far the most fast exchanging sites, 143, followed by the wild-type and acidic with 112 and 113, respectively. This is simply due to the greater number of labile hydrogens in the basic moieties relative to the carboxylic acid residues. If we consider the theoretical maximum including fast and slower exchanging hydrogens we see a decreasing number progressing from basic to WT to acidic. We set out to test whether gpHDX is able to differentiate between these surfaces and to determine if the exchange behaviour matched these trends. Native mass spectra were acquired for each of the GFP mutants and yield

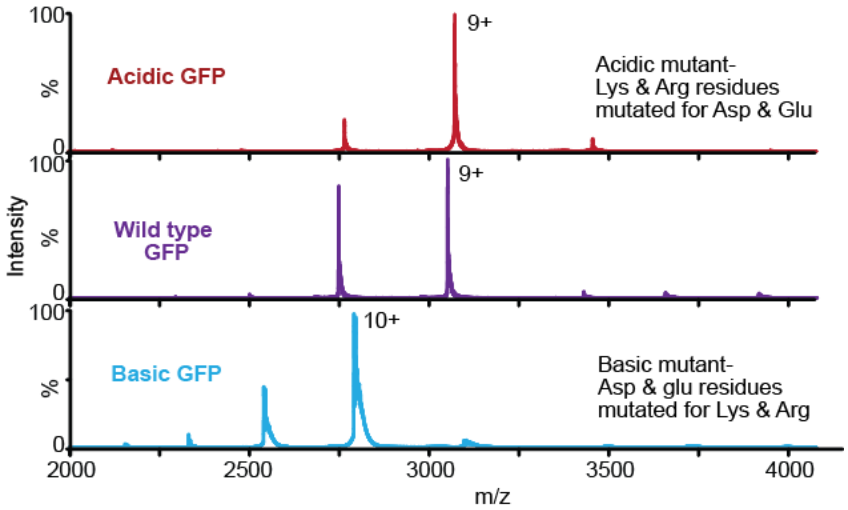


Figure 4.11. Native MS spectra of model GFP variants. Native mass spectra for three GFP mutants, acidic, wild type, and basic, possessing different surface chemistries.

narrow charge distributions centred on 9+ for the acidic and wild-type forms, the basic is offset by one charge giving an abundant 10+ charge state. The 10+ ions from all species could be readily probed for comparison whereas the basic required in-source charge reduction[19] to obtain a significant population of 9+ ions. All ions where first isolated in the quadrupole under stringent mass selection conditions to ensure the exclusive transmission of the $[M + nH]^{n+}$ ion and removal of all cationated ions. Species were then labelled, operating without ion trapping, in the trap ion guide through collisions with ND_3 gas and the uptake determined from the increase in m/z relative to an argon/NH_3 control under the same parameters.

The peak shape of the deuterated species remained largely symmetric but with the aforementioned envelope broadening. To determine the effect of labelling gas pressure on the exchange rate for the three GFP variants the partial pressure of ND_3 was increased from 1 µbar to 21 µbar. Gas-phase labelled acidic GFP 9+ spectra are shown in Figure 4.12A, shades of blue, relative to an argon control, black. A near-linear increase is observed for all species across this pressure range with little difference between charge states observed, Fig. 4.12B.

The most striking observation is the greater ability for the acidic species to exchange, across all pressures, despite the basic surface possessing a greater number of ion-neutral pairing sites and a significantly higher theoretical maximum uptake (basic: 230 D, WT: 198 D, acidic: 187 D). Suggesting that either basic sites significantly hinder exchange, or that carboxylic acid residues play a much more important role in exchange than often credited or anticipated.

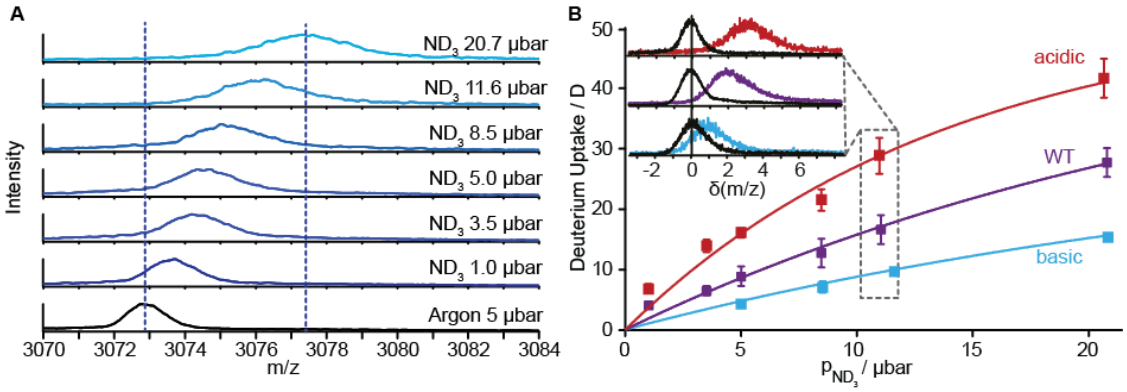


Figure 4.12. Gas-phase HDX discriminates between surface chemistries. A gpHDX labelled spectra of native acidic GFP 9+ with increasing ND_3 pressure, as indicated. Labelled spectra are shown in blue relative to the argon control in black. The m/z shift seen for 20.7 µbar corresponds to an uptake of 40 Da. **B** A plot of uptake against ND_3 pressure for all charge states of the acidic (red), WT (purple) and basic (blue) GFP mutants. An inset is shown of the mass spectra, on a m/z axis relative to Ar, at 11.6 µbar for the acidic and WT variants and 11.8 µbar for the basic mutant.

A simple charging model begins to explain observed H/D exchange levels

The faster exchange of the acidic surface is of particular interest as exchange is expected to be mediated by a positively charged site through the onium mechanism (§2.4), in addition to the limit of a single exchange event occurring with the labile

carboxylic hydrogen. Neighbouring moieties can stabilise this interaction and nearby labile side-chains may abstract a deuteron indirectly.[20] Calculations also show that a direct complex may be formed with a carboxyl site with comparable binding affinity, however subsequent exchange is thought to be slower.[21] As such we will consider charged amine, guanadino groups and neutral carboxylic acids as exchangeable sites.

Accurate prediction of the charge isomers can be a challenge[22] but, with Coulombic repulsion and stabilising interactions aside, charge will tend to reside where the gas-phase basicity, GPB, is greatest. Carboxylate residues, arginine, lysine, and histidine have the respective basicites:[23] 1328-1356, 1026, 978, 971 kJmol^{-1}. The N-terminus and C-terminus are considered to have a similar GPB to lysine and carboxylic acids. Assigning charge in this manner, carboxylate groups will exist in their neutral form followed by the charging of the Arg, Lys and His residues which possess a comparable GPB, ∼5% difference. If charge is assigned in this way the number of exchange competent sites is as follows, acidic: 60, WT: 43, basic: 37, Table 4.2. This corresponds to 86, 71 and 67 labile hydrogens for the acidic, WT and basic species respectively. Both the number of exchange sites and the labile hydrogens associated with them follow the observed trend as our experimental data. Histidines are omitted from these calculations as the number remains constant between each variant and the GPB is slightly lower than for lysine which does vary significantly between mutants.

In addition, in the case of small peptide systems, arginine has been shown to sequester hydrogens and prevent exchange. This is thought to be because arginine has a high GPB and limits charge migration away from this site. Small systems are however limited in the number of charge states and configurations they possess making it likely arginine can act as a proton sink. In the case of a folded protein surface, with many possible charge permutations and often more chargeable sites than the net charge, the restrictive effect of arginine on uptake seems less likely. If charge is assigned as before, but arginine residues are discounted from our calculations the acidic uptake becomes ever more favourable

Table 4.2. Exchangeable GFP residues upon charge assignment. Residues or moieties are shown in their neutral or charged form. Exchange competent sites are highlighted in grey. The number of labile hydrogens associated with these residues is also displayed. The number of exchangable sites including the number of labile hydrogens descends from acidic to basic variants. In the scenario where arginine residues are assumed to be exchange incompetent the number of exchangable species are displayed in brackets. The order remains the same only more pronounced.

	Acidic		WT		Basic	
	# residues	# labile Hs	# residues	# labile Hs	# residues	# labile Hs
-COOH	50	50	33	33	27	27
$-COO^-$	0	-	0	-	0	-
Lys^+	4	12	2	6	0	0
Lys	10	-	18	-	51	-
Arg^+	6	24	8	32	10	40
Arg	0	-	0	-	6	-
Total exchangeable	60 (54)	86 (62)	43 (35)	71 (39)	37 (27)	67 (27)

(acidic: 62 H, WT: 39 H, basic: 27 H), even when addressing the case of maximal uptake, the difference between variants is small and doesn't account for the observed trend (acidic: 163 H, WT: 166 H, basic: 166 H). This demonstrates how charge is distributed and assigned influences the rate of uptake; a simple count of the exchangeable species, even omitting the sequestering arginines, wouldn't yield the observed difference.

If we were to switch the polarity of the ion we would expect a decrease in the uptake magnitude as fewer positively charged sites are likely to exist, owing to the resulting net negative charge. In addition, a greater number of neutral carboxylic acid groups will become deprotonated and be rendered exchange incompetent. As discussed in Chapter 2, deuteron abstraction from ND_3 by the negatively charged carboxylic acid would result in an unstablised amino anion intermediate. Figure 4.13 corroborates these expectations having exposed the acidic and wild type variants to ND_3 whilst in negative polarity. The 8- charge state was sufficiently abundant for the acidic and wild type species to be probed and compared to the 8+ acidic ion. The negative ions were exposed to ND_3 over pressure range 3 - 84 μbar. Adducts were not observed over this range in negative mode. Both negative ion variants

display non-linear uptake and tend towards uptake saturation ∼25 Da and ∼10 Da, for the acidic and wild type respectively. Lower uptake is once again observed for the wild type variant relative to the acidic species. This is likely due to the greater number of carboxylic acid residues in their neutral form for the acidic mutant. The uptake for the negative ions is lower than for the positive ion counterpart across the same pressure range. As ammonia is highly basic forming an encounter complex with a positively charged species will be energetically favourable. Whilst the net charge is negative under negative polarity that is not to say positive charge centres do not exist. It is highly likely that it is this difference in the number of positively charged centres that results in a decrease in negative polarity.

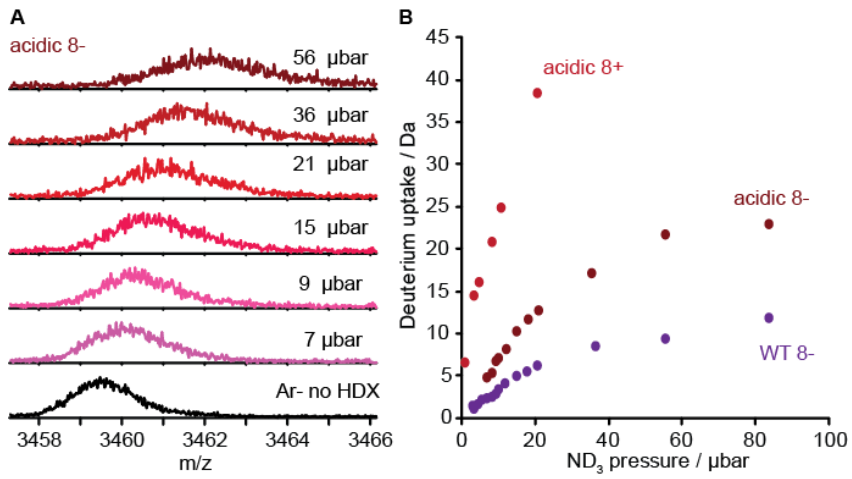


Figure 4.13. Impact of negative ion generation on GFP mutant exchange ability **A** Gas-phase labelled 8- ions (red) over the ND_3 pressure range indicated. **B** A plot of deuterium uptake against ND_3 pressure for the 8- ions of acidc and wild type (WT) GFP. The uptake for 8+ acidic GFP is shown for comparison, data also shown in figure 4.12.

Our data shows that gpHDX is capable of discriminating between variants of different surface chemistry and the resulting uptake can be rationalised using simple charge assignment. Prediction of a protein's uptake remains very difficult owing to the complexity and the inter-dependency of residues on the surface of a protein, but here we demonstrate the importance of carboxylic acid sites in promoting exchange. This demonstration that gpHDX is sensitive to local surface differences, irrespective of charge state, ion size, and secondary structure, founded the basis of our later work in probing gas-phase structure.

4.4 Conclusion

In this Chapter, I have described the implementation of gas-phase hydrogen/deuterium exchange to a native MS workflow. Peptides and small proteins were shown to have been labelled with high efficacy through the introduction of ND_3 to one of the instruments collision cell travelling wave ion guides. We also demonstrated the ability for the travelling wave potentials to effect the degree of uptake by changing the ion's velocity upon collision, though electric field structural perturbations cannot be ruled out.

The trap travelling wave conditions were found to need optimisation when a stringent quadrupole was applied and operated with trapped ions prior to ion mobility. Low wave heights, whilst may be desirable for labelling, can require low wave velocities to aid transmission for low ion currents exiting the quadrupole. If high wave heights are used instead, then low wave velocities may also be required to prevent ion roll-over into the mobility cell and manifesting as a high base-line.

Control experiments demonstrated that gas-phase adduct formation of the basic ammonia is of little concern, for these systems and under the given conditions, when ions continuously pass through the labelling cell without trapping in place. When contained ions were labelled in the trap a small amount of adducts were observed, though this could be suppressed with increased collision energies.

Finally, the methodology was applied on a series of model GFP variants with significantly different surface electrostatics and chemistries. Gas-phase labelling was highly sensitive to the different topologies which were comparable in size and of the same charge. The ability to distinguish between mutants was also evident in negative polarity though to a lesser degree. The degree of uptake the variants displayed was of some initial surprise with decreasing rate of exchange in the following order: acidic, wild-type and basic. The basic surface is significantly enriched with basic residues which contain many labile hydrogens, and likely to be a charge centre suitable for exchange, making the theoretical maximum uptake the reverse trend to that experimentally observed. This can be rationalised by assigning charge

according to gas-phase basicity and summing the number of exchangeable residues and associated labile hydrogens. Doing so in this manner reverses the expected exchange from the theoretical maximum. Previous reports on the behaviour of arginine residues to prevent exchange through sequestering charge may have a role and would only emphasise the predictions formed when charge has been assigned. Here we have shown that gas-phase HDX is capable of differentiating between different surfaces and forms the basis of the following chapters.

4.5 Methods

Gas-phase HDX modifications

Deuterated ammonia gas, ND_3, was purchased from CK Isotopes (Desford, UK) in a lecture bottle containing 10 L 99% deuterated ammonia gas purchased fitted with a CGA 180 CONCOA 422 series regulator (CK isotopes). Un-deuterated ammonia gas, NH_3 was purchased from CK gases (Ibstock, UK) in a 10 L lecture bottle. Gas-line modifications to leak NH_3/ND_3 gases into the trap and transfer ion guides were made from stainless steel 1/4 inch and 1/8 inch tubing. A 1/4 inch low flow metering valve with Vernier graduation (Swagelok, London, UK) were fitted to control high pressures from the regulator. A total of 1.5 m (3 x 0.5 m lengths) of 1/16 inch PEEKsil^{TM} capillary was fitted to restrict flow further and connected using a T-piece junction with a three-way switching valve. Trivial switching between NH_3 and ND_3 was enabled using a two-way switching valve (Swagelok). An additional Pirani pressure gauge (Edwards, Burgess Hill, UK) was fitted to the transfer collision cell to provide a local pressure read-out. ND_3 could be leaked into the trap and/or transfer cells using two three-way stainless steel switching valves (Swagelok).

For additional information regarding similar modifications on the low-pressure regions of the Synapt G1 and G2 instruments to enable gpHDX-MS, I refer the reader to previous papers.[1,5]

Materials

GFP variant constructs were re-cloned into pET vectors and proteins were expressed and purified similar to previously published protocols with more details provided below.[17,18] Plasmids were a gift from David Liu (HHMI, Dept. of Chem. and Chem. Biol., Harvard University, MA, USA). Protein was expressed in BL21(DE3) cells for 3 hours at 37 °C. Cells were lysed using a microfluidiser in 20 mM Tris (pH 8), 0.3 M NaCl. Protein was purified by immobilised metal-affinity chromatography using a HisTrap HP column (GE Healthcare, Amersham, UK) using 50 mM tris (pH 8), 20 mM imidazole, 0.3 M NaCl for protein binding and eluted with the same buffer containing 0.5 M imidazole.

For the basic variant, the eluate was diluted 5 fold into 50 mM tris (pH 8), and 2 M NaCl was added to adjust the final concentration to 0.25 M NaCl. The protein was then further purified by cation-exchange chromatography using a HiTrap SP HP column (GE Healthcare), and eluted using a gradient to 1 M NaCl. For the wild type and acidic variant, protein was dialysed into 50 mM tris (pH 8) and purified using anion-exchange chromatography using a HiTrap Q HP column (GE Healthcare) with an elution gradient to 1 M NaCl.

All GFP variants were further purified ready for ESI by size-exclusion chromatography in 0.2 M ammonium acetate (pH 6.8) using a Superdex 75 10/300 GL column (GE Healthcare).

All other samples and reagents were purchased from Sigma-Aldrich (Gillingham, UK) including lyophilised protein, peptide and sugar samples: leucine enkephalin, [Glu^1]-fibrinopeptide B (*human*), melitriose and ubiquitin (*bovine erythrocytes*). Peptides and sugars were dissolved in 50:50 methanol/water with 0.1% v/v formic acid. Ubiquitin was dissolved in LC-MS grade water and diluted to 10 μM in 200 mM ammonium acetate. Additional desalting steps were performed using Micro Bio-Spin columns as per the manufacturer's instructions (Bio-Rad Laboratories, Hercules, CA, USA).

Native mass spectrometry

Samples were ionised in positive ionisation mode, or negative were explicitly stated, in H_2O/MeOH for peptides and sugars or 200 mM NH_4OAc for proteins. Samples were prepared and loaded in borosilicate capillary emitters made in-house,[24] and ionized using a nanoflow Z-spray ESI source (Waters, Wilmslow, UK) mounted on a Synapt G1, or G2 for cone-labelling experiments Fig 4.3, HDMS mass spectrometer (Waters, Wilmslow, UK). Spectra were typically acquired when operating in ToF mode, or ion mobility mode when stated, according to published protocols.[25] All samples were sprayed with the source block at room temperature. Instrument parameters are listed below for all presented data. Where ND_3 pressures are given the corresponding negative control in Ar was at the same pressure.

Figure 4.2A LeuEnk, voltages: capillary 1.4 kV, sampling cone 10 V, trap activation 4 V, transfer activation 4 V. Pressures: backing pressure 2.3 mbar, trap ND_3 10 μbar, transfer Ar 6 μbar. Figure 4.2B GluFib, voltages: capillary 1.1 kV, sampling cone 40 V, extraction cone 5 V, trap activation 4 V, transfer activation 0 V. Pressures: backing pressure 2.3 mbar, trap ND_3 8.0 μbar, transfer Ar 8μbar. Figures 4.2C & 4.3A melitriose, voltages: capillary 2.0 kV, sampling cone 40 V, extraction cone 10 V, trap activation 20 V, transfer activation 4 V. Pressures: backing pressure 1.8 mbar, trap ND_3 4.7 μbar, transfer Ar 12μbar. Figure 4.2D ubiquitin, voltages: capillary 1.4 kV, sampling cone 10 V, extraction cone 3 V, trap activation 8 V, transfer activation 4 V. Pressures: backing pressure 3.8 mbar, trap ND_3 30 μbar, transfer Ar 20μbar. All spectra were processed using a smooth window of 1 times 1 (1×1), except ubiquitin which was (2×1).

Figure 4.3B (upper) melitriose, voltages: capillary 2.1 kV, sampling cone 175 V, extraction cone 10 V, trap activation 5 V, transfer activation 0 V. Pressures: backing pressure 1.8 mbar, trap ND_3 23 μbar, transfer Ar 12μbar.Figure 4.3B (lower) melitriose performed on a Synapt G2, voltages: capillary 1.5 kV, sampling cone 60 V, extraction cone 1 V, trap activation 5 V, transfer activation 0 V. Pressures:

backing pressure 2.6 mbar, make-up flow gas (ND_3/D_2O vapour) 50 mL/min, trap Ar 63 μbar, transfer Ar 19μbar. Spectra were smoothed by 1 × 1.

Figure 4.4 travelling wave heights were set to 6 V and 15 V in the trap and transfer and wave velocities were varied between 50 and 1100 $m\,s^{-1}$ voltages: capillary 2.8 kV, sampling cone 40 V, extraction cone 5 V, trap activation 12 V, transfer activation 4 V. Pressures: backing pressure 2.1 mbar, trap/transfer ND_3 10 μbar. Figure 4.5 transfer wave height was fixed at 6 V and the wave velocity varied from 25 - 400 $m\,s^{-1}$. Voltages: capillary 1.4 kV, sampling cone 10 V, extraction cone 2 V, trap activation 25 V, transfer activation 4 V. Pressures: backing pressure 3.7 mbar, trap Ar 18 μbar, transfer ND_3 19 μbar. Figure 4.5 was acquired in ion mobility mode with trap labelling and a pusher frequency of 180 μs. Trap wave velocities and heights were varied as indicated. Voltages: capillary 1.4 kV, sampling cone 10 V, extraction cone 3 V, trap activation 10 V, transfer activation 4 V. Pressures: backing pressure 3.9 mbar, set ND_3 trap 12 μbar, set transfer Ar 12 μbar. Spectra were smoothed by 1 × 1.

Figure 4.7 was acquired in ion mobility mode with trap travelling wave heights and velocities were varied as described in the figure caption and isolation windows are indicated. Voltages: capillary 1.4 kV, sampling cone 10 V, extraction cone 3 V, trap activation 8 V, transfer activation 4 V, IMS DC bias 15 V. Pressures: backing pressure 3.8 mbar, trap/transfer Ar 30 μbar. Spectra were not smoothed.

Figure 4.8 was acquired in ToF or ion mobility mode with NH_3 or Ar in the trap cell at 12 μbar. Trap activation voltages were varied as shown. Voltages: capillary 1.4 kV, sampling cone 10 V, extraction cone 3 V, transfer activation 4 V, IMS DC bias 15 V, trap velocity 200 $m\,s^{-1}$, trap height 0.2 V. Pressures: backing pressure 3.7 mbar, transfer Ar 12 μbar. Figure 4.9 was acquired under the same conditions but with the trap filled with ND_3. Attenuation was achieved using the attenuation lens prior to the trap cell or sample dilution. Spectra were smoothed by 5 × 1 and 15 × 1 in ToF and IMS modes respectively.

Figure 4.12: Native GFP ions were isolated for each charge state in positive mode and labelled in the trap cell under the ND_3 pressure indicated. Voltages: capillary 1.4 kV, sampling cone 10 V, extraction cone 3 V, trap activation 8 V, transfer activation 4 V, trap velocity 300 $m\,s^{-1}$, trap height 3 V. Pressures: backing pressure 3.8 mbar, transfer Ar 12 μbar. Negative mode spectra(Figure 4.12) were acquired in the same manner with the following differences: capillary 1.6 kV and extraction cone 3 V. Spectra were smoothed by 5 × 1.

Data analysis

Processing of mass spectra was carried out using MassLynx V4.1 software (Waters, Wilmslow, UK) with the mean filter smoothing indicated above. The deuterium content of the deuterium-labelled proteins was determined by the *m/z* shift relative to the non-deuterated control sample in the absence of deuteration (either NH_3 or Ar) and multiplied by the charge state of the peak. Peptide and sugar weighted centroids were calculated by copying the MassLynx spectral peak list into HDExpress[26,27] and using the inbuilt Excel (Microsoft, Redmond, WA, USA) macro.

4.6 Acknowledgements

I would like to thank my supervisor Justin Benesch for allowing me to follow my interests and develop the technique from the ground up. Advice from Ulrik Mistarz and Kasper Rand from the outset made establishing the method significantly less daunting. A special thank you to them also for hosting me for a number of weeks for a COST STEM research visit to support this work. I wish to thank Todd Mize for his practical support and advice in modifying the instrument. Kevin Giles and Jeff Brown provided helpful advice and discussions around T-wave transmission effects. To Georg Hochberg and Tim Allison for the advice I received during protein purification of the GFP mutants. Weston Struwe for advice when working with the saccharide systems. I thank previous group members, Art Laganowsky and Jon Hopper, for obtaining and cloning the GFP variants' plasmids. I thank Joseph Gault

for his critical review of this chapter. Finally to John Simmons for an interesting discussion around some of our findings.

References

[1] Rand, K. D.; Pringle, S. D.; Murphy, J. P.; Fadgen, K. E.; Brown, J.; Engen, J. R. Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Analytical chemistry* **2009**, *81*, 10019–28.

[2] Mistarz, U. H.; Rand, K. D. Installation , validation , and application examples of two instrumental setups for gas-phase HDX-MS analysis of peptides and proteins. *Methods* **2018**, 0–1.

[3] Mistarz, U. H.; Chandler, S. A.; Brown, J. M.; Benesch, J. L. P.; Rand, K. D. Probing the Dissociation of Protein Complexes by Means of Gas-Phase H/D Exchange Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2019**, *30*, 45–57.

[4] Breuker, K.; McLafferty, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase, 10^{-12} to 10^{2} s. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18145–52.

[5] Mistarz, U. H.; Brown, J. M.; Haselmann, K. F.; Rand, K. D. A Simple Setup for Gas-Phase H/D Exchange Mass Spectrometry Coupled to Electron Transfer Dissociation and Ion Mobility for Analysis of Polypeptide Structure on a Liquid Chromatographic Timescale. *Analytical chemistry* **2014**, *86*, 11868–11876.

[6] Uppal, S. S.; Beasley, S. E.; Scian, M.; Guttman, M. Gas-Phase Hydrogen/Deuterium Exchange for Distinguishing Isomeric Carbohydrate Ions. *Analytical Chemistry* **2017**, *89*, 4737–4742.

[7] Dookeran, N. N.; Harrison, A. G. Gas-phase H-D exchange reactions of protonated amino acids and peptides with ND_3. *Journal of Mass Spectrometry* **1995**, *30*, 666–674.

[8] Geller, O.; Lifshitz, C. A Fast Flow Tube Study of Gas Phase H/D Exchange of Multiply Protonated Ubiquitin. *The Journal of Physical Chemistry A* **2005**, *109*, 2217–2222.

[9] Wyttenbach, T.; Bowers, M. T. Gas phase conformations of biological molecules: the hydrogen/deuterium exchange mechanism. *Journal of the American Society for Mass Spectrometry* **1999**, *10*, 9–14.

[10] Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in travelling wave ion mobility resolution. *Rapid Communications in Mass Spectrometry* **2011**, *25*, 1559–1566.

[11] Morsa, D.; Gabelica, V.; De Pauw, E. Fragmentation and Isomerization Due to Field Heating in Traveling Wave Ion Mobility Spectrometry. *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 1384–1393.

[12] Ding, L.; Sudakov, M.; Kumashiro, S. A simulation study of the digital ion trap mass spectrometer. *International Journal of Mass Spectrometry* **2002**, *221*, 117–138.

[13] Ujma, J.; Giles, K.; Morris, M.; Barran, P. E. New High Resolution Ion Mobility Mass Spectrometer Capable of Measurements of Collision Cross Sections from 150 to 520 K. *Analytical Chemistry* **2016**, *88*, 9469–9478.

[14] Gabelica, V.; Marklund, E. Fundamentals of ion mobility spectrometry. *Current Opinion in Chemical Biology* **2018**, *42*, 51–59.

[15] Su, T.; Bowers, M. T. Ion-Polar Molecule Collisions. Proton Transfer Reactions of H_3^+ and CH_5^+ to the Geometric Isomers of Difluoroethylene, Dichloroethylene, and Difluorobenzene. *Journal of the American Chemical Society* **1973**, *95*, 1370–1373.

[16] Merenbloom, S. I.; Flick, T. G.; Williams, E. R. How hot are your ions in TWAVE ion mobility spectrometry? *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 553–562.

[17] Lawrence, M. S.; Phillips, K. J.; Liu, D. R. Supercharging proteins can impart unusual resilience. *Journal of the American Chemical Society* **2007**, *129*, 10110–10112.

[18] Thompson, D. B.; Cronican, J. J.; Liu, D. R. *Protein Engineering for Therapeutics Part B*, 1st ed.; Elsevier Inc., 2012; Vol. 503; pp 293–319.

[19] Hopper, J. T. S.; Sokratous, K.; Oldham, N. J. Charge state and adduct reduction in electrospray ionization-mass spectrometry using solvent vapor exposure. *Analytical biochemistry* **2012**, *421*, 788–90.

[20] Zhu, C.; Balta, B.; Aviyente, V.; Lifshitz, C. The Interaction of Protonated Diglycine with Ammonia: A Density Functional Theory Model Study. *The Journal of Physical Chemistry A* **2000**, *104*, 7061–7067.

[21] Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. Deuterium Exchange Reactions as a Probe of Biomolecule Structure. Fundamental Studies of Gas Phase H/D Exchange Reactions of Protonated Glycine Oligomers with D_2O, CD_3OD, CD_3CO_2D, and ND_3. *Journal of the American Chemical Society* **1995**, *117*, 12840–12854.

[22] Fegan, S. K.; Thachuk, M. A charge moving algorithm for molecular dynamics simulations of gas-phase proteins. *Journal of Chemical Theory and Computation* **2013**, *9*, 2531–2539.

[23] Marchese, R.; Grandori, R.; Carloni, P.; Raugei, S. On the zwitterionic nature of gas-phase peptides and protein ions. *PLoS computational biology* **2010**, *6*, e1000775.

[24] Hernández, H.; Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols* **2007**, *2*, 715–26.

[25] Kondrat, F. D. L.; Struwe, W. B.; Benesch, J. L. P. *Native Mass Spectrometry: Towards High-Throughput Structural Proteomics*; 2015; pp 349–371.

[26] Weis, D. D.; Engen, J. R.; Kass, I. J. Semi-Automated Data Processing of Hydrogen Exchange Mass Spectra Using HX-Express. *Journal of the American Society for Mass Spectrometry* **2006**, *17*, 1700–1703.

[27] Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. *Journal of the American Society for Mass Spectrometry* **2013**, *24*, 1906–12.

5

Probing the dissociation of protein complexes by means of gas-phase H/D exchange mass spectrometry

5.1 Abstract

Gas-phase hydrogen/deuterium exchange measured by mass spectrometry (gpHDX-MS) is a fast method to probe the conformation of protein ions. The use of gpHDX-MS to investigate the structure and interactions of protein complexes is, however, largely unharnessed. Ionising proteins under conditions that maximise preservation of their native structure (native MS) enables the study of solution-like conformation for milliseconds after electrospray ionisation (ESI), which allows for the heteroatom-bound non-amide hydrogens to rapidly exchange for a deuteron with ND_3 gas inside the mass spectrometer. Here, we explored the utility of using gpHDX-MS to examine protein-protein complexes and the feasibility of collision induced dissociation (CID) as a step towards greater localisation of the incorporated label. Protein complexes ranging from 24 kDa dimers to 395 kDa 24mers were analysed by gpHDX-MS with subsequent CID. We probed the effect of asymmetric charge-partitioning, during CID, on the redistribution of the labile deuterium label. In larger multimeric protein complexes the dissociated monomer showed a small but significant increase in deuterium. This indicates that intermolecular H/D migration occurs as part of the asymmetric partitioning of charge during CID. We discuss several models that may explain this increase in deuterium content and find that a model where only deuterium involved in migrating charge can account for most of the deuterium enrichment observed on the ejected monomer. Understanding this behaviour means the monomeric uptake can be measured accurately and be used as a proxy for the intact complex.

5.2 Introduction

Native mass spectrometry (MS) has become a powerful means to inform on the composition, stoichiometry, interaction, and subunit connectivity of protein complexes.[1–3] Using this methodology, protein complexes are ionised intact from solution, and the resulting native-like conformational states can be preserved in the gas-phase on a time scale ranging from milliseconds to seconds, enabling closer analysis of these conformations by MS.[4–7] Once inside the mass spectrometer, gaseous biomolecules can, in addition to the mass measurement, be probed by an array of gas-phase techniques.[8,9] These include ion mobility (IM) spectrometry,[4,10,11] and techniques that use gas-phase reactions such as proton-transfer reactions,[12,13] radical-based reactions,[14,15] or hydrogen/deuterium exchange (HDX), Chapter 4.

Having shown in the previous chapter that gas-phase HDX (gpHDX) is sensitive to the chemical make-up of a protein's surface we sought to explore how this could be applied to intact complexes with a view to informing on interfacial protection. Collision induced dissociation (CID), mechanistically discussed on page 45, is a powerful means of attaining subunit level information of a protein complex. By dissociation of a subunit, or ligand, one can often obtain the identity, binding stoichiometry and connectivity of the unit within the complex. An advantageous consequence of CID is the asymmetric charging concomitant with subunit ejection, allowing for accurate monomer assignment and spectral decongestion. The monomeric unit leaves with a significantly greater proportion of charge than its mass would suggest. The highly charged monomers, appearing at low m/z, can also be measured to a higher accuracy than for the intact complex owing to a narrower intrinsic isotope width, higher resolution at low m/z, increased ion transmission, and fewer adducts.

CID would therefore, in principle, be useful to inform on the gas-phase deuterium incorporation of each subunit. However, slow heating of ions through repeated bath gas collisions, is well known to result in proton migration, as discussed in Chapter 2. This presents a significant obstacle for using any collisional activation

to probe the incorporation of any labile label, such as deuterium. In this work we wanted to establish the level of such migration and whether its severity would prevent using CID products as a proxy for the intact incorporation.

In this chapter, we examine a series of protein homo-complexes, in the range 24 - 395 kDa, to show that gpHDX in combination with native MS can be used to study multimeric protein-protein complexes, including cytochrome c and α-lactalbumin dimers, transthyretin tetramers, HSP16.9 12mers, and HSP16.5 24mers. Gas-phase labelling followed by subsequent CID of the complexes made it possible to measure the deuterium content of both intact complex and dissociated subunits in the same experiment, and thereby approximate the number of heteroatom-bound non-amide hydrogens participating in the binding interfaces. In addition, upon analysis of larger multimeric protein complexes (12 – 24mers), a small but significant amount of intermolecular H/D migration was observed as part of the redistribution of charge during asymmetric partitioning, consistent with charge migration and the partial unfolding of the leaving subunit during activation.

5.3 Results and Discussion

5.3.1 Setup for gas-phase HDX-MS and CID

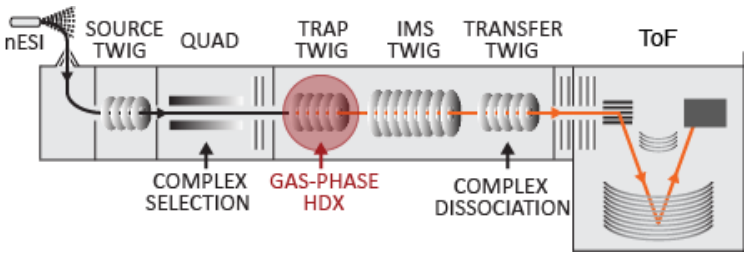


Figure 5.1. Gas-phase HDX-MS for studying protein-protein complexes. A Schematic of the mass spectrometer setup for performing gas-phase HDX-MS with subsequent complex dissociation by CID on a Q-ToF mass spectrometer equipped with travelling-wave ion guides (TWIG). Selected charge states of the complexes were isolated in the quadrupole, and gas-phase HDX-MS labelling was performed downstream in the trap TWIG by infusing ND_3 gas into the cell. Lastly, before mass analysis, the complex was dissociated in the transfer TWIG by CID with Ar. Figure modified from the co-authored publication Mistarz & Chandler *et al.*[16]

gpHDX combined with CID was performed on a modified quadrupole time-of-flight (Q-ToF) mass spectrometer (Waters Synapt), which allowed fast and

selective labelling of both intact protein complexes and their subunits (Figure 5.1). gpHDX was performed in the trap travelling wave ion guide (TWIG) which normally functions as a collision cell. By exchanging the standard Ar gas to ND_3 gas, it was possible to perform gas-phase HDX on a millisecond time scale as ions were pushed through the ion guide by the travelling wave (TW). By labelling in the trap TWIG with prior quadrupole selection, the possibility of subsequent complex dissociation with CID in the transfer TWIG was available (Figure 5.1). Experiments were performed by infusing ND_3 gas, NH_3 gas or Ar gas (reference/unlabelled sample) to the trap travelling-wave ion guide, as shown for the non-specific 9+ cytochrome C dimer, Figures 5.2A and 5.2B, respectively. The difference in the measured m/z of each ion from the labelled and unlabelled sample could be used to calculate its deuterium incorporation. No significant difference was observed for the relative distribution of monomer to dimer, charge state envelope, or adduct formation, between NH_3 gas and ND_3 gas in the trap.

5.3.2 Gas-phase HDX-MS and CID of protein-protein dimers

As an initial step, we sought to investigate how the labile deuterium is redistributed amongst the subunits of a dimer after CID. Cytochrome C, 12 kDa, forms non-specific dimers at high concentration.[17,18] By choosing an odd dimer charge state we can be sure that the peak corresponds exclusively to a dimer, as the monomer must adopt an integer number of charges. gpHDX was performed on the dimeric species in the trap TWIG with increasing collision energy in the transfer TWIG ranging 10 – 90 V. At the lowest collision voltage, 10 V, dimers were primarily present (>95%), and dissociation into monomers increased as a function of collision energy. At the highest collision voltages little dimer was observed in the mass spectrum (<0.5% dimer). gpHDX and subsequent CID were performed on the 9+ dimer of cytochrome C which dissociated both symmetrically into 5+/4+ monomers or more asymmetrically into 6+/3+ monomers. The increase in species m/z upon labelling by gpHDX is evident in Figure 5.3A.

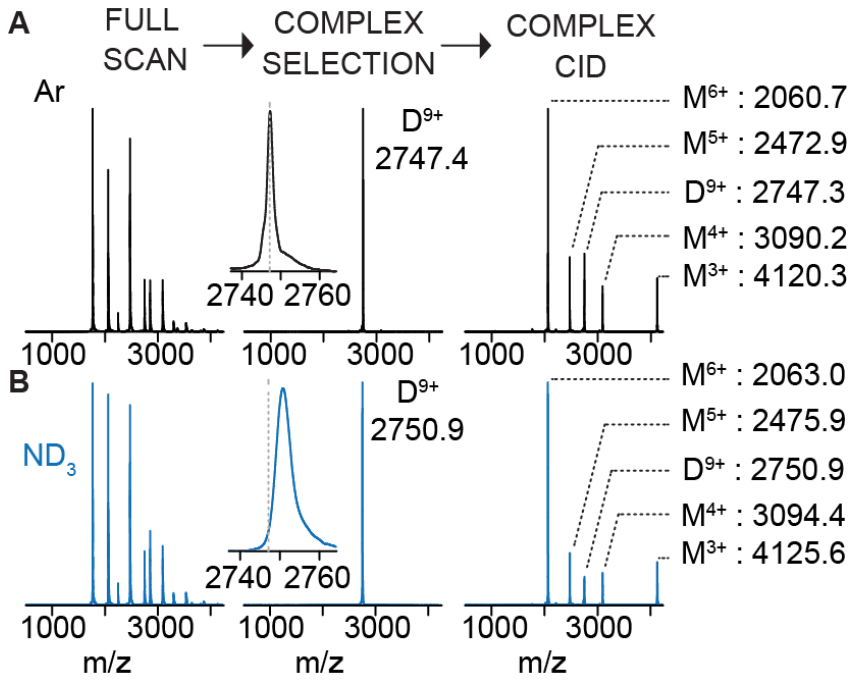


Figure 5.2. Gas-phase HDX and subsequent CID of a dimeric complex. Demonstration of gas-phase HDX and subsequent CID analysis for cytochrome C dimers (9+). Mass spectra are shown as a full range scan, quadrupole selected complex, and dissociated complex without **A** and with HDX **B**. An *m/z* shift relative to the unlabelled sample was observed and used to calculate deuterium uptake for both precursor dimer and dissociated subunits. Repeat experiments of initial data collection, data analysis and original figure created for the manuscript was provided by Ulrik Mistarz.[16]

The deuterium content for the intact dimer and dissociated monomer charge states are calculated for each species present across the transfer activation range. Fig. 5.3B. Uptake values are normalised to the number of subunits, N, present in the species. Intact cytochrome C showed a deuterium content of 51.4 - 55.2 D (25.7 – 27.6 D/subunit). The dissociated monomers contained approximately half the deuterium content of the precursor dimer, with a range of 25.1 - 28.8 D/subunit. This indicated that the deuterium label was equally retained by each subunit of the complex upon dissociation, irrespective of whether dissociation gave rise to symmetric (4+/5+) or asymmetric partitioning (3+/6+) of charge. Similar results were obtained for 11+ cytochrome C, which similarly dissociated predominantly asymmetrically (4+/7+). It should be noted for this small dimeric system, comparable deuterium content was measured for dissociated monomers of different charge.

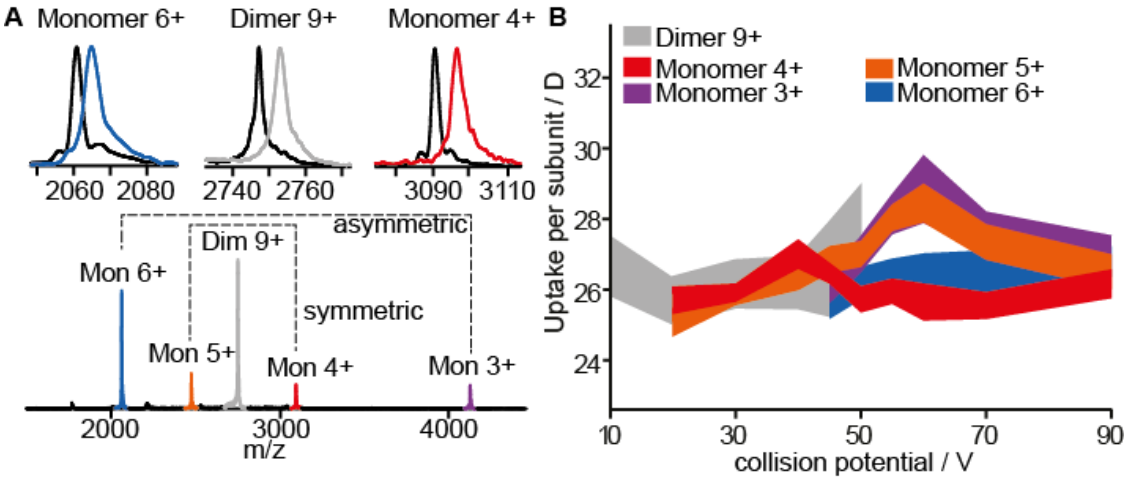


Figure 5.3. Deuterium distribution of after symmetric and asymmetric charge partitioning. A Mass spectra of the quadrupole selected 9+ cytochrome C dimer after gpHDX labelling in trap TWIG followed by dissociation in the transfer TWIG. Mass spectra of the 9+ dimer and associated dissociated monomers are shown for the symmetric case 4/5+ and the asymmetric case 3/6+ with insets providing a zoom for the latter. Dissociation spectra shown were acquired at 45 V in the transfer cell. Unlabelled species are shown in black with all labelled species displaying a mass shift shown in colour or grey. Deuterium content per subunit is shown, **B**, for the 9+ dimer (grey) and dissociated monomers of different charge states originating from the dimer (various colours). The deuterium content is shown as ribbon graphs where the midpoint and width correspond to the average deuterium uptake and 1 × SD, respectively. The experiments were performed in triplicate. Repeat experiments of initial data collection and data analysis performed by Ulrik Mistarz. Original figure created for the manuscript was provided by Ulrik Mistarz.

Identical experiments were performed for the 9+ and 11+ non-specific dimers of α-lactoglobulin as part of this study but were not the focus of my contribution.[16] The data were highly consistent with what was observed for cytochrome C. Thus, collisional activation and concomitant asymmetric charge-partitioning does not appear to induce measurable deuterium enrichment for the small homo-dimers studied.

5.3.3 Gas-Phase HDX-MS and CID to inform on protein-protein multimers

Transthyretin (TTR) is a 56 kDa homo-tetrameric protein formed from a dimer of dimers. TTR was ionised from native conditions and was observed primarily as a tetramer of charge states 13 - 15+. Gas-phase dissociation of the 15+ tetramer by CID (40 V) resulted in the asymmetric partitioning of charge per subunit. CID resulted primarily in monomers (6+/7+/8+/9+) and trimers (6+/7+/8+/9+),

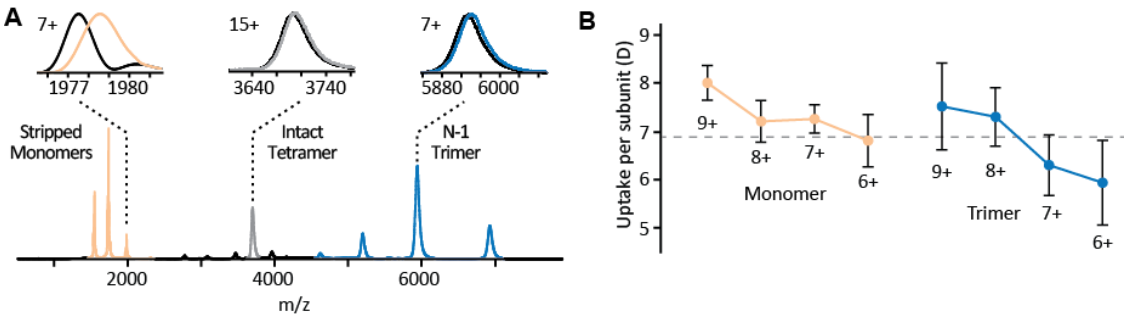


Figure 5.4. Dissociation of tetrameric transthyretin studied by gas-phase HDX-MS. Dissociation of tetrameric transthyretin studied by gas-phase HDX-MS. **A** Mass spectrum of quadrupole-selected 15+ tetramer of transthyretin following gas-phase HDX-MS and complex dissociation (magnified 10-fold for m/z >2500). Zoom-ins are of mass spectra of selected peaks, without (black) and with (coloured) gas-phase HDX-MS. Deuterium content per subunit is shown for precursor 15+ tetramer (dashed grey line) and dissociated monomers (orange) and trimers (blue) of different charge state **B**. Experiments were performed in at least three replicates, and the error bars denote the 1 × standard deviation. Repeat experiments of initial joint data collection, data analysis and the original figure created for the manuscript were provided by Ulrik Mistarz.

Figure 5.4A. Little difference between ND_3 and NH_3 as a collision partner was observed in the monomer:trimer ratio, charge state distribution, or salt- and solvent adduct formation. NH_3 was therefore found to be a preferential reference gas for the unlabelled control measurements for larger non-covalent protein-protein complexes. The deuterium content was measured at 27.6 D for the intact tetrameric complex, which corresponded to 6.9 D per subunit. A comparable deuterium content was measured for the dissociated monomers (6.8 – 8.0 D/subunit) and trimers (5.9 – 7.5 D/subunit) of different charge, Figure 5.4B. Dissociated TTR species (monomers and trimers) generally showed a small increase in deuterium content relative to the normalised intact tetramer (average ∼0.4 D/subunit). Furthermore, dissociated monomers showed a small average increase in deuterium of 0.5 D/subunit relative to dissociated trimers of the same charge. This could indicate a degree of inter-molecular H/D migration during dissociation resulting in minor deuterium enrichment of the ejected monomer subunit.

The experiments performed on TTR indicate that the expected H/D migration was occurring during CID, but for this asymmetric system as for cytochrome C and α-lactoglobulin the resulting subunit enrichment is limited.

5.3.4 Large Protein-Protein Multimers

The following sections are my further investigations and discussions building on the aforementioned work performed by Ulrik Mistarz and myself.

To further probe the effect of asymmetric charging, through CID, on subunit deuterium enrichment we studied two large monodisperse homo-oligomers from the small heat-shock protein (HSP) chaperone family. sHSPs are composed of low molecular mass monomers (15 - 40 kDa) that assemble into larger oligomerc complexes that can contain as many as 50 subunits in a mono- or polydisperse fashion.[19] These proteins have proven to be excellent systems for studying the mechanism of protein complex dissociation in the gas phase owing to their significant asymmetric charge partitioning.[20,21] Here, we have explored the gpHDX-CID methodology to study HSP 16.9, which forms a homomeric 12mer (201 kDa),[22] as well as HSP 16.5, a 395 kDa sHSP 24mer complex.[23]

Collision induced dissociation of HSP 16.9 gave rise to the highly asymmetric partitioning of charge through the sequential ejection of monomeric units. Initially, at collision voltages of 40 – 140 V, only singly stripped species were generated with the unfolded monomer chain leaving with more than half of the precursor charge. Further activation (160 – 180 V) resulted in the additional loss of a monomeric unit from the already stripped, N-1, species taking almost half of the remaining charge. For gas-phase labelling experiments the 32+ intact 12mer was isolated in the quadrupole and activated (160 V) to give both N-1 and N-2 species centred around the 18+ and 10+ charge state, respectively (Fig. 5.5A). The observed monomers, therefore, originated from both the intact 12mer and the N-1 precursors with the charge states following a bimodal distribution centred on 13+ and 9+, with the higher charge state series coming from the first dissociation event (12mer). HSP 16.5, 24mer, was also studied owing to its ability to dissociate monomers asymmetrically over two sequential dissociation events. The intact 48+ 24mer was quadrupole isolated, and the CID spectra were recorded over 70 – 180 V of activation. At 70 V HSP 16.5 dissociated into N-1 and N-2 stripped oligomers with a significant

amount of intact precursor still populated (30%), in addition to the corresponding monomers (Fig. 5.5B). The partitioning of charge was less asymmetric than for HSP 16.9 with the N-1 and N-2 monomers leaving with around a quarter and a third of the total charge, respectively. The monomers arose from all dissociation events and resulted in similarly charged species leading to a single charge state distribution for stripped monomers of HSP 16.5.

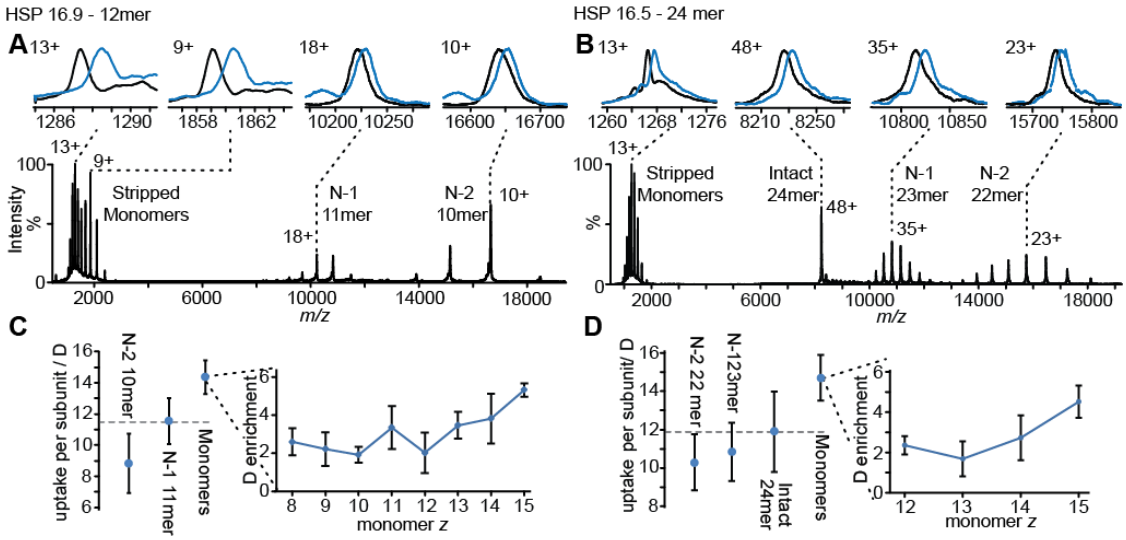


Figure 5.5. Gas-phase HDX and subsequent CID of sHSP oligomers. Mass spectra of quadrupole selected, gas-phase labelled, and collisionally activated HSP 16.9 32+ **A** and HSP 16.5 48+ **B**. The spectral intensity has been magnified 5-fold for m/z >6000. The most abundant charge state for each of the dissociation products is shown in the presence of ND_3 (blue) relative to the NH_3 control (black). The average deuterium content per subunit is shown for each of the dissociation products, for HSP 16.9 **C** and HSP 16.5 **D**. Symmetric partitioning of the intact uptake (dashed grey lines) is displayed for reference. The amount of D enrichment on the ejected monomer unit is shown for each observed charge state

To determine how deuterium from gas-phase HDX labelling partitioned on such asymmetric CID systems, HSP 16.9 and HSP 16.5 were labelled as described for other complexes following activation (gpHDX-CID), HSP 16.9 160 V and HSP 16.5 70 V. The most abundant charge state was highlighted, and the shift in m/z upon gas-phase labelling is shown in Figure 5.5A,B. The monomeric region provides high-quality data with little deviation between repeats, in part due to the instrument sensitivity over the range but also the stripped monomers are less prone to influence from salt adduct shifts while more variation in peak shape and the determined centroid were seen for the higher m/z species. Little peak deviation was seen

between experimental repeats often performed on different days for all dissociation products and intact oligomer, Fig. 5.6.

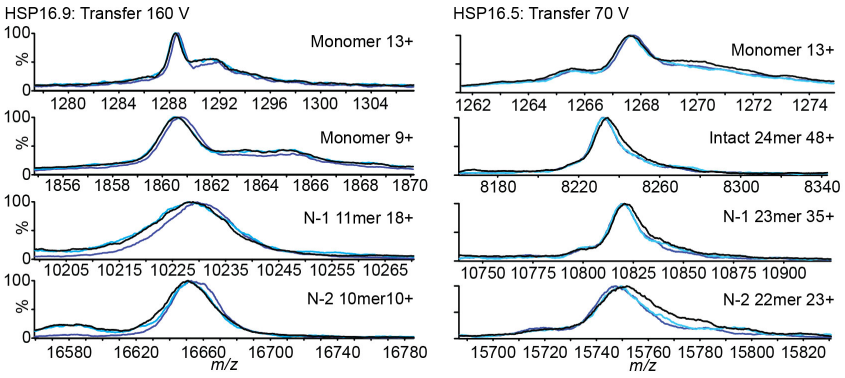


Figure 5.6. Reproducibility of gas-phase HDX-CID on sHSP oligomers. Example mass spectra for HSP 16.9 (left) and HSP 16.5 (right) showing the reproducibility of gas-phase labelled oligomers after CID. Three replicates for selected species and charges states are shown corresponding to the data in Figure 5.5

The uptake for HSP 16.9 and HSP 16.5 was determined for each dissociation product, shown in Figure 5.5C & D, and was averaged over all charges states across three repeats. The uptake for intact HSP 16.9 was 138 D or 11.5 D per subunit and was determined from the spectrum at a collision voltage (80 V) where the intact precursor remained with high signal-to-noise. This uptake value also represents the theoretical uptake of the dissociated monomers should the deuterium content of the complex partition perfectly symmetrically (i.e., the ejected subunit leaves with $^1/_N$ of the total deuterium uptake of the intact complex). The average deuterium per subunit for the N-2 oligomer, N-1 oligomer, and the monomeric species was found to be 9, 11.5 and 14 D/subunit (Fig. 5.5C). The deuterium content for symmetric partitioning of D is shown as a reference in dashed grey lines. The deuterium content was similarly determined for HSP 16.5 with the intact 24mer incorporating 12 D per subunit (288 D total), the monomers 15 D/subunit, N-1 oligomers 11 D/subunit, and the N-2 oligomers 10.5 D/subunit (Fig. 5.5D). The amount of deuterium on the monomeric species in each case was higher than if uptake were partitioned symmetrically, the difference between these two values was then defined as the monomeric enrichment. The enrichment is shown in Figure 5.5C & D for each monomer charge state and loosely increases with charge.

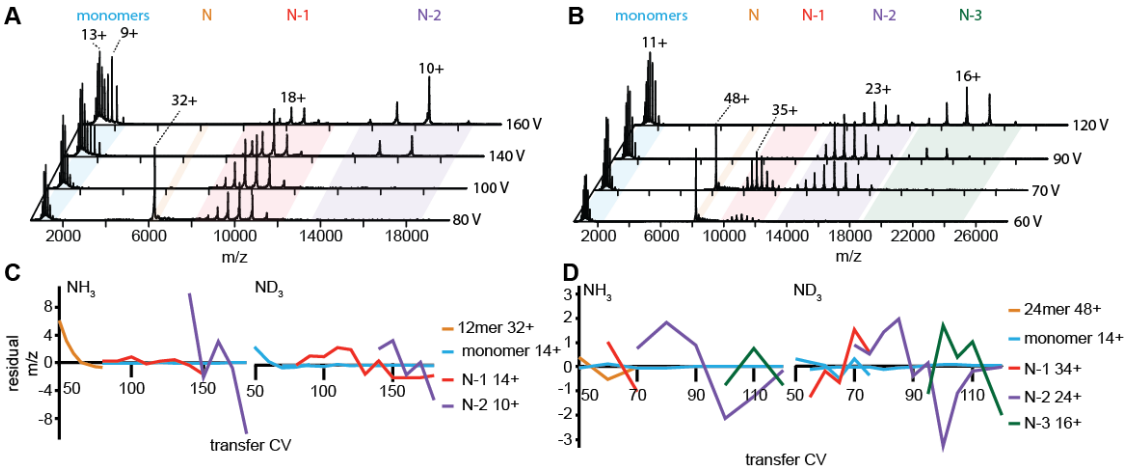


Figure 5.7. Collision energy dependency on *m/z* for gas-phase labelled sHSP oligomers. Gas-phase dissociation of HSP 16.9 and HSP 16.5 at increasing collision energies. The intact isolated complex was labelled with ND_3 gas in the trap TWIG with subsequent gas-phase dissociation in the transfer TWIG. Mass spectra are shown for **A** HSP 16.9 32+ 12mer (80 – 160 V CID) and **B** HSP 16.5 48+ 24mer (60 – 120 V CID). Sequential subunit loss occurs in both cases resulting in N-1, N-2 and, in the case of HSP 16.5, N-3 stripped oligomers in addition to the corresponding monomers. No systematic variation in *m/z* centroid was observed at increasing collision energy for either HSP 16.9 or HSP 16.5 under exposure to either NH_3 or ND_3. The deviation in peak centroid, relative to the value averaged across all collision voltages, is shown in both gases for **C** HSP 16.9 and **D** HSP 16.5. Monomers, show highly consistent behaviour across all collision voltages. (Oligomeric regions have been scaled by 5, >8000 *m/z*, for HSP 16.9 and scaled by 8 for collision voltages 70 – 120 V, >6000 *m/z*, for HSP 16.5)

As further validation that these differences are significant relative to any experimental or systematic error the collision energy dependence on peak position was determined for each species. The voltage was increased in the transfer as previously described between 60 - 160 V. The resulting spectra are shown in Figure 5.7A,B with the dissociated monomers and the stripped oligomers visible. The residual in *m/z* is provided for each species across the collision voltage range Figure 5.7C,D. This was calculated for three repeats obtained over many hours in both NH_3 and ND_3. Large variations can be seen for the stripped oligomers across the voltage range in a non-systematic manner for both un- and deuterated ammonia. The variation in the *m/z* residuals, the deviation in *m/z* from the average, for the monomers was near zero for all charge states and it is this attribute that makes using CID monomers a proxy for intact uptake highly appealing.

5.3.5 Deuterium enrichment of ejected subunits during CID of large complexes

For both HSP 16.9 and HSP 16.5, the dissociated monomers were ejected with more deuterium (D) than if each subunit left with a symmetric share. The result was therefore that the monomer became slightly enriched in deuterium and the stripped oligomers become slightly depleted. One likely explanation lies in the asymmetric redistribution of the net charge, with respect to mass. The charge per subunit was significantly lower for the intact complex compared to the stripped monomers. For example, the average charge localised to each subunit in the 48+ 24mer of HSP 16.5 was 2+, and the average charge state of dissociated monomers was 13+, representing an enrichment of charge of 11+ (Δz = 11). One likely method of charge enrichment during CID is the net migration of surface protons moving from the folded oligomer to the monomer, during unfolding and ejection of the monomer from the oligomer (Fig. 5.8A).[24] When the complex is labelled a proportion of the mobile charge will be D+. The fraction of deuterium from the pool of labile heteroatom-bound non-amide hydrogens on the surface of the intact complex will determine how much D enrichment occurs as charge migrates. If HSP 16.5 was fully saturated with deuterium, the D enrichment would simply be the number of charges that have migrated, Δz = 11. In this work, the surface was not fully exchanged as ND_3 pressures were kept low to prevent ammonia adducts from forming inside the spectrometer. The D enrichment is then the difference in charge, Δz, scaled by the fraction of deuterons from the pool of all heteroatom-bound non-amide hydrogens available (F_D), Eq. 5.1. This assumes that the rate of proton or deuteron site 'hopping' is comparable at these energies and that the flux of charge only occurs towards the ejecting subunit.

$$\begin{aligned} Enrichment_{monomer} &= \Delta z.F_D \\ \text{where, } \Delta z &= z_{monomer} - \frac{z_{intact}}{N} \end{aligned} \tag{5.1}$$

To estimate the degree of enrichment expected in these experiments several scenarios were considered to provide an upper and lower bound, Fig. 5.8B, Tables 5.1 and 5.2.

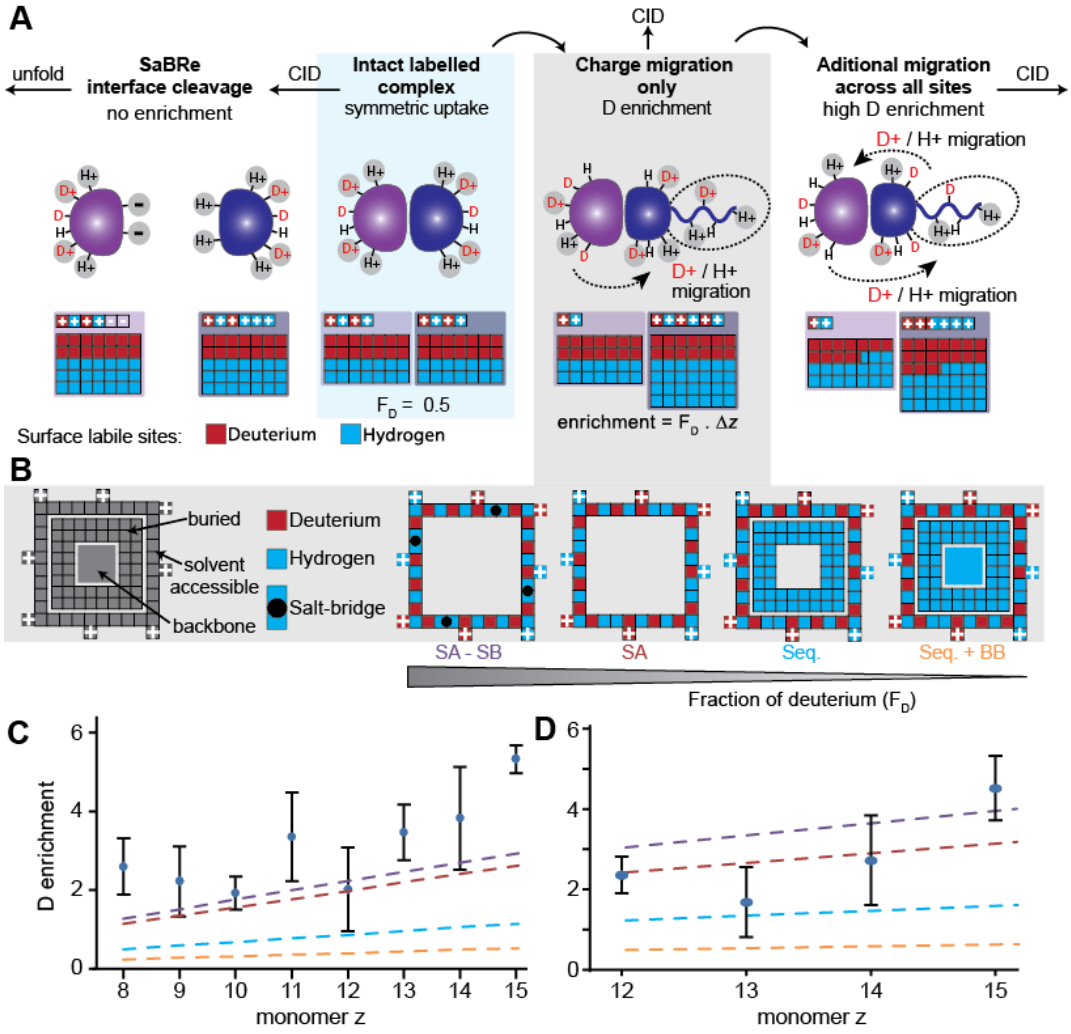


Figure 5.8. Migration of deuterium upon dissociation of protein complexes. **A** Schematic representing different dissociation mechanisms and how they may result in differing degrees of enrichment on the ejected subunit. A folded homo-oligomer will have symmetric deuterium uptake across each subunit, highlighted in light blue. Each site is represented by a block either occupied by hydrogen (blue) or deuterium (red), demonstrated for 50% deuterium saturation. Dissociation through salt-bridge rearrangement (SaBRe) or charge mediated unfolding will result in different enrichment levels, with the latter dependent on the fraction of deuterium in the system, panel **B**. Greater enrichment may occur if complete scrambling occurs across all sites, including newly exposed sites from unfolding. Charge is shown as white +/- and a net migration of 2+ is shown in each case. **B** A protein subunit is represented as a number of heteroatom-bound sites (squares) originating from different regions of the protein structure: surface accessible and buried non-amide hydrogens, and amide hydrogens associated with the back-bone. Upon activation a greater number of H/Ds may become mobile depending from where they originate. Illustrated are four cases; where only surface accessible side-chain H/Ds can migrate but not including hydrogens 'trapped' in salt-bridges (SA-SB), the case where only labile surface side-chain H/Ds can migrate (SA), where all labile Hs in the protein sequence are involved (seq.), and finally when all labile Hs from side-chains and the backbone amides can migrate (Seq.+BB). Progressing from SA-SB to Seq.+BB significantly decreases the fraction of D in the pool of migrating H/Ds as many exchange incompetent hydrogens are introduced to the system. The different models featured in **A** and **B** are shown for comparison with the observed monomer enrichment observed for HSP 16.9 **C** and HSP 16.5 **D** and discussed further in the text: Seq.+BB (orange), Seq. (blue), SA (red), and SA-SB (purple).

Summing all the heteroatom-bound hydrogens associated with the side-chains and backbone from the sequence provides an upper limit of the number of mobile Hs in the system, making F_D equal to 0.04 and 0.05 (Seq.+BB model), for HSP 16.9 and HSP 16.5, respectively (based on the measured uptake of 138 D and 288 D for the two intact proteins). If we then assume that amide hydrogens are not sufficiently mobile at the energies supplied for complex dissociation they can be removed from the calculation, the F_D increases to 0.09 and 0.12 (Seq. model). The observed enrichment lies above these lower enrichment limits (Seq. and Seq.+BB models, Fig. 5.8C & D) for both HSP 16.9 and HSP 16.5. This suggests that assuming all heteroatom-bound non-amide hydrogens, including those on the interior of the protein or at interfaces, are mobile is a poor approximation.

Table 5.1. Determination of the number of labile hydrogens expected for sHSP oligomers. Determination of the number of labile hydrogens expected for HSP16.9 and HSP16.5 based on the sequence (Seq.) and surface accessible (SA) residues from the crystal structures, 1GME and 1SHS respectively. The contribution from different residues are summarised for each model and the total determined for the intact complex including charge (z)

	HSP 16.9		HSP 16.5	
	number of labile Hs		number of labile Hs	
Residue	Seq.	SA	Seq.	SA
C	0	0	0	0
D	10	4	5	3
E	18	11	17	11
H	2	0	1	1
K	30	22	22	14
R	44	4	24	8
S	11	1	12	7
T	5	3	9	2
Y	0	0	2	0
termini	3	1	3	1
Subunit total	123	46	95	47
Complex total	1476	552	2280	1128
Complex total + z	1508	584	2328	1176

A more realistic model includes only labile Hs on the surface of the protein (surface accessible (SA) model); this was estimated for both proteins from the atomic coordinates of each protein and gives an F_D of 0.24 for both HSP 16.9 and HSP 16.5. The observed enrichment for HSP 16.5 broadly falls within both this surface-accessible model (SA) and the sequence model (Seq.) (Fig. 5.8D). The

Table 5.2. Approximation of the deuteron fraction of HSP 16.9 and HSP 16.5. The FD is based on the total number of labile hydrogens for different models discussed in the main text; sequence and the backbone (Seq.+BB), sequence (Seq.), surface accessible (SA) and surface accessible including salt-bridges (SA+SB).

Model	HSP16.9		HSP16.5	
	Labile Hs	F_D	Labile Hs	F_D
Seq.+BB	3320	0.04	5880	0.05
Seq.	1508	0.09	2328	0.12
SA	584	0.24	1176	0.24
SA+SB	548	0.25	936	0.30

deuterium enrichment levels for HSP 16.9 do not quite fall within these limits. The surface accessible model predicts an F_D of 0.24 while a linear fit of the observed data would suggest an F_D of 0.36 (Fig. 5.8C). While this is greater than expected, the enrichment remains well below the maximum allowed, F_D=1. This would suggest that fewer Hs are mobile than the crystal structure would predict, implying that a greater number of residues are either buried in the gas-phase or involved in strong interactions preventing migration away from that site.

Other factors that could affect H/D migration are differences in proton affinity of the sites, and solvation of sidechains with carbonyl groups of the backbone. The number of salt-bridges present in the crystal structure was determined and for each ion pairing interaction the proton count reduced accordingly, surface accessible minus salt-bridge (SA-SB) model. A total of 3 and 10 salt-bridge interactions were estimated for HSP 16.9 and HSP 16.5, respectively, giving F_D values of 0.25 and 0.30.

One other scenario to consider is that subunits do not become asymmetrically charged due to charge migration, but instead interfacial salt-bridges are asymmetrically broken in such a way that the ejected subunit leaves with more charge and the potential to subsequently unfold (SaBRe model).[25] For a labelled system, HDX at the interface of the complex would be unlikely, and so this mechanism would not provide any means of subunit deuterium enrichment, instead resulting in perfectly symmetric D partitioning (i.e. no monomeric enrichment) (Fig. 5.8A). This is inconsistent with the data for HSP 16.9 and HSP 16.5. However, it is conceivable

that both proton migration and the SaBRe mechanisms operate in tandem which would act to decrease the value of Δz in equation 5.1.

An upper limit to the amount of deuterium enrichment can be imagined if migration occurs in addition to the net charge movement (Fig. 5.8). In this scenario, complete scrambling occurs across all possible labile sites on the complex surface with H/D flux both towards and from the unfolded, yet attached, subunit. As the ejecting subunit unfolds it will reveal additional unlabelled, labile residues resulting in an excess of sites on the unfold-ed subunit. If complete equilibration of deuterium occurs over all sites, this will manifest in a deuterium enrichment of the ejected monomer. For HSP 16.9 the unfolded subunit contains 123 non-amide labile H sites, and the remaining oligomer has 473, therefore after complete H/D migration, the monomer will be left with 20.6% of the total uptake, rather than 8.3% for symmetric dissociation. This means the monomer would be enriched by 12.3% of the total uptake, 17 D. Our observed enrichment falls well below this level (not shown in Figure 5.8C & D) suggesting complete H/D migration is not occurring and/or the ejected subunit is not entirely unfolded.

Whilst the migration of surface charge, through the surface accessible and salt-bridge models, does not completely explain the observed data for HSP 16.9 the models do provide a means of explaining the large majority of enrichment with a relatively simple scenario. Other contributing factors may certainly be at play such as perturbation of structure in the gas phase and the scale and impact of salt-bridge and electrostatic interactions have on hindering H/D migration. Using a model where H/D flux is only towards the unfolding subunit appears valid as it accounts for most of the observed enrichment levels.

5.4 Conclusion

Here, we have shown that performing native MS and employing gas-phase HDX-MS in combination with CID can be used to provide uptake information on the structure of multimeric protein complexes. We analysed assemblies ranging from dimers to

24mers and showed that it was possible to label these by infusion of ND_3 into the trap T-wave ion guide of the mass spectrometer. Dissociation downstream of gpHDX allowed for the deuterium content to be determined at the subunit level and thus could be compared to that of subunits ionised from solution in an unbound state. This would both report on the exposed surface of free monomers and the intact complex, allowing inference of the protection from HDX provided by the interface.

Dimers of α-lactalbumin and cytochrome c were studied by gas-phase HDX-MS and dissociated at increasing CID energies. The deuterium content per subunit was comparable (± 2 D) for intact complex and dissociated monomers of various charge states. Increasing the activation energy for dissociation of the complexes, produced more dissociated monomers to intact dimer, but only affected the deuterium content to a smaller degree (<10%). This indicates that gas-phase HDX-MS of protein complexes can be performed at different activation energies for optimal dissociation of the complex. The dimers of α-lactalbumin and cytochrome c both additionally showed reduced gas-phase deuterium uptake compared to unbound monomers of the proteins. This indicated that the number of exchangeable heteroatom-bound non-amide hydrogens decreased upon dimerisation, possibly due to the formation of hydrogen bonding at the binding interface and changes in the higher-order conformation of the proteins.

The gas-phase HDX-MS and CID methodology was also used to study larger multimeric protein complexes. Tetramers of transthyretin were first studied, which dissociated asymmetrically in terms of charge and mass into monomer and trimers. The monomers contained a higher amount of charges per subunit (6+ – 9+) compared to trimers (2+ – 3+). Comparable deuterium content per subunit was observed for both intact precursor and monomers and trimers from dissociation. Dissociated monomers showed a small increase in deuterium content (0.6 D) relative to dissociated trimers. Additionally, dissociated monomers and trimers of higher charge states both showed a limited increase in deuterium content (0.4 – 1.6 D), compared to lower charged species. Due to the relatively small migration of charges

Δz=3.75, only a small intermolecular H/D migration was expected, which is in line with the <1 D enrichment of the dissociated monomers observed here. This indicated limited or localized intermolecular H/D migration upon asymmetric gas-phase dissociation of protein complexes, with increasing asymmetric charge migration leading to increasing H/D migration. This hypothesis was tested further on the larger 12mer HSP 16.9 and 24mer HSP 16.5 protein complexes which both dissociated very asymmetrically in terms of mass and charge. For dissociation of these large protein complexes, limited intermolecular H/D migration was clearly apparent. Dissociated monomers showed a relatively higher deuterium content than the intact complexes and the dissociated N-1 and N-2 complexes. The dissociated monomers additionally showed a small increase in deuterium content as a function of their charge state.

The generally accepted explanation for asymmetric dissociation by CID is the electrostatic charge partitioning model, where the single dissociated subunit unravels and undergoes proton migration prior to ejection from the residual complex.[24,26–28] The ejected subunit is therefore unfolded and bears a large percentage excess of charges, where the remaining subunits in the complex contain somewhat fewer charges. Based on the electrostatic charge partitioning model, four different models were proposed for quantifying the amount of potential inter-molecular migration of deuterium for dissociation of protein complexes that might be expected. i) The sequence model (Seq.+BB), ii) the sequence model not including backbone amide hydrogens (Seq.), iii) the surface accessible residue model (SA), and iv) the surface-accessible model that accounts for salt-bridge interactions (SA-SB). For HSP 16.5, the measured deuterium enrichment could best be explained by either the SA, SA-SB, and Seq. models. For HSP 16.9, the measured deuterium enrichment was higher than the proposed methods, but the SA and SA-SB models showed close correlations to the experimental values. Of the four proposed models, the SA and SA-SB models seemed to be most consistent with deuterium enrichment measured.

An alternative and more recent model for explaining the dissociation of protein complexes is the salt bridge rearrangement (SaBRe) model that proposed that

rearrangement and breaking of salt bridges at the protein complex interface are responsible for the asymmetric dissociation behaviour.[25] Gas-phase labelling would not occur on the binding interface of protein complexes, as hydrogen is shielded from HDX due to either hydrogen-bonding and/or being buried in the binding interface. No inter-molecular enrichment of deuterium is therefore expected according to the SaBRe model upon complex dissociation as any charged sites at the interface will therefore not likely contain any deuterium. A clear indication of intermolecular H/D migration was observed for the larger sHSP 12mer and 24mer protein complexes. From these data, it is therefore not likely that the SaBRe model alone is responsible for dissociation of non-covalent protein complexes. However, it cannot be excluded that salt bridge rearrangements play a part in the dissociation of protein complexes.

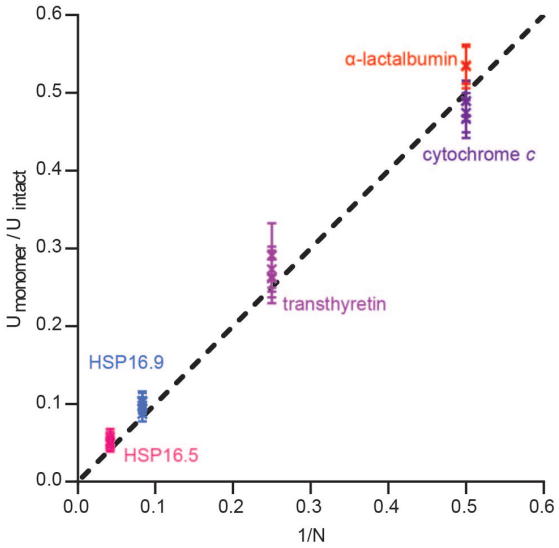


Figure 5.9. Deuterium levels partition symmetrically during CID of native protein complexes. Deuterium levels in native protein complexes following gas-phase HDX are partitioned in a highly symmetrical manner during CID. The fraction of monomer uptake ($U_{monomer}$) relative to intact complex uptake (U_{intact}) is plotted against the inverse of the number of subunits (N) in the intact complex. For perfectly symmetric D partitioning points would fall on the diagonal

The results presented here indicate that limited inter-molecular movement of H/D is occurring upon CID of non-covalent protein complexes, with the deviation from the symmetric partitioning of deuterium depending heavily on the charge asymmetry. The amount of migration, however, is far below the level expected if all labile sites were sampled by the mobile H/D pool over the course of a gas-phase

dissociation event. This is similar to previous studies on smaller peptide dimers labelled with deuterium at backbone amides sites in solution, which showed a smaller degree (5–10%) of intermolecular H/D migration upon CID.[29] In contrast, H/D levels at both backbone amides and side-chain sites are well known to migrate/scramble completely on an intra-molecular level for both peptides and proteins upon CID.[30–32]

Our results demonstrate that gas-phase HDX-MS with CID can be a powerful tool to provide structural information on protein-protein complexes and mechanistic insight into their dissociation. The method could be used to probe heterogeneous sample mixtures and proteins in a dynamic equilibrium, where solution HDX-MS and other biophysical techniques may not have adequate resolution or separation. Here we show that the deuterium content of the gas-phase dissociated monomer can be used as a proxy for the intact uptake, as only limited intermolecular H/D migration is occurring upon CID even in cases of very asymmetric charge partitioning (Fig. 5.9). This is of significant value for systems where the intact precursor cannot be cleanly isolated or carries significant adducts which make deuterium uptake determination difficult, and for which time-aligned experiments exploiting IMS is advantageous.[33] Further insight and development of the charge migration models used here could allow for a correction to be applied to account for intermolecular H/D migration upon complex dissociation. This could facilitate the further study of constituent parts of a heterogenous protein-protein complex within the mass spectrometer.

5.5 Methods

Materials

Deuterated ammonia gas, ND_3, was purchased from either AGA Gas (Pullach, Germany) in a lecture bottle containing 98% ND3 in a 10 L gas fitted with a Z14850 regulator (Sigma Aldrich, Schnelldorf, Germany) for experiments conducted at the University of Copenhagen, or a lecture bottle containing 10 L 99% deuterated ammonia gas purchased from CK Isotopes (Desford, UK) fitted with a CGA 180

CONCOA 422 series regulator (CK isotopes) for experiments performed in our laboratory. HSP 16.9 (*wheat*) was expressed and purified as described previously,[34] and HSP 16.5 (*Methanocaldococcus jannaschii*) was purified using the following method. In brief, the protein was expressed in BL21(DE3) cells for 3 hours at 37 °C. Cells were lysed using a microfluidiser in 20 mM tris (pH 8), 50 mM NaCl, 0.1 mM EDTA. Protein was first purified by anion exchange chromatography using two 5 ml Hi-Trap Q columns (GE Healthcare) using 20 mM tris (pH 8.0) for protein binding and eluted with a line-ar gradient to 1 M NaCl. Fractions were pooled and concentrated to which an equal volume of 2 M ammonium sulphate was added before filtration. The filtrate was then loaded onto a 5 ml hydrophobic interaction chromatography (HIC) phenyl sepharose column (GE Healthcare) in 20 mM tris, 50 mM NaCl, 1 M ammonium sulphate (pH 8.0) and eluted using a linear gradient in the absence of ammonium sulphate. Final purification and buffer exchange was performed using size exclusion chromatography (SEC) on a Superdex 200 Increase 10/300 column (GE Healthcare) with a 200 mM ammonium acetate solution (pH 8). All other samples and reagents were purchased from Sigma-Aldrich including lyophilised protein samples: Transthyretin/prealbumin (*human plasma*), α-lactalbumin (*bovine milk*), and cytochrome c (*equine heart*). The lyophilised proteins were unless stated otherwise, dissolved in LC-MS grade water and diluted into a suitable concentration in 200 mM ammonium acetate. Additional desalting steps were performed utilizing Micro Bio-Spin columns as per the manufacturer's instructions (Bio-Rad Laboratories, Hercules, CA, USA). Cytochrome c and α-lactalbumin dimer experiments were performed at 75 – 100 μM (dimer), and further diluted to 10 μM (monomer) for monomer control experiments. Transthyretin, HSP 16.9, and HSP 16.5 samples were prepared and analysed at 0.5 – 5 μM (oligomer).

Native mass spectrometry of protein-protein complexes

Protein samples were ionized in positive ionization mode from native-like conditions (200 mM ammonium acetate). Samples were prepared and loaded in borosilicate

capillary emitters (either: Proxeon/Thermo or made in-house), and ionized using a nanoflow Z-spray ESI source (Waters, Wilmslow, UK) mounted on either a Synapt G1 or a Synapt G2 HDMS mass spectrometer (Waters, Wilmslow, UK) operating in ToF mode, according to published protocols.[35] The source and spray settings were tuned to obtain the best possible signal-to-noise and HDX efficiency: capillary 1.4 kV, sampling cone 10 – 120 V (120 V for in-source CID experiments), extraction cone 3 V, accelerating potential into the trap collision cell 8 – 40 V, and source block temperature 20 °C.

Gas-phase HDX-MS and gas-phase dissociation of protein-protein complexes

Gas-phase HDX-MS was performed by leaking ND_3 gas into a trap travelling wave ion guide (TWIG) of the mass spectrometer, and thus leading to labelling of the ions as a function of the ND_3 pressure in the TWIG and the time the ions traversed the TWIG (governed to a large extent by the trap T-Wave velocity as demonstrated in 4). Gas pressures, T-Wave heights and T-Wave velocities of the TWIGs were adjusted for optimal ion transmission and labelling: trap ND_3 pressure 12 – 14 μbar, trap T-Wave height 7.0 V, and trap T-Wave velocity 150 m/s (for experiments with α-lactalbumin, cytochrome c, and transthyretin). Alternatively, trap ND_3 pressures of 22.2 μbar (HSP 16.9) or 42 μbar (HSP 16.5), with trap T-Wave height 0.1 – 0.2 V, and trap T-Wave velocity 200 m/s. By manually setting the gas regulator metering valve to obtain similar TWIG pressures of ND_3, with an average deviation in the m/z of deuterated species at ∼0.01%, between replicates. These minor differences in uptake probably reflect day-to-day thermal errors inherent to the Pirani gauge. We also note that all pressures reported in the mass spectrometer software are always "nitrogen-equivalent", and conversion to absolute pressure values for ND_3 would require a correction factor depending on the thermal conductivity of ND_3 versus N_2. The protein complexes were dissociated by post-source CID (referred to simply as CID) in the transfer TWIG to measure the labelling of individual components of the protein complex, at suitable activation energy for the complex. Control experiments, in which no labelling occurs, were performed with either Ar or NH_3.

For experiments with or without HDX, the gas pressure and T-Wave velocity setting are kept in the standard operating range for performing native mass spectrometry of protein complexes.[4] The T-Wave height setting is increased to ensure no ion "roll-over" is occurring in the T-Wave.[36,37] See Figure 4.1 on p. 122 for an overview of modifications performed on the mass spectrometer to enable gas-phase HDX-MS.

Data analysis

Processing of mass spectra was carried out using MassLynx V4.1 software (Waters, Wilmslow, UK) with a suitable mean filter smoothing and subsequent Gaussian fit using an in-house developed script using Excel 2013 (Microsoft, Redmond, WA, USA). The deuterium content of the deuterium-labelled proteins and protein-protein complexes was determined by subtracting the apex of the fitted Gaussian peak (corresponding to the m/z), with the non-deuterated control sample in the absence of deuteration (either NH_3 or Ar) and multiplied by the charge state of the peak.

Proton migration model estimations

To determine the fraction of deuteriums from the total number of labile sidechain heteroatom-bound hydrogens, the number of surface accessible labile side-chain sites had to be estimated. This was performed through the generation of the solvent accessible surface through PyMOL visualization software (Schrödinger, NY, USA). The X-ray crystal structures used for determination were 1GME and 1SHS, respectively for HSP 16.9 and HSP 16.5. The N-terminal residues in each oligomer are buried in the centre of the complex with only a small portion of the N-terminal structure visible. As such, structures were reduced to the alpha-crystallin domain and the C-terminal region, residues 44 and 43 onwards for HSP 16.9 and HSP 16.5 respectively. Spheres of 1 van der Waals radii were added to each of the side-chain hetero-atoms for the following residues: Lys, Arg, His, Tyr, Asp, Glu, Thr, Ser, and Cys. If the sphere was present at, or protruding from, the solvent-accessible surface the site for that residue was deemed exposed and included in the calculation. Salt bridges were also determined using the polar contact feature in the PyMOL software with a distance cut-off set at 3.6 Å. If a salt-bridge interaction was present

the labile hydrogen count was reduced by 1. It should be noted that the true shielding/restrictive effect of a salt-bridge may be greater than this and could result in up to 3 hydrogens being removed from the mobile pool. Our calculation does not account for protonated side-chains solvated by the backbone carbonyl group picoseconds after ionization.[7] Backbone carbonyl group solvation in the gas-phase does not appear to render these hydrogens exchange incompatible, as peptides have shown to exchange most heteroatom-bound non-amide hydrogens during gas-phase HDX-MS.[38]

5.6 Acknowledgements

I thank Ulrik Mistarz and Kasper Rand, University of Copenhagen, first and foremost whose efforts would not have brought this shared work to publication. For clarification of contributions, Ulrik Mistarz was funded by a COST short-term scientific mission to join our lab to begin this work. Most of the preliminary work surrounding the small protein-protein systems was acquired and interpreted together. Ulrik Mistarz then performed additional repeat experiments at his home institution from which the data on cytochrome C, α-lactoglobulin and transthyretin presented here were acquired and analysed, as such these systems are only discussed in brief for the sake of completeness. Figures 5.2, 5.3 and 5.4 are modified from the manuscript edition and were initially created by Ulrik Mistarz, all other figures are solely my work. This chapter is an adaptation of the co-authored publication Mistarz *et al.* 2018, I have focussed on my contribution in probing large complexes through gpHDX-CID and the development of the theoretical basis and discussion around subunit enrichment during CID. I thank Justin Benesch for his critical discussions surrounding the mechanisms of subunit enrichment. To Dale Shepherd for his historic provision of in-house expressed and purified HSP 16.9. Finally, to Jeff Brown for instrumental assistance and advice over the years.

References

[1] Konijnenberg, A.; Butterer, A.; Sobott, F. Native ion mobility-mass spectrometry and related methods in structural biology. *Biochimica et Biophysica Acta - Proteins and Proteomics* **2013**, *1834*, 1239–1256.

[2] Mehmood, S.; Allison, T. M.; Robinson, C. V. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annu. Rev. Phys. Chem.* **2015**, *66*, 453–474.

[3] Loo, J. A. Studying Noncovalent Protein Complexes by Electrospray Ionization Mass Spectrometry. *Mass Spectrometry Reviews* **1997**, *16*, 1–23.

[4] Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nature protocols* **2008**, *3*, 1139–52.

[5] Badman, E. R.; Hoaglund-Hyzer, C. S.; Clemmer, D. E. Monitoring Structural Changes of Proteins in an Ion Trap over ∼10-200 ms: Unfolding Transitions in Cytochrome c Ions. *Analytical Chemistry* **2001**, *73*, 6000–6007.

[6] Sobott, F.; Hernández, F. H.; McCammon, M. G.; Tito, M. a.; Robinson, C. V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Analytical Chemistry* **2002**, *74*, 1402–1407.

[7] Breuker, K.; McLafferty, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase, 10(-12) to 10(2) s. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18145–52.

[8] Green, M. K.; Lebrilla, C. B. Ion-molecule reactions as probes of gas-phase structures of peptides and proteins. *Mass spectrometry reviews* **1997**, *16*, 53–71.

[9] Chandler, S. A.; Benesch, J. L. Mass spectrometry beyond the native state. *Current Opinion in Chemical Biology* **2018**, *42*, 130–137.

[10] Bohrer, B. C.; Merenbloom, S. I.; Koeniger, S. L.; Hilderbrand, A. E.; Clemmer, D. E. Biomolecule Analysis by Ion Mobility Spectrometry. *Annual Review of Analytical Chemistry* **2008**, *1*, 293–327.

[11] Jurneczko, E.; Barran, P. E. How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *The Analyst* **2011**, *136*, 20–28.

[12] McLuckey, S. A.; Stephenson, J. L. Ion/ion chemistry of high-mass multiply charged ions. *Mass Spectrometry Reviews* **1998**, *17*, 369–407.

[13] Gross, D. S.; Schnier, P. D.; Rodriguez-Cruz, S. E.; Fagerquist, C. K.; Williams, E. R. Conformations and folding of lysozyme ions in vacuo. *Proceedings of the National Academy of Sciences* **1996**, *93*, 3143–3148.

[14] Cooper, H. J.; Hakansson, K.; Marshall, A. G. The role of electron capture dissociation in biomolecular analysis. *Mass Spectrometry Reviews* **2005**, *24*, 201–222.

[15] Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Analytical Chemistry* **2000**, *72*, 563–573.

[16] Mistarz, U. H.; Chandler, S. A.; Brown, J. M.; Benesch, J. L. P.; Rand, K. D. Probing the Dissociation of Protein Complexes by Means of Gas-Phase H/D Exchange Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2019**, *30*, 45–57.

[17] Jurchen, J. C.; Garcia, D. E.; Williams, E. R. Further studies on the origins of asymmetric charge partitioning in protein homodimers. *Journal of the American Society for Mass Spectrometry* **2004**, *15*, 1408–1415.

[18] Jurchen, J. C.; Williams, E. R. Origin of Asymmetric Charge Partitioning in the Dissociation of Gas-Phase Protein Homodimers. *Journal of the American Chemical Society* **2003**, *125*, 2817–2826.

[19] Hilton, G. R.; Lioe, H.; Stengel, F.; Baldwin, A. J.; Benesch, J. L. P. *Molecular Chaperones*; Springer Berlin Heidelberg, 2012; pp 69–98.

[20] Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chemistry & biology* **2006**, *13*, 597–605.

[21] Benesch, J. L. P.; Ruotolo, B. T.; Sobott, F.; Wildgoose, J.; Gilbert, A.; Bateman, R.; Robinson, C. V. Quadrupole-Time-of-Flight Mass Spectrometer Modified for Higher-Energy Dissociation Reduces Protein Assemblies to Peptide Fragments. *Analytical Chemistry* **2009**, *81*, 1270–1274.

[22] Van Montfort, R. L. M.; Basha, E.; Friedrich, K. L.; Slingsby, C.; Vierling, E. Crystal structure and assembly of a eukaryotic small heat shock protein. *Nature Structural Biology* **2001**, *8*, 1025–1030.

[23] Kim, K. K.; Kim, R.; Kim, S. H. Crystal structure of a small heat-shock protein. *Nature* **1998**, *394*, 595–9.

[24] Popa, V.; Trecroce, D. A.; McAllister, R. G.; Konermann, L. Collision-Induced Dissociation of Electrosprayed Protein Complexes: An All-Atom Molecular Dynamics Model with Mobile Protons. *Journal of Physical Chemistry B* **2016**, *120*, 5114–5124.

[25] Loo, R. R.; Loo, J. A. Salt Bridge Rearrangement (SaBRe) Explains the Dissociation Behavior of Noncovalent Complexes. *Journal of the American Society for Mass Spectrometry* **2016**, *27*, 975–990.

[26] Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. *Journal of the American Chemical Society* **1994**, *116*, 5271–5278.

[27] Felitsyn, N.; Kitova, E. N.; Klassen, J. S. Thermal Decomposition of a Gaseous Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behavior. *Analytical Chemistry* **2001**, *73*, 4647–4661.

[28] Sinelnikov, I.; Kitova, E. N.; Klassen, J. S. Influence of Coulombic Repulsion on the Dissociation Pathways and Energetics of Multiprotein Complexes in the Gas Phase. *Journal of the American Society for Mass Spectrometry* **2007**, *18*, 617–631.

[29] Modzel, M.; Stefanowicz, P.; Szewczuk, Z. Hydrogen scrambling in non-covalent complexes of peptides. *Rapid Communications in Mass Spectrometry* **2012**, *26*, 2739–2744.

[30] Demmers, J. A. A.; Rijkers, D. T. S.; Haverkamp, J.; Killian, J. A.; Heck, A. J. R. Factors Affecting Gas-Phase Deuterium Scrambling in Peptide Ions and Their Implications for Protein Structure Determination. *Journal of the American Chemical Society* **2002**, *124*, 11191–11198.

[31] Jørgensen, T. J. D.; Gårdsvoll, H.; Ploug, M.; Roepstorff, P. Intramolecular migration of amide hydrogens in protonated peptides upon collisional activation. *Journal of the American Chemical Society* **2005**, *127*, 2785–2793.

[32] Mistarz, U. H.; Bellina, B.; Jensen, P. F.; Brown, J. M.; Barran, P. E.; Rand, K. D. UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Analytical Chemistry* **2018**, *90*, 1077–1080.

[33] Shepherd, D. A.; Marty, M. T.; Giles, K.; Baldwin, A. J.; Benesch, J. L. Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins. *International Journal of Mass Spectrometry* **2014**,

[34] Lee, G. J.; Pokala, N.; Vierling, E. Structure and in Vitro Molecular Chaperone Activity of Cytosolic Small Heat Shock Proteins from Pea. *Journal of Biological Chemistry* **1995**, *270*, 10432–10438.

[35] Kondrat, F. D. L.; Struwe, W. B.; Benesch, J. L. P. *Native Mass Spectrometry: Towards High-Throughput Structural Proteomics*; 2015; pp 349–371.

[36] Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid communications in mass spectrometry : RCM* **2004**, *18*, 2401–14.

[37] Mistarz, U. H.; Rand, K. D. Installation, validation, and application examples of two instrumental setups for gas-phase HDX-MS analysis of peptides and proteins. *Methods* **2018**, *144*, 113–124.

[38] Rand, K. D.; Pringle, S. D.; Murphy, J. P.; Fadgen, K. E.; Brown, J.; Engen, J. R. Gas-phase hydrogen/deuterium exchange in a traveling wave ion guide for the examination of protein conformations. *Analytical chemistry* **2009**, *81*, 10019–28.

6

Probing the protein unfolding pathway with gas-phase hydrogen/deuterium exchange MS

6.1 Abstract

Understanding the states proteins populate away from their native form is important for understanding protein folding and it's implication for disease. Little is known to what extent the (un)folding funnel of a protein in the gas-phase and native solution-state overlap. Native proteins undergo structural transitions on transferral to the gas-phase and additional manipulation or activation can cause further exploration of the protein's free energy landscape. Directly monitoring structural transitions as a protein evolves in the gas phase is difficult. To this end, this chapter describes our development of a combined ion mobility-gpHDX approach to selectively probe protein conformers at a global and side-chain level.

We use the gpHDX methodology described in Chapter 4 to probe the collision induced unfolding (CIU) of a series of GFP mutants. Sequential isolation, mobility separation and labelling allowed us to determine not only the collision cross section (CCS), but also the conformer specific uptake as a function of energy. Our results suggest that as the protein CCS increases due to activation, the labile sites associated with ND_3 exchange begin to self-solvate and restrict exchange. We propose the systems studied here undergo a gas-phase inversion as they become activated, acting to bury electro-negative and -positive sites. Ion mobility and gpHDX were shown to be sensitive to different aspects of protein ion structure and provide new insights into the gas-phase unfolding profile of proteins when coupled.

6.2 Introduction

Native mass spectrometry (MS) in its simplest and purist form can accurately inform on the mass of a complex and its constituent parts through gas-phase manipulations.[1–3] Introduction of intact protein complexes into the gas phase allows one to manipulate ions with high fidelity, providing unrivalled selectivity over the species one wishes to probe. Having this level of control has led to the development of native MS for understanding three-dimensional structure and assembly.[4,5] The most common approach to obtaining structural insight comes through the incorporation of ion mobility (IM) spectrometry, §2.1.6. This experiment reports an arrival time distribution, a measure that is sensitive to the overall shape of a protein, which can be used to calculate a collision cross-section (CCS) and may therefore distinguish between alternative three-dimensional models.[6] Manipulation through activating collisions with the buffer gas results in the slow heating of protein ions, triggering structural perturbations and ultimately, in the case of a complex, the ejection of a subunit or bound ligand. Understanding the structural changes at a local level has proven difficult to probe experimentally and yet crucial to understand how activation affects protein structure.[7] Performing IM on proteins as a function of activation within the mass spectrometer reveals the global structural transitions they undergo, from collapse to unfolding,[8,9] often termed collision induced unfolding (CIU). The number of transitions can correlate well with the number of domains in the native protein,[10] but their structural assignment remains difficult[11] with little general applicability. Furthermore, IM is typically only sensitive to changes if they affect the CCS of the overall protein, and do so by more than a few percent.[12] Such unfolding profiles measured by IM have found significant utility in recent years in assessing the stabilising effect of bound ligands and have potential in a more applied pharmaceutical context.[9,13] As a consequence, there is a considerable need to develop a means to probe structure on a more local level, to interrogate the properties and changes that occur on the surface of the protein.

A potential means to address this need could be by performing gas-phase HDX (gpHDX). This approach involves introducing a deuterated vapour into the mass spectrometer to exchange the isotopic label with labile side-chain hydrogens on the accessible surface of the protein ion as described in the previous two chapters. Previous studies have exploited isotopic exchange to probe limited aspects of gas-phase protein structure[14] and have recently been applied to observing complex interactions.[15] An important early study utilised an IM separation and simultaneous gpHDX to try and understand the efficacy of the technique for probing conformations of denatured proteins. These experiments were hampered by the concurrent nature of the separation and labelling approach and as such cannot decouple the impact CCS has on uptake.[16]

In this chapter we describe the use of gpHDX to interrogate the conformational landscape of native proteins inside the mass spectrometer. Having previously demonstrated in Chapter 4 the sensitivity to surfaces of different chemical environments using model GFP systems, we further probe these proteins as they evolve in the gas phase upon activation. We discover that the variation in uptake as a function of activation is extremely information rich, revealing changes not reported by the CCS alone. By performing sequential IM-gpHDX-MS, we are able to extract CCS-selected uptake profiles, allowing us to assign structural transitions of proteins that are indistinguishable by IM alone, and represent a dramatically improved specificity to the unfolding profile. Furthermore, changes observed through gpHDX can be ascribed to the self-solvation of exchange-competent residues coupled to the inversion and expansion of the structure.

6.3 Results and Discussion

6.3.1 Gas-phase HDX reveals structural changes during activation

Having demonstrated that gas-phase labelling is sensitive to the surface chemistry and composition of a native protein system we set out to use the method to probe how a protein surface evolves upon activating collisions. The GFP mutants,

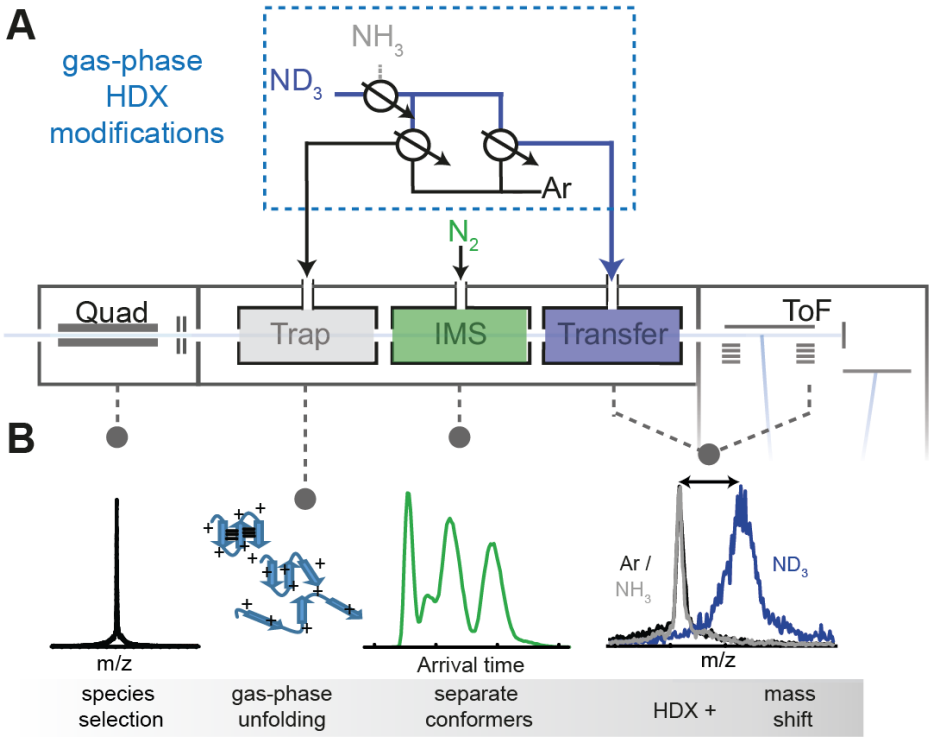


Figure 6.1. Gas-phase HDX modifications to allow conformer specific labelling. A Schematic depicting the gas-modifications to a Waters Synapt G1 ion mobility mass spectrometer. Deuterated ammonia (ND_3), or undeuterated (NH_3) for control experiments, is leaked into the transfer ion guide for labelling downstream to a mobility separation. **B** CIU-gpHDX work flow. Quadrupole isolation allows for high sample specificity including elimination of adduct species. Gas-phase collisional activation in the trap ion guide causes protein unfolding which can be subsequently separated and the arrival time distribution of conformers determined. The transfer ion guide is infused with deuterated ammonia vapour to exchange with the conformer surface and the mass-to-charge ratio subsequently determined.

introduced in Chapter 4, were exposed to activating collisions in the trap ion guide, after which the arrival time, a measure of size, was determined for the resulting species through travelling wave (TW) IM, §2.1.6. Introduction of ND_3 labelling gas post-IM, in the transfer cell optics, then allowed the generated conformers to exchange before MS analysis, Fig. 6.1.

We first enforced a strict quadrupole selection of a single charge state so that any shifts in m/z can be fully attributed to the labelling reaction. At higher collision energies salt adducts may be lost, thus reducing the perceived amount of uptake; it was therefore crucial in these studies to be stringent at the early isolation stage. This is demonstrated for acidic GFP where a full range scan is shown in Fig. 6.2A followed by the narrow isolation of the 10+ charge state in B. The $[\mathrm{M} + 10\,\mathrm{H}]^{10+}$ is

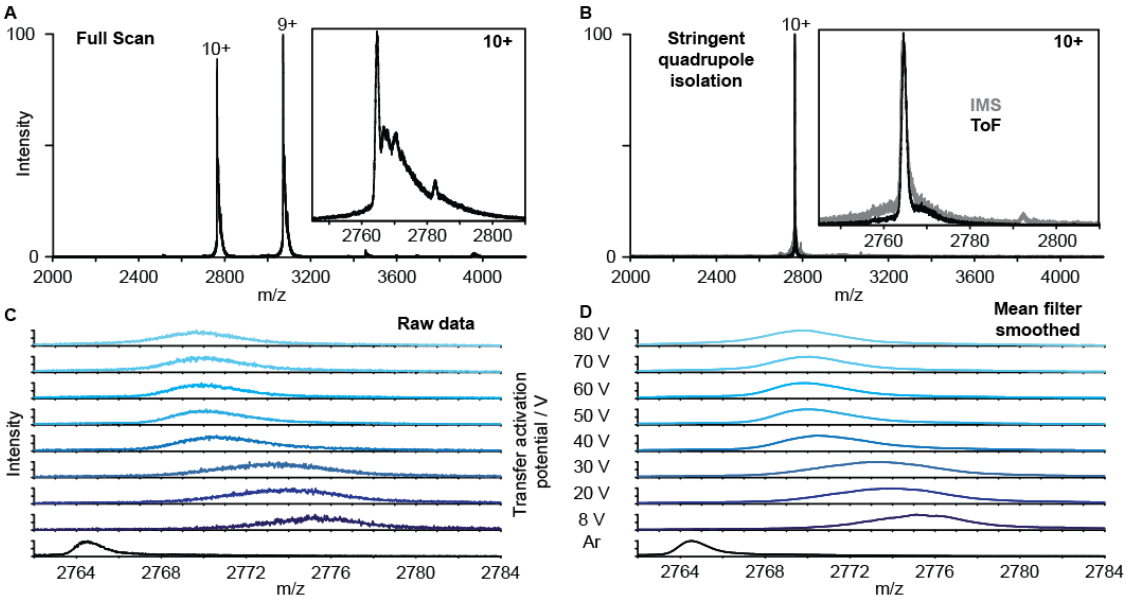


Figure 6.2. Obtaining CIU-gpHDX native mass spectra. A narrow quadrupole window is applied to give a narrow peak corresponding to only the protonated 10+ ion of acidic GFP, **A** & **B**. Ions are activated in the trap collision cell, pass through the mobility cell and are labelled in the transfer region. Labelling manifests as a shift in *m/z* (shades of blue) relative to the Ar control (black), **C**. Incorporation is seen to decrease as the trap collision potential is increased, 8 - 80 V, as indicated. Spectra are then smoothed using a mean filter of 20 × 1, with no shift in peak centroid observed due to processing, **D**.

almost exclusively transmitted when operating in both ToF (black) and IM (grey) modes.

Ions were activated into the trap cell at potentials between 8 and 80 V, the latter being the point at which fragmentation begins to occur. Ions then passed through the IM cell at a pressure of 0.57 mbar where the effective temperature of the ion becomes thermalised due to ∼100,000 gentle collisions with the nitrogen bath gas.[i] ND_3 was introduced at a partial pressure of 12 μbar[ii] to the transfer region for labelling to occur on the previously activated species. The injection voltage into the transfer cell is kept deliberately low, 8 V, as to not induce further conformational changes or increase the energy of the system. Figure 6.2C shows the raw spectra of the activated then labelled 10+ acidic GFP over the trap voltage range indicated.

[i]The number of collisions with nitrogen in a 18.2 cm long drift cell, 57 mbar at room temperature, is calculated for the 10+ ion with a diameter of 55 Å.

[ii]The final pressure in both the trap and transfer cells are around 30 μbar due to nitrogen leakage from the high pressure mobility cell. Thus, the relative ion velocity will be slower when compared to ToF mode.

Uptake was particularly large when ion activation was low, ~110 D at 8 V. As the activation was increased the uptake began to decrease and was accompanied by a small narrowing of the peak width. The data smoothing applied throughout these studies is shown in Figure 6.2D for clarification that processing has little impact on the centroid position or width.

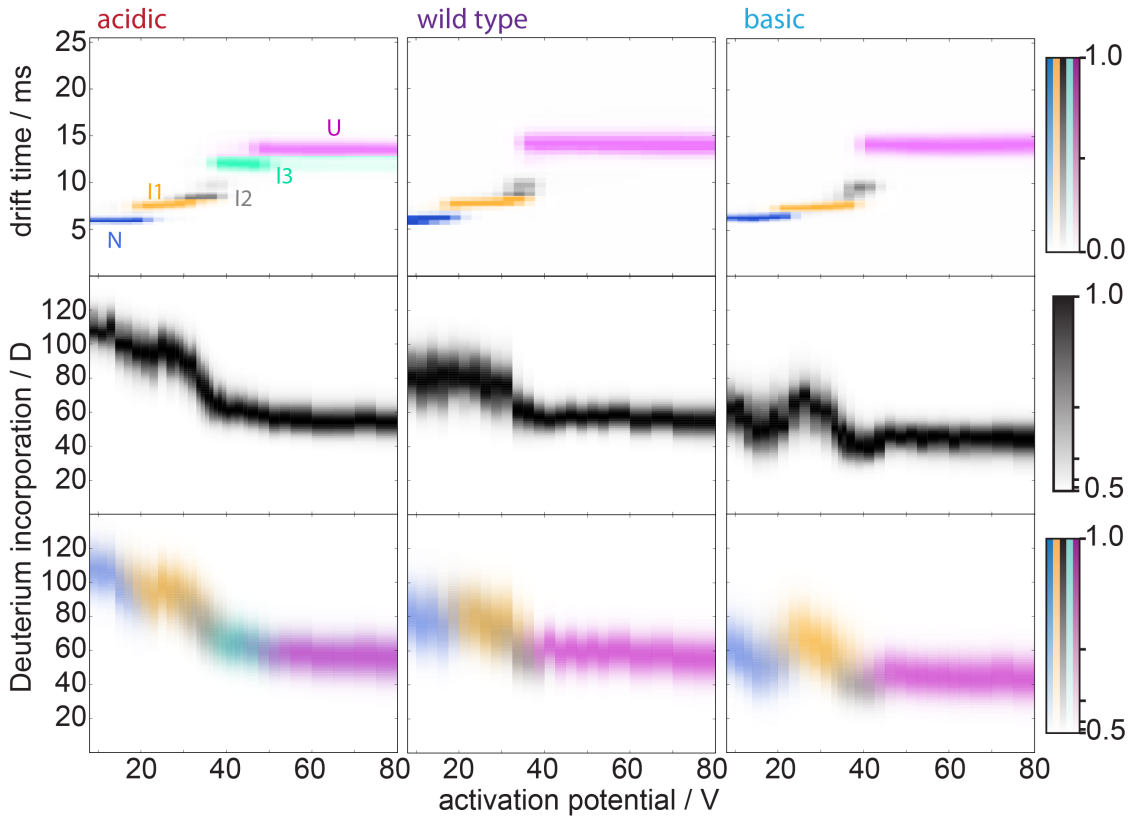


Figure 6.3. Gas-phase HDX and ion mobility report on conformer structural changes. **Top**, Collision induced unfolding plot using TWIMS to determine the drift time of the gas-phase conformers for the 10+ ion of acidic, WT and basic GFP variants. **Middle**, the total gpHDX uptake of all species present across the activation range. **Lower**, The drift-time selected uptake across the unfolding profile. Conformer extracted uptake distributions are shown in the colour corresponding to the conformer states in the upper panels. The coloured labels, N, I1, I2, I3 and U, correspond to the different conformer states identified (native, intermediates and unfolded).

To understand how the structural transitions were coupled to activation we refer to the IM dimension which was simultaneously acquired. Figure 6.3(top row) shows how ion size evolves as a function of energy for the WT and two GFP mutants.

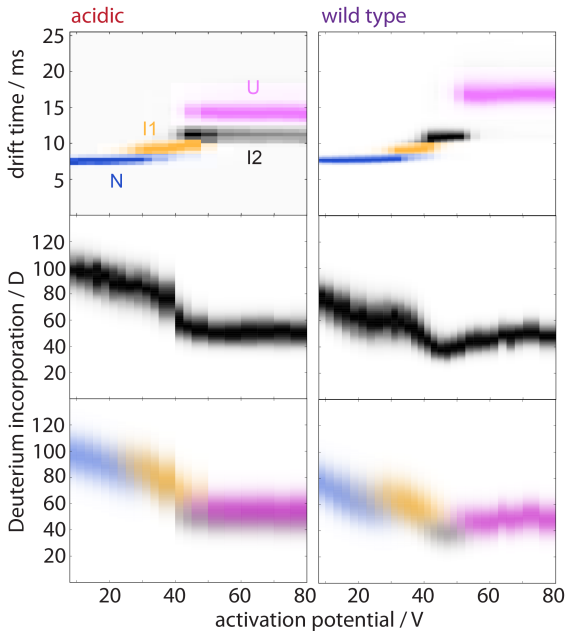


Figure 6.4. Collision-induced unfolding and labelling of GFP 9+ ions. Top, Collision induced unfolding plot using TWIMS to determine the drift time of the gas-phase conformers for the 9+ ion of acidic and WT GFP variants. **Middle**, the total gpHDX uptake of all species present across the activation range. **Lower**, The drift time selected uptake across the unfolding profile. Conformer extracted uptake distributions are shown in the colour corresponding to the conformer states in the upper panels. The coloured labels, N, I1, I2, I3 and U, correspond to the different conformer states identified (native, intermediates and unfolded).

The drift time, determined using a TW-IM instrument, is plotted against the prior collision voltage in the trap cell. Each GFP variant has a conserved native drift time of ∼6 ms. Upon activation, the variants begin to emerge at longer times from the drift cell due to, it is assumed, more extended structures.[10,17] All GFP variants display discrete gas-phase intermediates, indicated by the various colours, that have well-defined activation induced transitions. Little variation in arrival time was observed for a given conformer over its activation range. All 10+ ions of the GFP variants display comparable CIU profiles and possess similar transition onsets and drift times (N = 5.9 - 6.3 ms, I1 = 7.5 - 8.0 ms, I2 = 9.4 - 9.8 ms, U = 13.6 - 14.3 ms, [I3 = 12 ms]). The wider arrival time range seen for the most unfolded state is due to the acidic species which has a drift time of 13.6 ms, while the WT and basic fall within the much narrower range of 14.2 - 14.3 ms. The shorter arrival time of the acidic species may suggest this unfolded state is more structurally dissimilar to the other mutants than the other intermediate states on the CIU

pathway. The acidic mutant also has an additional conformeric state, I3, which coexists with the most extended form common to all mutants. The coexistance of two conformers at higher energies is also seen for the 9+ charge state of the acidic variant, Fig. 6.4(top row) where I2 is more pronounced and coexists with the most extended form across the entire activation range. Despite significantly different surface topologies (§4.3.6) all variants unfold in a broadly conserved manner, with the exception of the acidic variant at its most unfolded states. These subtleties in the unfolding profiles suggest that the surface composition is important in the structural landscape during activation.

We then considered how the label incorporation varied as the protein underwent these structural changes. Spectra akin to that shown in Figure 6.2D, were stacked and plotted as a two-dimensional histogram with uptake against the activation potential, Fig.6.3(middle). The relative transparency of the bars represent the intensity in the uptake (or m/z) dimension. Intensity is plotted non-linearly, as described in the methods, to emphasise the centroid peak shifts associated with the broad spectra. In the x-axis, data is binned according to the collision voltage interval, acquired every 1 - 5 V with smaller steps close to an IM transition. Spectra were processed and data plotted in a semi-automated manner using an external python script we developed, feeding, in part, from outputs from PULSAR software.[18] Further details are provided in the the Methods section, p. 201. These gpHDX unfolding histograms provide an ideal means of visualising and comparing the uptake profile to that of the conventional ion mobility unfolding plot.

The uptake for the initial low energy structure follows the same trend between GFP variants as described in Chapter 4, Fig.4.12 on p. 138, with uptake values of 110 Da, 80 Da and 60 Da for the 10+ charge states of the acidic, WT, and basic GFP variants respectively. The experimental difference here being the additional IM trapping procedure prior to labelling rather than a continuous ion beam. Surprisingly, a significant reduction in the proteins' deuterium incorporation is seen upon activation despite IM confirming the protein is expanding. The magnitude observed and the profile is highly dependent on the variant studied and the charge

state. (Compare with the 9+ in Figure 6.4). The decrease in deuterium is significant across all species but the magnitude of change is greatest for the acidic system. The acidic mutant sees a decrease of 55 D from the lowest to highest voltages employed, whilst the basic variant has a decrease of 15 D, with a greater number of features within the uptake profile.

The decrease in uptake for a more expanded structure is additionally unusual given that more extended states, by their very definition, have a greater CCS. The larger CCS should result in more reaction prone collisions and yet this does not manifest as an increase in uptake but quite the opposite. As a result, if we were to normalise the uptake per collision one would see an enhanced negative difference between the native and unfolded states.

6.3.2 gpHDX is sensitive to surface side chain structure not observed in IM

To uncover how each conformer observed in the IM dimension contributes to the overall mass change profile the difference in m/z, and ultimately uptake, was extracted across a slice in drift time to give the conformer specific deuterium incorporation. The conformer extraction analysis was also semi-automated using a modified plug-in to PULSAR and external script. Conformer specific uptake can be seen in Fig.6.3 & 6.4(lower) for the 10+ and 9+ charge states of GFP variants, respectively, where the colour corresponds to the conformer observed in IM, Figs. 6.3 & 6.4(top). It becomes clear how features in the uptake unfolding plots correspond to discrete IM transitions but also that the trace is far from step-wise but rather a continuum between transitions is observed. This is in stark contrast to IM, with the implication being that structural changes occur that restrict exchange but do not manifest as a detectable change in CCS. Given the nature of gpHDX and its localisation to the surface side-chains it may be that we are sensitive to local surface perturbations and interactions not detectable through a mobility measurement.

To gain a better understanding about the relationship between the CIU and HDX profiles we extracted the step-changes in each profile, in Da or ms, and

the gradient, a measure of state variability, for each of the conformers over their activation range, Fig. 6.5A. If we first consider the change at a transition by plotting the change in uptake against the change in drift time for all charge states and mutants, Fig. 6.5B. The drift time is plotted as a corrected drift time scaled by the charge squared, z^2, to account for the charge state dependencies in a TW-IM separation, see p. 41.[19] The trivial observation is that IM only ever observes an increase in drift time during unfolding irrespective of GFP variant or charge state, examples do exist however in the literature of gas-phase compaction.[20,21] Drift time differences are as great as +5 ms which is large relative to the mobility peak resolution. However, the change in uptake upon conformer transition is both positive and negative, falling in both the upper and lower quadrants, with neither dominating. Certain transitions across the 9/10+ variants tend to cluster depending on the transition identity. These are highlighted in Figure 6.5B with the initial unfolding step (N-I1) resulting in a positive uptake difference between 0 - 11 Da, the second transition (I1-I2) often displaying a large decrease, 9 - 15 Da, and the final unfolding step (I2-U) largely increasing once again, 7 - 10 Da. The fact that these uptake differences form loose groups depending on the transition may help describe the structural events occurring.

Secondly, the gradient for each conformer was plotted for each dimension, where the gradient is the change in drift time or deuteration over the voltage range (Δdt/ΔV or ΔD/ ΔV). It can be seen from Figure 6.5C, also visually in Figures 6.3 & 6.4, that the gradient for each conformer in mobility space is near zero. This shows that each conformer has an unchanging size within the resolution limits of our IM spectrometer. The axes limits were generously chosen to be $\sim$2 % of the total drift time range which would act to highlight any changes. The gpHDX conformer changes suggest quite a different story however. The gradients plotted, for all variants and conformers, are almost exclusively negative but span a large range from –1.7 - +0.1 D/V. If we extract these values according to which conformer they belong then we find it is predominantly the native state which covers a large gradient range. As the state becomes more unfolded, the distribution of gradients

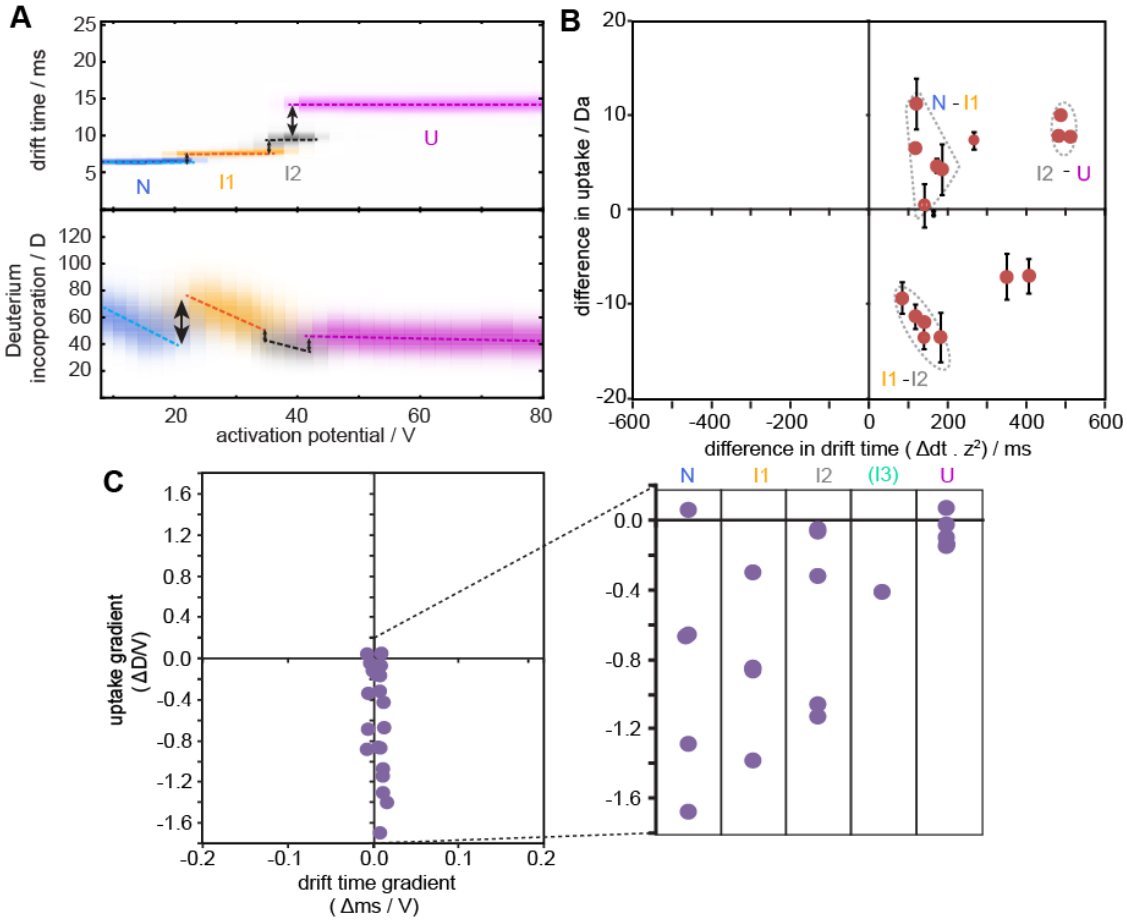


Figure 6.5. Step and gradient extractions from IM and gpHDX dimensions. A Extraction of the step changes and gradients for each conformer, demonstrated for the CIU-gpHDX basic GFP 10+. **B** A plot of the difference in uptake upon conformational transition for all mutants, transitions and charge states, against the corrected difference in drift time. Scaled by the square of charge, z^2, to account for the faster migration of higher charge states in a TWIM device. **C** Comparison of the gradients of each conformeric state, with the extracted uptake gradient shown for each conformer as a highlighted panel.

narrows until it centres near zero for the most extended conformer. The gradient reports on the amount of change that can occur for a particular conformer as the activation increases. The IM dimension tells us little is occurring at a global level until the next step change during the transition to a new conformer. The gpHDX dimension, however, tells us that exchange for a particular conformer often becomes increasingly reduced over the states activation range. As gpHDX ultimately reports on side-chain reactivity we are now informing on many local structural or electrostatic perturbations that additively reduce the surface's propensity for exchange, that is until the next step transition which triggers the process again for

the new conformeric state.

6.3.3 Validation of CIU-gpHDX behaviour

Replicates and inter-lab study

To validate this behaviour, the GFP variants were subjected to multiple repeats with the acquisition trap voltage randomly scrambled to avoid any systematic error, such as gas pressure deviations, over the course of the ramp. The three repeats for acidic GFP 9+ charge state are shown in Figure 6.6 as 2D histograms, processed in the same manner as previously described. Little difference was observed between the repeats in either the unfolding profile when uptake is conformer selected or otherwise.

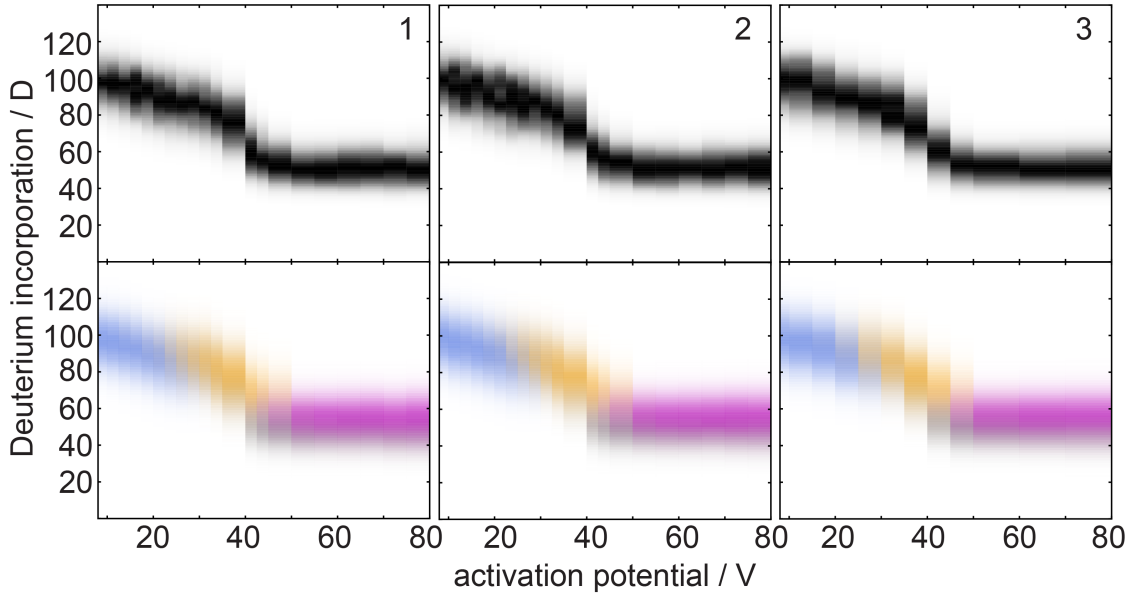


Figure 6.6. Reproducibility of CIU-gpHDX. CIU-gpHDX unfolding plots of acidic GFP 9+ charge state performed in triplicate showing high reproducibility. Colours correspond to the different conformer states identified in Figure 6.4. Collision voltages were ramped at a maximum interval of 5 V over 8 - 80 V.

To offer further support, experiments were replicated during an inter-lab study with the Rand laboratory at the University of Copenhagen. Gas-phase HDX experiments were performed as closely to our methods as practically possible. The 9+ WT GFP was subjected to gas-phase labelling in an analogous set-up to our own laboratory with activation in the trap and subsequent labelling in the transfer.

Some experimental differences do exist due to the difference in instrumentation used. Our experiments were conducted on a Waters Synapt HDMS generation 1 instrument whilst the gpHDX modifications in the replicate laboratory were on a 2nd generation spectrometer. Whilst the core architecture remains the same, differences do exist between the two instruments[22] in the pressures employed and the voltages required. The G2 instrument typically operates at a higher ion mobility pressure which will also influence the labelling and activation cells. In addition the waveform applied to the travelling waves differs between the generations, the G1 operates with a voltage applied to electrode pairs repeated every 6 electrodes whilst the G2 uses four lenses to generate the wave amplitude and repeats every 4 electrodes. These subtle difference will result in a different wave motion and levels of activation during the brief exposure to the high field wave front. Other, differences include the additional helium cell post-trap immediately before the IM region and greater bias applied to inject ions into the higher pressure cell. The final most notable difference between the experiments was the ND_3 pressure used, with 12 μbar used in our laboratory and 75 μbar when we performed the experiments in the Rand laboratory. The higher pressure was due to the lack of a restriction capillary and needle valve on the G2 instrument. Despite the high pressure no adducts were found when controls were performed with NH_3.

Differences do exist between the two CIU-gpHDX profiles, Fig. 6.7, but encouragingly the general form is consistent. An additional conformer was discernible on the higher IM resolution instrument which is highlighted in green. Importantly the large decrease in uptake seen at around 30 V is present in both replicates and corresponds to the loss of I1 and the formation of I2 and/or I3. The fact that the final uptake is comparable, ∼40 D, in both cases should not be over interpreted as the ND_3 pressure is significantly different. Like-wise that the native state does not have the same initial uptake is not of concern. What is unusual however is that the 2nd generation data do not display a large negative gradient for the native state but remains relatively flat. This could be due to a difference in early-stage activation between the instruments resulting in subtly different gas-phase structures

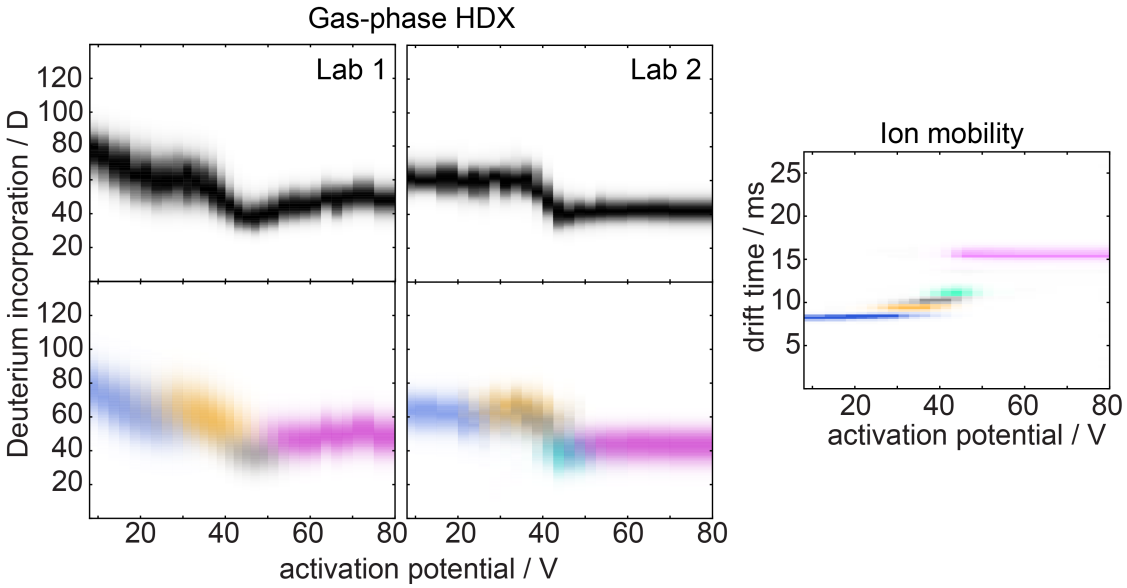


Figure 6.7. CIU-gpHDX unfolding plots resulting from an inter-laboratory study. Wild-type GFP 9+ ions were studied using the CIU-gpHDX methodology on two analogous set-ups in our laboratory (Lab 1) and the Rand laboratory, Copenhagen (Lab 2), **left**. Experiments were performed on a Waters Synapt G1 and G2 spectrometers as closely as the instruments would allow. Experimental differences are discussed in the text. Ion mobility wave and gas parameters were identical and the ion mobility unfolding plot is shown (**right**). An additional conformer state (green) was resolvable on the higher IM resolution G2 instrument. Colours indicated in the lower spectra correspond to the conformers seen in the IM plot with uptake distributions extracted over the respective drift time space.

particularly at a local level. The native state is likely to be most sensitive to small differences in early activation.

Activation in the absence of ion mobility

CIU-gpHDX exploits the architecture of the ion mobility spectrometer through sequential activation, separation and labelling steps. This is advantageous as it decouples the energetic effects of activation away from the labelling cell by virtue of hundreds of thousands of collisions in the high pressure intermediate IM cell. Ions become close to thermalised due to the low energy collisions, with only the travelling waves providing additional propulsion through the cell. Figure 6.8 explores how the uptake changes with energy when ions have been previously (A) or simultaneously (B) activated, in the absence of the trapping or higher pressures associated with IM.

Firstly, WT and acidic GFP ions were activated in the trap followed by labelling in the transfer cell, under an ND_3 pressure of 12 μbar, over the trap voltage range

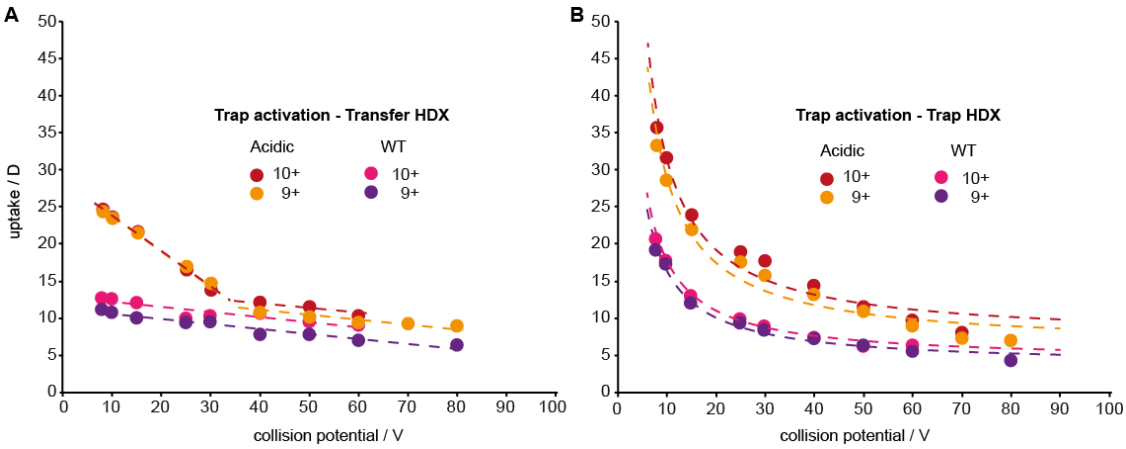


Figure 6.8. Trap activation of GFP variants with simultaneous and subsequent labelling. GFP WT and acidic variants were activated in the trap collision cell, 8 - 80 V, and either labelled subsequently in the transfer, **A**, or simultaneously in the trap, **B**, under a partial ND_3 pressure of 12 µbar. Experiments were operated in ToF mode without ion trapping. The uptake is plotted against the collision voltage for each mutant and charge state. **A** When ions are subsequently labelled a linear decrease is observed, WT, or a two-stage linear, acidic, decrease is observed. When ions are simultaneously accelerated and labelled uptake approximately follows a $^1/_V$ dependency with voltage not seen in **A**. Fits are provided with dashed lines in the corresponding colour.

of 8 - 80 V, Fig.6.8A. This is analogous to the CIU-gpHDX experiments presented previously but without the ion mobility parameters at play.. The uptake is plotted against the collision potential resulting in a linear or bi-linear decrease. The wild type GFP displays a single linear decrease and at all collision voltages exchanges less the acidic counterpart. The wild type has an initial uptake of 11 D and 12 D for the 9+ and 10+ respectively. The acidic mutant has an initial uptake of 25 D decreasing to a comparable level to the wild type, ∼10 D, at around 40 V. The acidic mutant rather than having a single linear decrease appears to have a degree of bi-linearity. The point of inflection appears between 30 - 40 V for both charge states. These profiles are not unlike those seen under IM conditions in Figures 6.3 & 6.4 (Middle), with the wild type appearing more linear than the acidic for both charge states. It is important to note also that the gradients are shallowest at the higher voltages, as seen in the CIU-gpHDX data.

The second experiment, performed in the absence of IM, was the simultaneous activation and labelling in the trap, Fig. 6.8B. Ions were once again activated over 8 - 80 V whilst under a ND_3 labelling pressure of 12 µbar. Little difference in

uptake was observed between the isolated 9+ or 10+ ions. Most notably, the uptake also decreases as a function of collision voltage for both variants with the acidic maintaining a higher uptake over the entire duration. Interestingly, the curves follow a $1/V$ dependency, plotted in dashed lines, reasonably well. As kinetic energy, E_{kin}, is proportional to the acceleration voltage applied and E_{kin} is proportional to the velocity squared, ν^2, these fits also follow a $1/\nu^2$ relationship. Earlier, on page 129, Chapter 4, it was discussed that the capture cross section of an ion is also related to $1/\nu^2$. As ions are accelerated into the trap cell where ions will be labelled, any decrease in the capture cross section could act to reduce the uptake. That all charge states for both the acidic and WT GFP have good fits makes the decreasing capture cross section a likely candidate for the uptake decrease during simultaneous labelling. A decrease in the number of collisions due to a shorter residence time with increased ion velocity could also contribute to this trend, as mentioned in§ 4.3.3. As this chapter has discussed in detail, the structure is also changing with activation making the experiment hard to deconvolute.

Absolute differences between the data collected when operating in and out of ion mobility mode are likely, due to the reasons discussed earlier, namely, the increase in pressure in the activation and labelling cells, and the ion trapping involved in gating prior to IM. The differences in uptake magnitude in these experiments, simultaneous versus subsequent labelling, are likely do to the energy the ions enter the cell and the travelling waves employed for the two different regions, detailed in the methods. It is encouraging that all activation-labelling experiments experience a decrease in uptake with activation irrespective of mode or ion species. The $1/V$ dependency is only observed when ions are simultaneously labelled providing confidence that ions are indeed energetically thermalised in the ion mobility cell in all other cases.

6.3.4 Absolute CCS determination for the activated states of GFP

To determine the degree of unfolding we examined the unfolding intermediates for the WT GFP using a linear drift cell IM instrument. Linear drift cells are capable

of providing an absolute CCS from first principles without external calibration. This is advantageous as accurate calibration of TW-IM data requires an external calibrant whose mobility is comparable to that of the species in question,[23] which becomes difficult for unfolded species of unknown size. As the drift times of the other variants closely resemble the WT, the absolute CCSs determined for the WT conformers can be used to calibrate the acidic and basic mutants.

CCS values for WT GFP were determined according to published protocols[12,24,25] and is briefly explained here. Ions were released into the IM drift cell and migrated under a linear electric field (V), 50 - 120 V. A Gaussian fit is performed to the arrival time distribution to determine the centroid, Fig. 6.9A (inset). A plot of the arrival time centroid against $^1/_V$ is used to determine the inverse of mobility, $^1/_K$, from the slope gradient, demonstrated for the 8+, 9+ and 10+ charge states of the native conformation of WT GFP in Figure 6.9A. The intercept corresponds to the time-zero, t_0, which accounts for time spent outisde the mobility cell before reaching the pusher region. The mobility can then be used to determine the CCS given the known ion mass, gas pressure and temperature, Eq. 2.6. This was performed for the native state to give a CCS of 2477 Å^2, 2481 Å^2, 2493 Å^2 for the 8+, 9+, and 10+ respectively, with errors <5 Å^2. The same procedure was carried out on the most unfolded state, at a trap collision voltage of 70 V, to give CCS values of 2740 Å^2, 3483 Å^2, and 3593 Å^2 for the the 8+, 9+, and 10+, respectively.

A linear calibration curve was then generated using the corrected drift times (dt'), adjusted to account for the time not spent inside the drift tube, and the corrected CCS (CCS'), adjusted to account for the reduced mass and *m/z* dependence, from the native and unfolded states. This produces a sufficiently good linear fit, $R^2 = 0.99$, to be used to calibrate the intermediate unfolded states and the acidic and basic mutants, Fig.6.9B. The determined CCS for each gas-phase intermediate, for all mutants and charge states, are provided in Figure 6.9C. The gas-phase intermediates for the 9+ and 10+ charge states for all mutants have a very similar CCS. This is with the exception of the most unfolded state for the 9+ acidic GFP, which more closely resembles the I3 state seen in the basic 9+ and the acidic 10+.

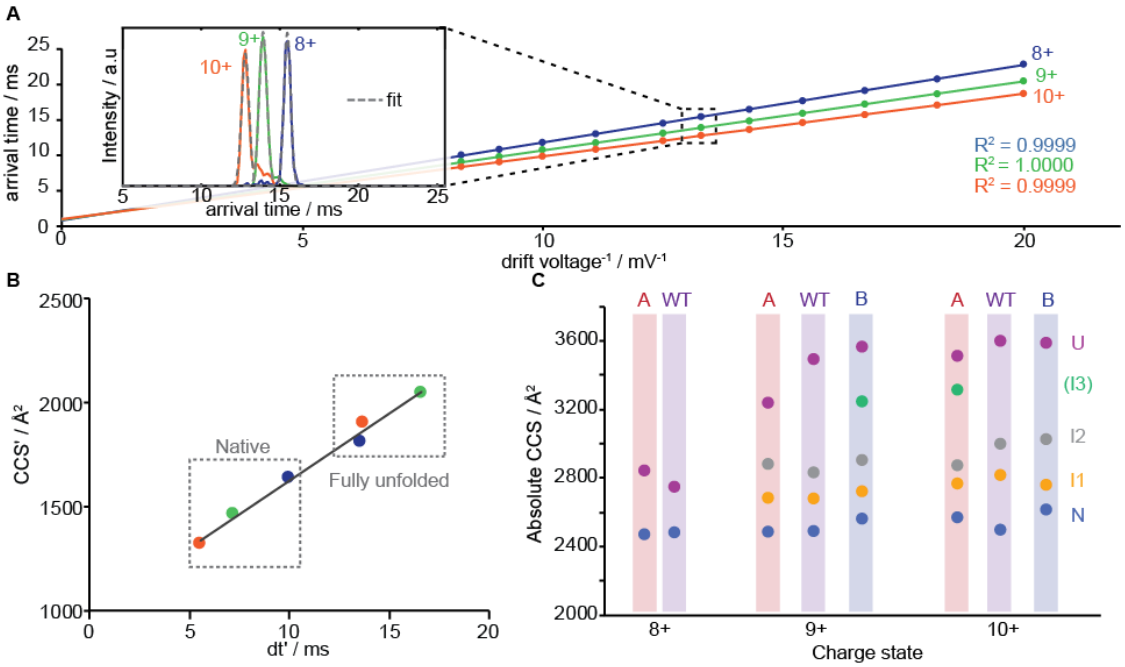


Figure 6.9. Absolute CCS determination of unfolded intermediates. Absolute CCS values were determined for the native and most unfolded states of WT GFP 8+, 9+ and 10+ charge states. **A** A plot of arrival time against the inverse of drift voltage for the 8+, 9+ and 10+ charge states. The gradient is inversely related to the mobility, K, of the ion and the intercept corresponds to time-zero, t_0, which is the time spent outside of the drift cell before detection. The inset shows the raw mobilograms for the 8+, 9+ and 10+ ions at a drift voltage of 75 V. Gaussian fits to determine the arrival time centroid are shown as a dashed grey line overlay. **B** The determined corrected CCS values are plotted against the corrected drift times to produce a linear calibration curve for the native and most unfolded states. **C** The calibration curve was used to calibrate TWIMS data for the remaining intermediate unfolded states of WT GFP and all states of the acidic and basic variants. CCS values are shown for each charge state and variant with the intermediate label assigned earlier in this chapter indicated. For reference the theoretical CCS using a projection approximation for WT GFP is ~2250 Å^2.

The largest CCS a protein poly-peptide chain can adopt is to be fully extended. To understand just how unfolded our most unfolded gas-phase states were the extended maximum was estimated by treating the chain as a rigid cylinder. The dimensions for the 238 residue long chain are 833 - 952 Å, assuming an average residue-residue length of 3.5 - 4 Å,[26] and of average radius 5.5 Å.[27] This corresponds to a surface area of 28976 - 33089 Å^2. Cauchy's geometric relationship between surface area, A_{surf}, of a convex shape and the rotationally averaged projected area, PA, is simply, $PA = \frac{A_{surf}}{4}$. The maximum CCS of the extended GFP can adopt is therefore 7244 - 8282 Å^2. The structures we observe are significantly below this upper limit range and at most are 50% of this value, suggesting significant residual

structure is maintained even in the most unfolded forms.

6.3.5 Gradual side-chain rearrangements during activation

The decreases observed in uptake with prior activation are likely to occur from protection of the labile sites either in the form of steric shielding or by the sequestration of hydrogens in strong interactions such as charge solvation. Evidence from previous gpHDX studies, experimental and theoretical, suggests that electrostatic interactions are able to disrupt and prevent an exchange event.[15,28–30] Shortly after the initial stages of electrospray, side-chains are expected to undergo gas-phase collapse resulting in self-solvation through neighbouring residues or the peptide backbone.[7] An ensemble of interactions will be formed to compensate for the lack of hydration as residual water is lost. When ions are activated, increased energies allow the ion to sample nearby regions of the energy landscape involving the rearrangement of surface side-chains. IM measurements are typically only sensitive to structural changes which alter the CCS by a few percent,[12] most of the proposed surface rearrangements will fall significantly below this. At the onset of a large global structural change, of which IM is sensitive to, much of the interaction network could become 'reset' and the rearrangement of interactions to find a new minima can begin for the new conformer with further activation providing the means to explore such motions. The tendency for the unfolded states to display less of a change in uptake than the prior states could support this as a more unfolded state may be more restricted in the number of interactions it can form.

Determination of the CCS for the unfolded states was discussed in the previous section. The most unfolded state for GFP WT 10+ has a CCS of 3600 Å^2, around a 50% increase relative to the native state, if we calculate the theoretical value based on a rigid cylinder for the fully elongated peptide backbone we would obtain a value 7244 - 8282 Å^2. Whilst this is a crude approximation it does provide an upper limit to the CCS that unfolded GFP can adopt. Clearly collisional activation does not cause the complete elongation during Coulombic unfolding. However, to what extent the protein unravels is difficult to determine. Some simulation studies suggest

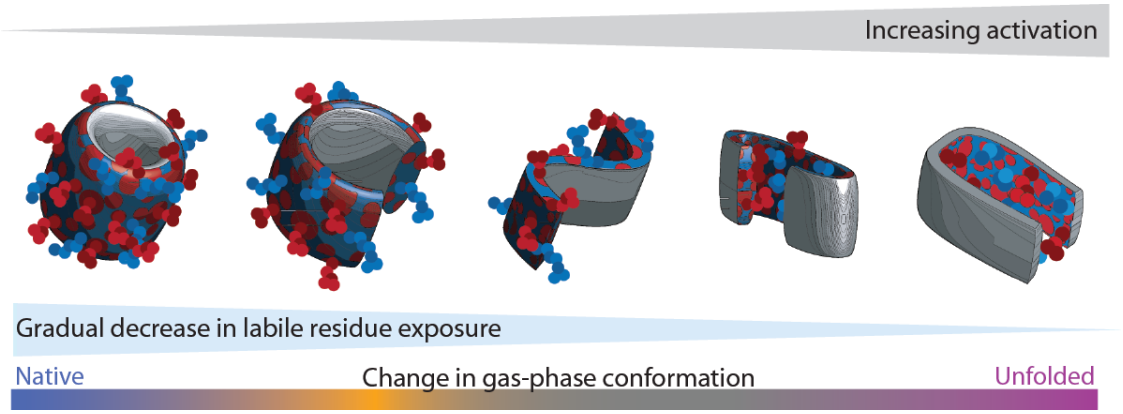


Figure 6.10. Protein surface inversion during gas-phase activation. A hypothesised cartoon schematic of how the surface of the GFP beta barrel might evolve after activation in the gas phase. *Acidic* and *basic* residues that are capable of exchange are depicted in red and blue. Grey areas correspond to non-exchanging hydrophobic regions. As activation is increased the polar, exchangeable, residues begin to self-solvate and the structure gradually becomes inverted as evidenced by gpHDX. Concurrently the structure expands due to poor folding and Coulombic repulsion, supported by ion mobility.

an elongated charged string begins to form from a leaving N-terminus.[31] Elongation of the protein could also cause a decrease in the exchange ability of side-chains as distances between stabilising neighbours becomes greater.[32] Whilst this may certainly be involved, for a given IM conformation the degree of separation between two stabilising neighbours will be limited as the amount of backbone motion will be restricted.

We propose that for these proteins the unfolding is more akin to an inversion than unravelling.[33] Inversion, for a soluble protein, would act to bury the labile hydrogens involved in gpHDX as the chargeable and polar residues self-solvate, Fig. 6.10. In doing so an expansion in CCS would be observed and crucially account for the HDX trends presented here. Simulations in other studies corroborate these hypotheses as they observe the polar residues begin to self-solvate and bury as a function of activation in the gas phase.[34] gpHDX offers a highly complementary chemical probe of the ions surface to the mobility measurement.

6.4 Conclusion

In this chapter we have demonstrated how a gas-phase HDX labelling strategy can be used to selectively probe native protein structure and its structural evolution as it is activated *in vacuo*. Understanding how proteins unfold remains difficult to experimentally probe and as ion mobility experiments continue to utilise the information gained from CIU, often little can be gleaned from the intermediate structures. Gas-phase HDX data presented here suggests the technique is chemically sensitive to these structural differences at a global and local level by reporting on side-chain protection.

We have demonstrated how the resulting CIU-gpHDX profile provides a unique fingerprint complementary to ion mobility which is sensitive to charge state and surface chemistry. The method was shown to be highly reproducible across many in-house repeats and between laboratories. A significant non-linear decrease in uptake was observed as activation is increased with inflections and features in the profile corresponding to the conformer transition seen in the ion mobility dimension. Extracted uptake centroids and gradients from the unfolding plots demonstrated the potential to inform on the structural transitions that ion mobility is insensitive to. Transitions in gpHDX can result in both a positive or negative change in uptake propensity whilst IM only ever observes further unfolding. Additionally, for a given state the uptake does not always remain constant with a tendency in the more native structures to decrease of the conformer's life-time.

These findings in addition to the absolute cross-sections make us believe the polar, exchangeable, side-chains are burying to increase self-solvation, ultimately preventing exchange. The activation provides a means for the side-chains to explore the energy landscape and form, or break, charge-solvating or salt-bridge interactions. When activation reaches a critical point a large-scale structural change is observed in the mobility dimension, after which the process of side-chain exploration begins again. This may be why gpHDX appears as a continuum of changes with some points of inflection whilst IM is step-wise.

To fully utilise the technique and validate our findings one would desire higher structural localisation still, at a level close to that provided by enzymatic digestion in solution. The labile nature of the deuterium label makes fragmentation methods very difficult to exploit as described in Chapter 5. Specific scission of the backbone through rapid fragmentation techniques offer a lot of promise to gas-phase HDX in their ability to fragment before H/D scrambling can occur.[35–37] If this is realised it presents a means of obtaining peptide level information on the gas-phase structural changes we can currently observe. In doing so hypotheses can be tested and our fundamental understanding of how ion structure develops significantly improved.

6.5 Methods

Materials

All materials are as described in Chapter 4. GFP variants are the same as those described in §4.3.6 from the same expression and purification.

Native MS and gpHDX

All gpHDX experiments were performed on a Synapt G1 Q-IM-ToF (Waters, Wilmslow, UK) modified for gpHDX and high mass transmission, Chapter 4, with the exception of the inter-lab study using a Synapt G2 Q-IM-ToF (Waters), Fig. 6.7. Protein samples were buffer exchanged into 200 mM ammonium acetate using Micro Bio-Spin columns (Bio-Rad Laboratories, Hercules, CA, USA). All CIU-gpHDX experiments were acquired at room temperature using nano-electrospray in positive polarity. Ions were isolated using a quadrupole selection followed by trap cell activation and transfer gpHDX. Trap activation potentials were varied from 8 - 80 V and for unfolding experiments were acquired every 1 - 5 V. Argon was used as a collision partner (trap), nitrogen as a mobility drift gas (IM) and ND_3 as a HDX reagent (transfer). Specific instrument parameters are detailed in the following: capillary 1.4 kV, sample cone 10 V, extraction cone 3 V, transfer activation 4 V, IM DC bias 15 V, trap TW velocity 200 $m\,s^{-1}$, trap TW height 0.2 V, IM TW velocity 260 $m\,s^{-1}$, IM TW height 7.5 V, transfer TW velcoity 248 $m\,s^{-1}$, transfer

TW height 10 V. Pressures were set without the IM gas on to determine partial pressures. Pressures: backing pressure 3.9 mbar, set trap Ar 12 μbar, set transfer ND_3 12 μbar, final IM N_2 0.57 mbar. Stringent quadrupole selections for each charge state were set using LM and HM values of >15. The travelling wave voltages were monitored using an external oscilloscope to ensure no electrical "cross-talk" was present between the trap collision cell and transfer labelling cell.

The data acquired for the interlab study was performed on a Synapt G2 (Waters) instrument with parameters as close as practically possible to those above. Parameters were as follows: capillary 1.4 kV, sample cone 10 V, extraction cone 2 V, transfer activation 4 V, IM DC bias 25 V, trap TW velocity 310 $m\,s^{-1}$, trap TW height 6 V, IM TW velocity 200 $m\,s^{-1}$, IM TW height 7.5 V, transfer TW velocity 250 $m\,s^{-1}$, transfer TW height 10 V. Pressures: backing pressure 3.7 mbar, set trap Ar 5 μbar, set transfer ND_3 75 μbar, final IM N_2 1 mbar.

Data acquired in ToF mode, Fig. 6.8, was as previously described with an ND_3 pressure of 12 μbar in the indicated labelling cell, either the trap or transfer. Travelling wave conditions were set to 300 $m\,s^{-1}$ and 3 V in the trap and 248 $m\,s^{-1}$ and 0.2 V in the transfer.

Absolute CCS values were determined for GFP ions on a modified linear drift tube instrument (Waters).[24] Nitrogen was used as an IM collision partner at a pressure of 1.8 torr at 26 °C. The following conditions were used: capillary 1.4 kV, sample cone 10 V, extraction cone 4 V, trap activation 10 V. The IM drift voltage was varied from 50 - 200 V with a IM DC bias of 20 V.

Data processing

Processing of mass spectra was carried out using MassLynx V4.1 software (Waters, Wilmslow, UK) with the mean filter smoothing of 20 × 1. CIU unfolding plots were generated using PULSAR software with a histogram bin size defined as 2.5 V in the x-axis and 0.5 ms in the y-axis. Arrival time distributions were calculated using PULSAR and a model fit to the data as described previously.[18] We adapted PULSAR to allow the mass distribution to be extracted for a given mobility region,

allowing conformer specific m/z distributions to be calculated. An external Python script was developed to allow an 2D-histogram to be produced for the uptake dimension as a function of collision energy. Each collision voltage *slice* is normalised to 1 and then interpolated to give a histogram of uptake against collision voltage for all species present (i.e. no conformer extraction). Additionally, the external script extracts the uptake distribution for each conformer at a given collision slice and performs the interpolation. The uptake distribution for a given conformer is normalised such that the sum of all conformer distributions is equal to 1 within a collision voltage slice. The weighting of each conformer is determined by the relative conformer abundance observed in the IM dimension. The final histograms presented were scaled to a power function ($y = x^8$), and re-normalised, to highlight differences between the broad uptake distributions, the corresponding intensity is shown in the colour bar in figure 6.3.

6.6 Acknowledgments

I thank Justin Benesch first and foremost for many engaging discussions around the work presented in this Chapter, in what turned out to be the most intellectually stimulating and satisfying work of my DPhil. To Tim Allison for his help in adapting PULSAR and assistance in creating external scripts to automate data analysis and visualisation. I also thank Max Hantke for helping me improve my python scripts. I very much appreciate advice and discussions with Kevin Giles and Jeff Brown on travelling wave technicalities. I thank Jon Hopper, Tim Allison and Georg Hochberg for assistance during protein expression and provision of the GFP plasmids. I thank Mark Agasid for his critical review of this chapter.

References

[1] Van Den Heuvel, R. H. H.; Heck, A. J. R. Native protein mass spectrometry: From intact oligomers to functional machineries. *Current Opinion in Chemical Biology* **2004**, *8*, 519–526.

[2] Benesch, J. L. P.; Aquilina, J. A.; Ruotolo, B. T.; Sobott, F.; Robinson, C. V. Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. *Chemistry & biology* **2006**, *13*, 597–605.

[3] Skinner, O. S.; Haverland, N. A.; Fornelli, L.; Melani, R. D.; Do Vale, L. H.; Seckler, H. S.; Doubleday, P. F.; Schachner, L. F.; Srzentić, K.; Kelleher, N. L.; Compton, P. D. Top-down characterization of endogenous protein complexes with native proteomics. *Nature Chemical Biology* **2018**, *14*, 36–41.

[4] Shepherd, D. A.; Marty, M. T.; Giles, K.; Baldwin, A. J.; Benesch, J. L. Combining tandem mass spectrometry with ion mobility separation to determine the architecture of polydisperse proteins. *International Journal of Mass Spectrometry* **2015**, *377*, 663–671.

[5] Li, H.; Nguyen, H. H.; Loo, R. R.; Campuzano, I. D.; Loo, J. A. An integrated native mass spectrometry and topdown proteomics method that connects sequence to structure and function of macromolecularcomplexes. *Nature Chemistry* **2018**, *10*, 139–148.

[6] Marklund, E. G.; Degiacomi, M. T.; Robinson, C. V.; Baldwin, A. J.; Benesch, J. L. Collision Cross Sections for Structural Proteomics. *Structure* **2015**, *23*, 791–799.

[7] Breuker, K.; McLafferty, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase, 10(-12) to 10(2) s. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18145–52.

[8] Chandler, S. A.; Benesch, J. L. Mass spectrometry beyond the native state. *Current Opinion in Chemical Biology* **2018**, *42*, 130–137.

[9] Niu, S.; Rabuck, J. N.; Ruotolo, B. T. Ion mobility-mass spectrometry of intact protein-ligand complexes for pharmaceutical drug discovery and development. *Current Opinion in Chemical Biology* **2013**, *17*, 809–817.

[10] Zhong, Y.; Han, L.; Ruotolo, B. T. Collisional and Coulombic unfolding of gas-phase proteins: high correlation to their domain structures in solution. *Angewandte Chemie (International ed. in English)* **2014**, *53*, 9209–12.

[11] Eschweiler, J. D.; Martini, R. M.; Ruotolo, B. T. Chemical probes and engineered constructs reveal a detailed unfolding mechanism for a solvent-free multidomain protein. *Journal of the American Chemical Society* **2017**, *139*, 534–540.

[12] Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nature protocols* **2008**, *3*, 1139–52.

[13] Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. Collision induced unfolding of isolated proteins in the gas phase: past, present, and future. *Current Opinion in Chemical Biology* **2018**, *42*, 93–100.

[14] Green, M. K.; Lebrilla, C. B. Ion-molecule reactions as probes of gas-phase structures of peptides and proteins. *Mass spectrometry reviews* **1997**, *16*, 53–71.

[15] Mistarz, U. H.; Brown, J. M.; Haselmann, K. F.; Rand, K. D. Probing the Binding Interfaces of Protein Complexes Using Gas-Phase H/D Exchange Mass Spectrometry. *Structure* **2016**, *24*, 310–318.

[16] Valentine, S. J.; Clemmer, D. E. H/D Exchange Levels of Shape-Resolved Cytochrome c Conformers in the Gas Phase. *Journal of the American Chemical Society* **1997**, *119*, 3558–3566.

[17] Badman, E. R.; Hoaglund-Hyzer, C. S.; Clemmer, D. E. Monitoring Structural Changes of Proteins in an Ion Trap over ∼10-200 ms: Unfolding Transitions in Cytochrome c Ions. *Analytical Chemistry* **2001**, *73*, 6000–6007.

[18] Allison, T. M.; Reading, E.; Liko, I.; Baldwin, A. J.; Laganowsky, A.; Robinson, C. V. Quantifying the stabilizing effects of protein–ligand interactions in the gas phase. *Nature Communications* **2015**, *6*, 8551.

[19] Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry* **2008**, *80*, 9689–9699.

[20] Devine, P. W.; Fisher, H. C.; Calabrese, A. N.; Whelan, F.; Higazi, D. R.; Potts, J. R.; Lowe, D. C.; Radford, S. E.; Ashcroft, A. E. Investigating the Structural Compaction of Biomolecules Upon Transition to the Gas-Phase Using ESI-TWIMS-MS. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 1855–1862.

[21] Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V. Charge-state dependent compaction and dissociation of protein complexes: Insights from ion mobility and molecular dynamics. *Journal of the American Chemical Society* **2012**, *134*, 3429–3438.

[22] Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in travelling wave ion mobility resolution. *Rapid Communications in Mass Spectrometry* **2011**, *25*, 1559–1566.

[23] Allison, T. M.; Landreh, M.; Benesch, J. L.; Robinson, C. V. Low Charge and Reduced Mobility of Membrane Protein Complexes Has Implications for Calibration of Collision Cross Section Measurements. *Analytical Chemistry* **2016**, *88*, 5879–5884.

[24] Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Analytical chemistry* **2010**, *82*, 9557–65.

[25] Thalassinos, K.; Grabenauer, M.; Slade, S. E.; Hilton, G. R.; Bowers, M. T.; Scrivens, J. H. Characterization of Phosphorylated Peptides Using Traveling Wave-Based and Drift Cell Ion Mobility Mass Spectrometry. *Analytical Chemistry* **2009**, *81*, 248–254.

[26] Ainavarapu, S. R. K.; Brujić, J.; Huang, H. H.; Wiita, A. P.; Lu, H.; Li, L.; Walther, K. A.; Carrion-Vazquez, M.; Li, H.; Fernandez, J. M. Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. *Biophysical Journal* **2007**, *92*, 225–233.

[27] Waugh, D. F. Proteins and Their Interactions. *Reviews of Modern Physics* **1959**, *31*, 84–93.

[28] Rand, K. D.; Pringle, S. D.; Morris, M.; Brown, J. M. Site-Specific Analysis of Gas-Phase Hydrogen/Deuterium Exchange of Peptides and Proteins by Electron Transfer Dissociation. **2012**,

[29] Lee, S. W.; Lee, H. N.; Kim, H. S.; Beauchamp, J. L. Selective binding of crown ethers to protonated peptides can be used to probe mechanisms of H/D exchange and collision-induced dissociation reactions in the gas phase. *Journal of the American Chemical Society* **1998**, *120*, 5800–5805.

[30] Wyttenbach, T.; Paizs, B.; Barran, P.; Breci, L.; Liu, D.; Suhai, S.; Wysocki, V. H.; Bowers, M. T. The Effect of the Initial Water of Hydration on the Energetics, Structures, and H/D Exchange Mechanism of a Family of Pentapeptides: An Experimental and Theoretical Study. *Journal of the American Chemical Society* **2003**, *125*, 13768–13775.

[31] Fegan, S. K.; Thachuk, M. Controlling Dissociation Channels of Gas-Phase Protein Complexes Using Charge Manipulation. *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 722–728.

[32] Geller, O.; Lifshitz, C. A Fast Flow Tube Study of Gas Phase H/D Exchange of Multiply Protonated Ubiquitin. *The Journal of Physical Chemistry A* **2005**, *109*, 2217–2222.

[33] Wolynes, P. G. Biomolecular folding in vacuo!!!(?). *Proceedings of the National Academy of Sciences* **1995**, *92*, 2426–2427.

[34] Landreh, M.; Marklund, E. G.; Uzdavinys, P.; Degiacomi, M. T.; Coincon, M.; Gault, J.; Gupta, K.; Liko, I.; Benesch, J. L.; Drew, D.; Robinson, C. V. Integrating mass spectrometry with MD simulations reveals the role of lipids in Na+/H+ antiporters. *Nature Communications* **2017**, *8*, 1–9.

[35] Brodie, N. I.; Huguet, R.; Zhang, T.; Viner, R.; Zabrouskov, V.; Pan, J.; Petrotchenko, E. V.; Borchers, C. H. Top-Down Hydrogen-Deuterium Exchange Analysis of Protein Structures Using Ultraviolet Photodissociation. *Analytical Chemistry* **2018**, *90*, 3079–3082.

[36] Mistarz, U. H.; Bellina, B.; Jensen, P. F.; Brown, J. M.; Barran, P. E.; Rand, K. D. UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Analytical Chemistry* **2018**, *90*, 1077–1080.

[37] Pan, J.; Heath, B. L.; Jockusch, R. A.; Konermann, L. Structural interrogation of electrosprayed peptide ions by gas-phase H/D exchange and electron capture dissociation mass spectrometry. *Analytical Chemistry* **2012**, *84*, 373–378.

7

Concluding remarks

The research I have presented in this thesis demonstrates the versatile role mass spectrometric methods have in structural biology. I have shown how H/D exchange (HDX) labelling strategies can be applied using two approaches: probing protein structures in solution, or directly in the gas-phase. The more mature of these methods, solution HDX, has a proven track-record of reporting on the dynamics and three-dimensional structure of protein conformers and assemblies. In collaborative studies detailed in Chapter 3, I demonstrated the sensitivity that solution labelling has to quantitatively probe the dynamics of a light-induced conformational change. Here, I not only identified a highly localised region of induced flexibility, but, owing to the presence of EX1 kinetic behaviour, was able to determine the rate at which the conformational change occurs. Additionally, by performing a *pulsed* labelling experiment and monitoring the return to the *dark-state*, I was able to demonstrate the reversibility of the change and its unusually long-lived nature.

In this chapter, I also showed how labelling can be used to infer the interface in a homo-dimeric complex. To identify buried interfacial residues, I compared the exchange kinetics of the "dimeric" ancestor of haemoglobin to its "monomeric" form, generated through mutation and titration. Regions were identified that were less protected than in the intact dimer and were shown to be consistent across

all of the "monomeric" systems studied. From these data, I was able to suggest that the interface was consistent with the *packing* interface observed in the extant haemoglobin hetero-tetramer.

The solution HDX work performed on the ancestral haemoglobin also highlighted the potential complications of labelling a solution ensemble. Native mass spectrometry showed that the species present were not simply a monomer and dimer, but also that they existed in their *apo* and *holo* forms. For these studies, where the sample complexity remained low, this was of little concern as we were able to account for the behaviour using distributions determined by native MS. That native MS can identify each species with high resolution and manipulate each ion with increased fidelity presents a compelling case for performing the labelling reaction inside the spectrometer.

To this end, Chapter 4 explains my implementation of gas-phase HDX (gpHDX) in an ion mobility (IM) mass spectrometer. In doing so, I demonstrated the efficacy of the HDX reaction for small protein and peptide systems, whose uptake may be tuned using ion manipulations. I then explored how gpHDX may be used to probe ion structure using a model system of charged GFP variants. I demonstrate that gpHDX is readily able to distinguish between ions of different surface chemistries but of the same charge and size. Interestingly, the exchange behaviour between variants did not display the same trend as the theoretical uptake would suggest, but rather that carboxylic acid residues have a dominant role in exchange.

Having demonstrated the sensitivity of gpHDX to the surface of different proteins, I then used the technique to probe the gas-phase conformational space of proteins. Proteins undergo unfolding transitions in the gas-phase during collisional activation: collision induced unfolding (CIU). Whilst these unfolding trajectories can offer a characteristic fingerprint, little is experimentally known of the structural transitions they undergo. In Chapter 6, I used gpHDX to label collisionally activated gas-phase conformers. By coupling the method to an IM instrument, we were able to extract the conformer specific uptake. This revealed that, as ions undergo activation,

side-chains become increasingly more protected, even whilst the global size of the ion does not change. The uptake and CIU profiles represent two orthogonal approaches for probing gas-phase structure at a global and local level. The data presented here suggests that, for the proteins studied, the exchangeable side chains undergo many gradual rearrangements as a function of energy that act to shield them from exchange. As the energy is increased, the ion undergoes a stepwise increase in mobility corresponding, I argue, to a larger-scale backbone rearrangement. That gpHDX and IM appear to report on different aspects of protein structure make this approach highly complementary.

If protein complexes are activated further still, then dissociation of a subunit can occur, collision induced dissociation (CID). In Chapter 5, I explored how gpHDX may be used to probe the dissociation pathway and understand how deuterium is partitioned during the process. I showed that, for small systems, if a protein complex is labelled and a subunit is dissociated, the ejected subunit leaves with approximately its share of the original incorporation. For much larger systems, the ejected subunit had a small but significant increase accounted for by the charge-precipitated unfolding mechanism of CID. Understanding these processes further would allow one to use the monomer uptake as a proxy for the intact complex, which is particularly important for heteromeric complexes.

Recent advances in "top-down" methodologies make the prospect of coupling rapid fragmentation to workflows presented here tantalisingly close. The ability to localise uptake to a near-residue level will not only allow for detailed gas-phase structural studies, but will, importantly, open up gas-phase labelling as a viable alternative to traditional solution techniques.