

Electronic Thesis and Dissertation Repository

---

7-20-2015 12:00 AM

## Complementary Mass Spectrometry Methods for Characterizing Protein Folding, Structure, and Dynamics


Siavash Vahidi, *The University of Western Ontario*

Supervisor: Dr. Lars Konermann, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Chemistry

© Siavash Vahidi 2015

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

 Part of the [Analytical Chemistry Commons](#), [Biochemistry, Biophysics, and Structural Biology Commons](#), and the [Physical Chemistry Commons](#)

---

### Recommended Citation

Vahidi, Siavash, "Complementary Mass Spectrometry Methods for Characterizing Protein Folding, Structure, and Dynamics" (2015). *Electronic Thesis and Dissertation Repository*. 3011.  
<https://ir.lib.uwo.ca/etd/3011>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

COMPLEMENTARY MASS SPECTROMETRY METHODS FOR  
CHARACTERIZING PROTEIN FOLDING, STRUCTURE, AND DYNAMICS

(Thesis format: Integrated Article)

by

Siavash Vahidi

Graduate Program in Chemistry

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies  
The University of Western Ontario  
London, Ontario, Canada

© Siavash Vahidi 2015

## **Abstract**

Proteins are involved in virtually every biochemical process. A comprehensive characterization of factors that govern protein function is essential for understanding the biomedical aspects of human health. This dissertation aims to develop complementary mass spectrometry-based methods and apply them to solve problems pertaining to the area of protein structure, folding, and dynamics.

Chapter 2 uses fast photochemical oxidation of proteins (FPOP) to characterize partially disordered conformers populated under semi-denaturing conditions at equilibrium. In FPOP,  $\cdot\text{OH}$  generated by laser photolysis of  $\text{H}_2\text{O}_2$  introduces oxidative modifications at solvent accessible side chains. By contrast, buried sites are protected from radical attack. Using apomyoglobin (aMb), it was demonstrated that under optimized conditions undesired effects can be almost completely eliminated and detailed structural information can be obtained.

Chapter 3 combines FPOP with submillisecond mixing to enable studying early events in protein folding in a kinetic fashion. aMb served as a model system for these measurements. Spatially-resolved changes in solvent accessibility follow the folding process. Data revealed that early aMb folding events are driven by both local and sequence-remote docking of hydrophobic side chains. Assembly of a partially formed scaffold after 0.2 ms of folding is followed by stepwise consolidation that ultimately yields the native state. The mixer used improved the time resolution by a factor of 50 compared to earlier FPOP experiments. In conjunction with slower mixing techniques, folding pathways from fractions of a millisecond all the way to minutes are monitored.

Chapter 4 uses ion mobility mass spectrometry (IM-MS) to explore the structural relationship between semi folded solution and gas phase protein conformers. Collision cross sections (CCSs) provide a measure of analyte size. Mb was used as model system because it follows a sequential unfolding pathway that comprises two partially disordered states. IM-MS data showed that the degree of gas phase unfolding is not strongly correlated with the corresponding solution. Gas phase unfolding as well as collapse events can lead to disparities between gaseous and solution structures for partially unfolded proteins. IM-MS data on non-native conformers should therefore be interpreted with caution.

Chapter 5 uses HDX-MS to examine the role of conformational dynamics for the function of multi-protein molecular machines such as  $F_0F_1$  ATP synthase. HDX-MS monitors backbone deuteration kinetics in the presence of  $D_2O$ . Disordered segments exchange more rapidly than those in tightly folded regions. Measurements of spatially-resolved deuterium are performed using LC-MS. It was found that the H-bonding network of key power transmission elements is insensitive to PMF-induced mechanical stress. Unexpectedly, HDX-MS reveals a pronounced destabilization of the  $\gamma$  C-terminus during rotational catalysis under proton motive force (PMF). The behavior of  $\gamma$  is attributed to kinetic friction within the apical rotor bearing.

**Keywords:** protein structure | folding intermediates | conformational dynamics | mass spectrometry | kinetic | rapid mixing | covalent labeling | hydroxyl radicals | hydrogen/deuterium exchange | electrospray ionization | ion mobility spectrometry

## Co-Authorship Statement

The work in Chapters 2, 3 and 4 were published in the following articles respectively:

Vahidi S, Stocks BB, Liaghati-Mobarhan Y, & Konermann L (2012) Mapping pH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* 84:9124-9130. Reproduced with permission. ©American Chemical Society

Vahidi S, Stocks BB, Liaghati-Mobarhan Y, & Konermann L (2013) Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling. *Anal. Chem.* 85:8618-8625. Reproduced with permission. ©American Chemical Society

Vahidi S, Stocks BB, & Konermann L (2013) Partially Disordered Proteins Studied by Ion Mobility-Mass Spectrometry: Implications for the Preservation of Solution Phase Structure in the Gas Phase. *Anal. Chem.* 85:10471-10478. Reproduced with permission. ©American Chemical Society

The work in Chapter 5 has been incorporated into the following article:

Vahidi S, Bi Y, Dunn SD, & Konermann L (2015). H/D Exchange Mass Spectrometry Reveals Friction-Mediated Torsional Stress During  $F_0F_1$  ATP Synthase Operation. *In preparation.*

The original draft for each of the above articles was prepared by the author. Subsequent revisions were by the author and Dr. Lars Konermann. The computer code used in Chapter 4 for generating random protein structures was written by Dr. Konermann. Protein expression/purification and functional assays appearing in Chapters 5 were completed by the laboratory of Dr. Stanley D. Dunn. All other experimental work and data analysis were performed by the author under the supervision of Dr. Konermann.

## **Acknowledgments**

My advisor Prof. Lars Konermann is a fantastic teacher, a gentleman, and a true scholar with an endless passion for science. Working with Lars was the main driving force behind my decision to come to Western. He deserves full credit for inconspicuously instilling in me a rational framework within which I could address scientific questions. Lars has managed to establish a brand of science that is distinctly his and a research program that deserved every bit of my best effort. Part of my immense respect for Lars comes from neither of us needing to go easy on the other's opinion to have a good relationship. I am incredibly grateful for the position I held for five years in the group and sad to give up my desk. I take solace in the permanent nature of being a former Konermann Lab member.

I am thankful to all 27 members of the Lab whose time overlapped with mine. In particular, the senior PhD students and postdocs who populated the Lab upon my arrival had a profound impact on me. I was lucky to have a desk strategically positioned next to Bradley Stocks' and pick his brain on daily basis. It was not a coincidence that the beginning of our collaboration and friendship marked the end of my struggles with oxidative labeling of proteins. Thanks for taking me under your wings and teaching me the nuances of maintaining and troubleshooting lab equipment and the art of performing robust experiments.

The Biochemistry Department has been a great source of support. Many thanks to Drs. Gary Shaw and James Choy for devoting many hours of their time to me for career advice and unselfishly supporting my future career direction. I would like to extend my sincere thanks to Prof. Stanley Dunn and Yumin Bi for their friendship, biochemical extraordinaire, seemingly endless supply of exotic proteins, and Stan's many teachings of Americanisms.

My efforts at Western were actively supported by my father, mother and brother. Thank you for filling me up with encouragement, affection, and vision. Finally, a big thank you to my better half, Sara, for taking me as I am, keeping me sane throughout my struggles in grad school, and fostering my interest in the protein world.

# Table of Contents

<b>Title</b> .....	<b>i</b>
<b>Abstract</b> .....	<b>ii</b>
<b>Co-Authorship Statement</b> .....	<b>iv</b>
<b>Acknowledgments</b> .....	<b>v</b>
<b>Table of Contents</b> .....	<b>vi</b>
<b>List of Symbols and Abbreviations</b> .....	<b>x</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1 The protein structure-function paradigm .....	1
1.1.1 <i>The Basics of Protein Structure and Folding</i> .....	1
1.1.2 <i>Protein Structural Hierarchy</i> .....	2
1.1.3 <i>Factors Contributing to Protein Stability</i> .....	3
1.1.4 <i>Protein Folding Mechanisms</i> .....	4
1.1.5 <i>Protein Dynamics</i> .....	8
1.1.6 <i>Protein Misfolding</i> .....	9
1.2 Experimental Methods to Study Protein Structural and Motional Properties.....	9
1.2.1 <i>X-ray Crystallography</i> .....	9
1.2.2 <i>Cryogenic Electron Microscopy (Cryo-EM)</i> .....	11
1.2.3 <i>Nuclear Magnetic Resonance (NMR) Spectroscopy</i> .....	12
1.2.4 <i>Optical spectroscopy</i> .....	14
1.2.5 <i>Computational Methods</i> .....	17
1.3 Fundamentals of Mass Spectrometry .....	18
1.3.1 <i>Electrospray Ionization</i> .....	18
1.3.2 <i>Mass Analyzers</i> .....	20
1.4 Structural Mass Spectrometry of Proteins .....	23
1.4.1 <i>Hydrogen/Deuterium Exchange (HDX)</i> .....	24
1.4.2 <i>Covalent Labeling</i> .....	29
1.4.3 <i>Crosslinking</i> .....	32
1.4.4 <i>Native Electrospray Mass Spectrometry (Native ESI-MS)</i> .....	33

1.4.5	<i>Ion Mobility Spectrometry (IMS)</i> .....	34
1.5	Integrative Structural Biology.....	37
1.6	Scope of Thesis.....	38
1.7	References.....	39
	<b>Chapter 2: Mapping pH-Induced Protein Structural Changes under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry .....</b>	<b>62</b>
2.1	Introduction.....	62
2.2	Experimental Section.....	65
2.2.1	<i>Materials</i> .....	65
2.2.2	<i>Optical Spectroscopy</i> .....	66
2.2.3	<i>Oxidative Labeling</i> .....	66
2.2.4	<i>LC/ESI-MS</i> .....	68
2.2.5	<i>Data Analysis</i> .....	69
2.3	Results and Discussion .....	69
2.3.1	<i>Optical Spectroscopy</i> .....	69
2.3.2	<i>Validation of FPOP Strategy Using Test Peptides</i> .....	71
2.3.3	<i>Intact Protein Analyses</i> .....	74
2.3.4	<i>Solvent Accessibility Measurements Using Normalized Oxidation Levels</i> .....	76
2.3.5	<i>Structure of Native aMb</i> .....	78
2.3.6	<i>The pH 4 Intermediate</i> .....	82
2.4	Conclusions.....	83
2.5	References.....	85
	<b>Chapter 3: Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling .....</b>	<b>91</b>
3.1	Introduction.....	91
3.2	Experimental Section.....	95
3.2.1	<i>Materials</i> .....	95
3.2.2	<i>Submillisecond Mixer Design</i> .....	96
3.2.3	<i>Mixer Characterization</i> .....	97
3.2.4	<i>Oxidative Labeling</i> .....	99



3.2.5	<i>LC/ESI-MS and Data Analysis</i> .....	102
3.3	Results and Discussion .....	103
3.3.1	<i>Global Solvent Accessibility Changes</i> .....	104
3.3.2	<i>Solvent Accessibility of Individual Protein Segments</i> .....	107
3.3.3	<i>Submillisecond Events</i> .....	107
3.3.4	<i>Folding Events Beyond 1 ms</i> .....	112
3.4	Conclusions.....	113
3.5	References.....	115
	<b>Chapter 4: Partially Disordered Proteins Studied by Ion Mobility-Mass Spectrometry: Implications for the Preservation of Solution Phase Structure in the Gas Phase</b> .....	<b>122</b>
4.1	Introduction.....	122
4.2	Experimental Section.....	126
4.2.1	<i>Materials and Sample Preparation</i> .....	126
4.2.2	<i>Ion Mobility-Mass Spectrometry</i> .....	126
4.2.3	<i>Modeling</i> .....	129
4.3	Results and Discussion .....	129
4.3.1	<i>Mb Unfolding in Solution Monitored by Optical Spectroscopy</i> .....	129
4.3.2	<i>Mb Unfolding Probed Via ESI Charge State Distributions</i> .....	130
4.3.3	<i>Ion Mobility Results</i> .....	132
4.3.4	<i>Collision Cross Sections of Model Structures</i> .....	136
4.3.5	<i>Effects of Solution Phase Structure on Gas Phase Conformation</i> .....	137
4.3.6	<i>Intensity-Weighted CCS Distributions</i> .....	141
4.4	Conclusions.....	143
4.5	References.....	144
	<b>Chapter 5: H/D Exchange Mass Spectrometry Reveals Friction-Mediated Torsional Stress During F<sub>0</sub>F<sub>1</sub> ATP Synthase Operation</b> .....	<b>153</b>
5.1	Introduction.....	153
5.2	Experimental Section .....	155
5.2.1	<i>Materials</i> .....	155
5.2.2	<i>Preparation of inside-out membrane vesicles</i> .....	155
5.2.3	<i>Functional Enzyme Assays</i> .....	157

5.2.4	<i>Hydrogen/Deuterium Exchange</i> .....	157
5.2.5	<i>Liquid Chromatography-Mass Spectrometry</i> .....	159
5.2.6	<i>Peptide Mapping</i> .....	161
5.2.7	<i>Sequence coverage and proteomic profile of membranes</i> .....	161
5.2.8	<i>HDX Data Analysis</i> .....	162
5.3	Results and Discussion .....	163
5.4	References.....	173
	<b>Chapter 6: Conclusions</b> .....	<b>177</b>
6.1	Summary .....	177
6.2	Future Directions .....	180
6.2.1	<i>Characterization of the Labeling Pulse in FPOP</i> .....	180
6.2.2	<i>Pushing the HDX-MS Envelope</i> .....	182
6.3	References.....	183
	<b>Appendix I-Permissions</b> .....	<b>187</b>
	<b>Curriculum Vitae</b> .....	<b>191</b>

## List of Symbols and Abbreviations

ADP - adenosine diphosphate

AFM - atomic force microscopy

AMP-PNP - adenosine 5'-( $\beta,\gamma$ -imido)triphosphate

ATP - adenosine triphosphate

CCS - collision cross section

CD - circular dichroism

CEM - chain ejection model

CEST - chemical exchange dependent saturation transfer

CID - collision-induced dissociation

CPMG - Carr-Purcell-Meiboom-Gill

CRM - charge residue model

Cryo-EM - cryogenic electron microscopy

cyt *c* - cytochrome *c*

DC - direct current

DESI - desorption electrospray ionization

*E. coli* - *Escherichia coli*

ECD - electron capture dissociation

ECD - electron capture dissociation

ESI - electrospray ionization

ETD - electron transfer dissociation

FCCP - carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone

FPOP - fast photochemical oxidation of proteins

FRET - Förster resonant energy transfer

FT-ICR - Fourier transform ion cyclotron resonance

GEE - glycine ethyl ester

HDX - hydrogen/deuterium exchange

IDP - intrinsically disordered proteins

IEM - ion evaporation model

IMS - ion mobility spectrometry

LC - liquid chromatography

MAIV - matrix assisted ionization vacuum

MD - molecular dynamics

MS - mass spectrometry

Mb - myoglobin

N - native state

NMR - nuclear magnetic resonance

NOESY - nuclear Overhauser effect spectroscopy

$\cdot\text{OH}$  - hydroxyl radical

PDB - protein data bank

PEP - phospho(enol)pyruvic acid

PMF - proton motive force

PTM - post translational modifications

RF - radio frequency

SAXS - small-angle X-ray scattering

SDS PAGE - sodium dodecyl sulfate polyacrylamide gel electrophoresis

$\tau$  - radical lifetime

$[\theta]$  - mean residue ellipticity

$t_{1/2}$  - half-life

TEMPO - 2,2,6,6-Tetramethylpiperidin-1-yl)oxyl

tricine - N-(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine

TROSY - transverse relaxation optimized spectroscopy

TWIM - travelling-wave ion mobility

UPLC – ultra performance liquid chromatography

XFEL - X-ray free electron lasers

XL - cross-linking

# Chapter 1: Introduction

## 1.1 The protein structure-function paradigm

### 1.1.1 The Basics of Protein Structure and Folding

Proteins are the molecular devices through which the genetic information encoded in the DNA is put into action. They are highly abundant and occur in all cells and come in a wide variety of sizes; from relatively small proteins to huge macromolecules with molecular weights in the MDa range (Figure 1.1A&B).

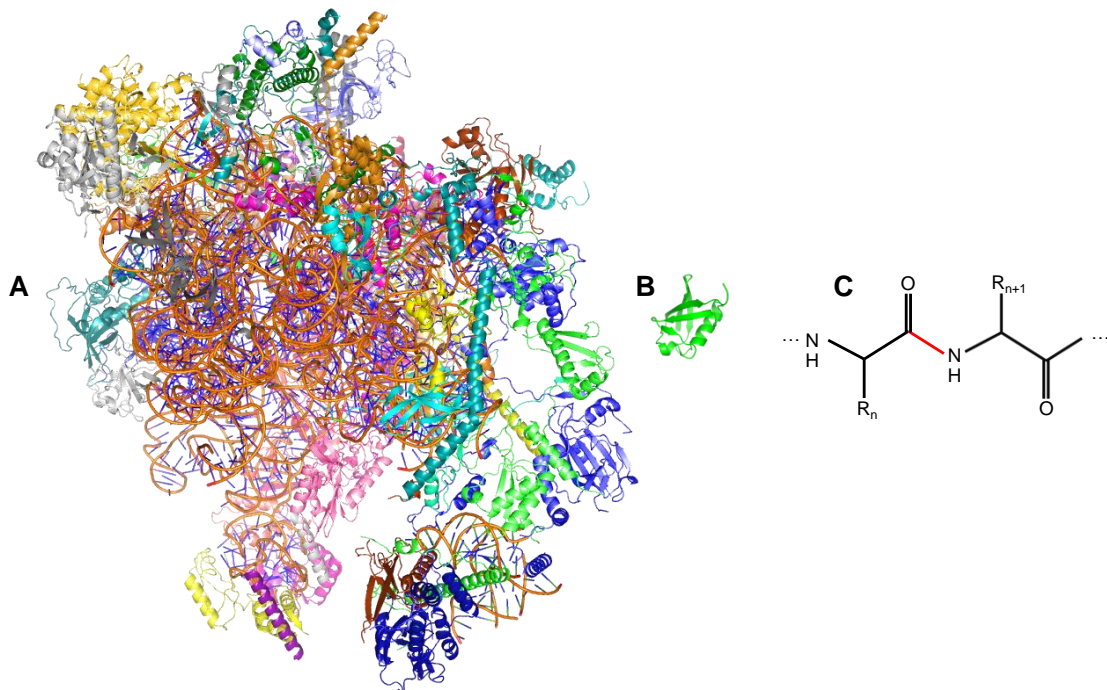


Figure 1.1 X-ray structures of (A) the 2.8 MDa human mitochondrial ribosome (PDB ID: 3J9M), and (B) the 8.5 kDa human ubiquitin (PDB ID: 1UBQ). The ribosome is a ribonucleoprotein complex. The structures are depicted to scale. (C) Structure of a two-residue polypeptide segment. The peptide bond is drawn in red.

Proteins exhibit astonishing diversity in terms of their biological function. Tasks such as cellular respiration, enzyme catalysis, signal transduction, protein synthesis and degradation are performed by proteins (1). Despite their structural and functional diversity, all proteins are made up of the same ubiquitous 20 naturally occurring amino acids (Figure 1.1C). The protein structure-function paradigm holds that proteins need to *fold* into a specific three-dimensional structure to carry out their functions. How proteins are able to find this structure, called the native state (N), remains an active area of research (2).

### *1.1.2 Protein Structural Hierarchy*

The native structure of proteins can be split into four categories. The primary structure of the protein refers to the amino acid sequence of the polypeptide chain (Figure 1.1C). The amide and carbonyl groups are the same for all amino acids. Therefore the side chains (R groups) embody key properties such as size, electric charge, and hydrophobicity (1). The many amino acids comprising a protein are connected via peptide bonds (Figure 1.1C - highlighted in red), which exhibits partial double bond character. Except for the peptide bond, which almost always adopts a trans configuration, free rotation around other bonds along the protein backbone paves the way for the existence of countless three-dimensional structures. In addition to the choice of amino acids, the cellular machinery is armed with post translational modifications (PTMs) to further regulate structure and function (3). Secondary structure refers to local structural arrangements such as  $\alpha$ -helices,  $\beta$ -sheets, etc. (1). Tertiary structure of proteins defines the association of the secondary structural elements and the three-dimensional shape of the protein. Within protein

complexes, quaternary structure determines the overall geometry and arrangement of individual subunits relative to one another within protein complexes (1).

### *1.1.3 Factors Contributing to Protein Stability*

What are the driving forces behind protein folding? The second law of thermodynamics posits that the native state of protein is the conformation with the lowest free energy (4). At first glance, protein folding might seem to contradict this because conversion of an unfolded chain to a compact structure is concomitant with a significant loss of conformational entropy. The following factors contribute to the stability of the folded state of proteins.

The hydrophobic effect and interactions are widely considered to be the most important factor contributing to the stability of native proteins (5, 6) Six of the twenty naturally occurring amino acids are highly hydrophobic, with another eight having intermediate hydrophobicity. Globular proteins form tightly packed structures, within the cores of which the hydrophobic amino acids side chains are buried and sequestered from the surrounding water. By contrast, amino acids with charged and polar side chains are often found on the protein's surface where they interact with the surrounding water. Theoretical studies suggest that hydrophobic forces alone are adequate to cause the folding of proteins with *ca.* 200 amino acids (7).

Hydrogen bonding is the hallmark of the secondary structural elements of proteins (8). In  $\alpha$ -helices the amide hydrogen (N-H) is hydrogen bonded to the oxygen of the carbonyl group of the amino acid four residues earlier.  $\beta$ -sheets also participate in complex networks of hydrogen bonds. Some question the importance of hydrogen



bonding because these moieties can form hydrogen bonds with water even in the unfolded state (9).

van der Waals interactions or “self-solvation” is often cited as an important factor (10). The tight packing of atoms within the compact structure of folded protein implies the importance of close-range interactions. These include forces between pairs of permanent and induced dipoles.

Salt bridges originate from electrostatic interactions between the negatively charged carboxylate ( $\text{RCOO}^-$ ) of Asp and Glu residues and the cationic side chains of Lys and Arg (11). Other residues with ionizable side chains such as His can also participate, depending on solution pH and their pKa. Salt bridges on the surface of proteins do not contribute much to stability. This is because they are heavily solvated by water. Shielding of salt bridges from water and burying them in the interior is energetically unfavorable. These factors severely limit the ability of salt bridges to stabilize the native state. Some proposals even outline a destabilizing role for salt bridges (12).

Conformational (chain) entropy is by far the strongest force opposing the formation of the folded protein (13). Protein folding is concomitant with a large loss in chain entropy as the polypeptide coils up into its compact native state from its many open denatured configurations.

#### *1.1.4 Protein Folding Mechanisms*

In a series of elaborate experiments that culminated in the 1972 Nobel Prize, Christian Anfinsen showed that the native structure is determined only by the protein's amino acid

sequence (14). The Anfinsen dogma postulates that the native structure is a unique, stable and kinetically accessible structure that represents the minimum free energy of the protein and its surrounding solvent. This finding gave birth to protein folding research.

The astounding speed of protein folding has been of special interest since the early days. Using conservative estimates, a relatively small protein with 100 amino acids can fold in  $\sim 10^{150}$  different ways. What is the mechanism via which disordered polypeptide chains avoid the numerous unproductive avenues and find their way to the native states (15, 16)? Do proteins fold by pre-determined “folding pathways” (17)? In 1969 Cyrus Levinthal pointed out that protein folding does not take place based on a trial and error conformational search (18). Even in the unrealistic scenario where a protein could try one different conformation every  $10^{-50}$  seconds, folding would still take  $10^{100}$  seconds. This value is 80 orders of magnitude longer than the age of the universe and contradicts the experimentally observed ms-s folding times. This disparity in the experimental and theoretical folding times represents one facet of the so-called “protein folding problem”.

Statistical thermodynamics work from the Dill (19) and Wolynes (20) groups in the late 80s addressed Levinthal’s paradox by introducing the energy landscapes concept. Hyperdimensional free energy maps were conceptualized to display the energetics preferences of polymer chains. For foldable polymer models, it was shown that a small number of compact, low-energy conformational ensembles populate the bottom of funnel-shaped energy landscapes. By contrast, there are numerous high energy unfolded conformers that correspond to the rim of the funnel (Figure 1.2) (19). Protein folding proceeds via conformational diffusion of the chain on these funneled surfaces. A biased

conformational search ensures that structures with increasingly lower free energies are adopted. Calculations have shown that steps favourable by as little as  $k_B T$  ensure rapid folding times, consistent with experimental observations (21).

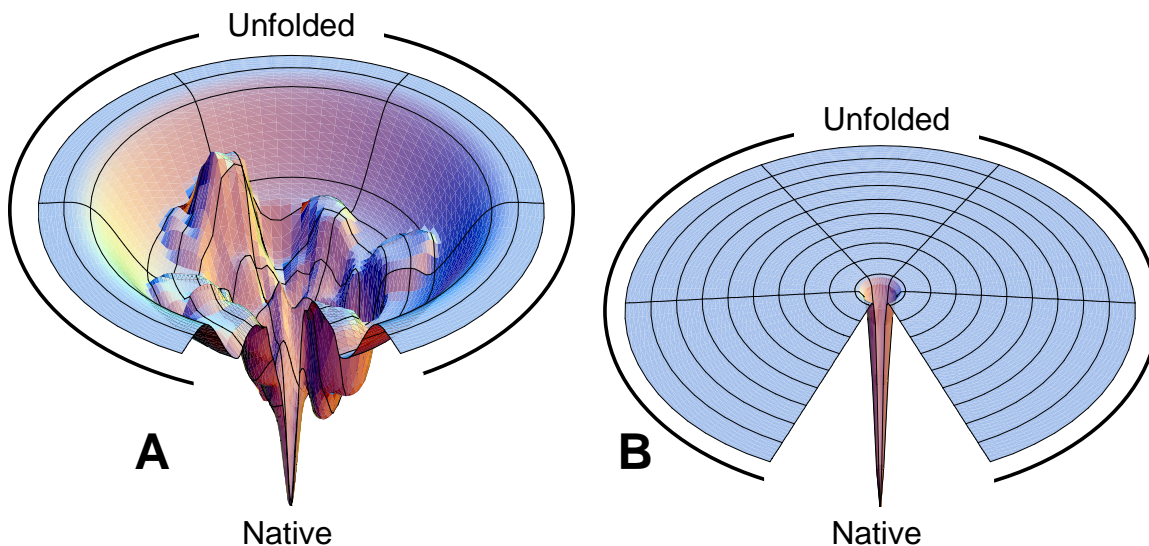


Figure 1.2 Protein folding landscapes. (A) Free energy landscape with multiple folding pathways with a surface biased towards the native state (N). The ruggedness of the funnel leads to the formation of transient protein folding intermediates (B) Folding on the Levinthal “golf course” entails an unbiased and random search for the native state. Images taken from <http://dillgroup.stonybrook.edu> and reproduced under a Creative Commons Attribution 4.0 International License.

While this framework addressed critical questions such as the speed of folding, a molecular picture of how the disordered protein transitions into the native state was lacking. A popular approach to remedy this is characterizing the properties of the transient folding intermediates (local minima on the bumpy funnel) which are populated during the folding process (22, 23). This approach relies on the tenet that snapshots of the chain during folding would give insights into a general rule-book based on which proteins fold. The combination of hydrogen/deuterium exchange and quench flow mixing

(see Section 1.4.1) has proven to be highly effective (23). In the case of some proteins, conformers that highly resemble those found in kinetic experiments can be populated under semi-denaturing solution composition (24). Once the similarity of the structures is validated, this approach allows detailed structural characterization under equilibrium conditions (25).

Several protein folding models have been proposed that attempt to capture the structural transitions that ultimately yield the native state (26). The diffusion-collision model divides the protein into several microdomains small enough for all conformations to be searched through rapidly (27). Coalescence of these unstable microdomains yields the native state. The framework model assigns a dominant role to local forces (28). This model predicts that the formation of secondary structure elements precedes chain collapse and tertiary interactions. The nucleation-condensation model proposes that the folding process initiates from a nucleus that consists primarily of adjacent residues (22). Folding proceeds because this nucleus is not stable unless it makes tertiary interactions with the surrounding residues. Secondary and tertiary native state contacts grow concurrently as the nucleus grows rapidly. Fersht and coworkers have proposed a unifying folding mechanism that envisages the framework and nucleation-condensation models to be different manifestations of an underlying common mechanism (29). Proteins may appear to fold via either model depending on the stability of folding intermediates and the height of the transition barriers. The hydrophobic collapse model posits that hydrophobic forces result in polypeptide chain collapse early in the folding process and long-range interactions precede or are concomitant with the formation of local contacts such as secondary structure elements (30, 31).

### *1.1.5 Protein Dynamics*

The advent of high resolution structural methods over the past half-century have led to a great many atomic-resolution models of proteins essential to life. Although studies of static three-dimensional structures determined via crystallography and other methods are crucial (see Section 1.2), proteins in solution are highly dynamic. It is becoming increasingly clear that a complete description of biological function must take into account conformational motions that can manifest over a broad range of time and length scales. Hence, the structure-function paradigm has to be extended to include dynamics (32, 33). This is akin to the situation where a few photographs snapped during a hockey game do not reflect the entire process of scoring goals. Proteins undergo a variety of conformational changes that enable them to function as catalysts (34), ion transporters (35), signaling switches (36), etc. Homologous protein structures exhibit disparate dynamics, spanning order of magnitude in time, thereby underscoring the importance of this aspect (37).

The concept of energy landscapes can also be applied within the context of the native protein. Although the energy well corresponding to the native state in Figure 1.2 looks deep, ruggedness within this energy well gives rise to substates that are separated by barriers with varying heights. Similar to protein folding, the energy difference between the states and the height of the barriers separating them gives rise to protein dynamics on various time scales (38). Slow dynamics correspond to  $\mu\text{s}$ - $\text{ms}$  timescale interconversion between substates separated by energy barriers of several  $k_{\text{B}}T$ . Ligand/substrate binding and allosteric events often manifest themselves over

milliseconds. Proteins are not static within these substates. They undergo numerous small-amplitude fluctuations around the average structure on the ps-ns timescale. The energy to cross the barrier for these faster dynamics is readily available under ambient conditions as the interconverting states are separated by less than  $1 k_B T$ . This gives rise to loop motions and sidechain rotations.

### *1.1.6 Protein Misfolding*

Studying the structural and motional properties of proteins goes beyond a curiosity-driven scientific endeavour. Only proteins that fold properly possess the stability to function robustly in the cellular environment. Close to 50 disorders such as Alzheimer's disease, spongiform encephalopathies, Parkinson's disease, Senile systemic amyloidosis, and type II diabetes have been linked with the misfolding of normally soluble, peptides and proteins, and their subsequent conversion into insoluble toxic aggregates (39). While protein folding started as curious scientific discovery, contemporary protein research is motivated by this biomedical aspect and its implications for human health (40). Understanding of the fundamentals that govern biomolecular structure, folding, and interactions will aid in the development of therapeutics.

## **1.2 Experimental Methods to Study Protein Structural and Motional Properties**

### *1.2.1 X-ray Crystallography*

X-ray crystallography is the foremost technique for determining the atomic structure of proteins. The importance of this method cannot be overstated: crystallographic projects have earned an astounding 28 Nobel Prizes (41). Kendrew and coworkers resolved the

first protein structure in 1958 (42). The Protein Data Bank (PDB) currently contains more than 100,000 structures, of which *ca.* 90% were determined by X-ray crystallography.

In the beginning of the 20<sup>th</sup> century, Max von Laue and the Braggs worked out much of the theory behind the interaction of X-ray beams and material (43). X-rays that pass through a crystal scatter off the electrons in the crystal lattice and then interfere with each other. In some instances, the electromagnetic waves undergo constructive interference; in others, they cancel each other out. The result is a diffraction pattern from which the electron density that scattered the original beam can be calculated. This is followed by model building and structure refinement. The average structure resolution in the PDB is around 2 Å. Distinguishing individual atoms within the sample is possible for resolutions of *ca.* 1 Å (41). Resolution suffers in images of complex samples due to imperfections and incoherent motions within crystals (43, 44).

X-ray crystallography requires several milligrams of relatively pure material. Some proteins, in particular those containing dynamic regions, are not amenable to X-ray diffraction methods because they do not form crystals. The excision of such domains gives truncated constructs that sometimes form crystals more readily. However, this strategy may perturb the protein structure. Another source of artefacts comes from within the crystals. Protein-protein contacts in the tightly packed environment of crystals have been shown to differ from those in solution (45). Crystallography of membrane proteins continues to lag far behind that of soluble targets (41). X-ray data provide snapshots of the protein ground state, and generally do not report on dynamics. An interesting development has been synergistic utilization of NMR-based order parameters and X-ray models to obtain complementary information regarding both structure and dynamics (38).

Some of the limitations mentioned are expected to be overcome with the more widespread implementation of X-ray free electron lasers (XFEL) (46, 47). Some early signs of success have already emerged. A notable example is the determination of the photosystem II structure at 1.95 Å using serial femtosecond crystallography (48, 49).

### *1.2.2 Cryogenic Electron Microscopy (Cryo-EM)*

Cryo-EM has been revolutionized over the past decade and holds great potential as a method for the structure determination of large biomolecular assemblies (50). The advent of electron guns with enhanced brightness and coherence, along with advances in specimen handling and processing methods have improved resolution from several nanometers to the sub-nanometer regime (51). The biggest advantage of cryo-EM over X-ray crystallography is that structural information is obtained from small quantities of noncrystalline material. It is best suited for probing large (>200-300 kDa), rigid protein complexes that possess some form of symmetry. Several studies have revealed the architecture of systems that have long been intractable by X-ray crystallography. These include 6.9 Å structure of V-type ATP synthase from *Saccharomyces cerevisiae* (52), 6.2 Å structure of the  $\alpha$ -subunit of F-type ATPase, (53) and 11 Å structure of chromatin fibers (54). A hybrid approach that is being increasingly applied entails fitting atomic-level models of domains and smaller fragments obtained from other methods into cryo-EM density maps. This methodology marries the best attributes of the two methods and has proven to be highly valuable for studying a number of challenging targets. A pair of ground-breaking studies from the Baumeister (55) and the Martin (56) groups focusing on the yeast 26S proteasome highlight the potential of this methodology.



Despite these notable examples, studying smaller and flexible targets by cryo-EM remains challenging. Degradation of the structural integrity of the specimen during analysis is one of the biggest hurdles the method faces. Similar to X-ray crystallography, information about protein conformational dynamics are currently unavailable via cryo-EM. Continued improvement in the areas described above promises that cryo-EM will play an increasingly prominent role in structural biology (57).

### *1.2.3 Nuclear Magnetic Resonance (NMR) Spectroscopy*

NMR spectroscopy is uniquely suited for characterizing biomolecular structure and dynamics (58). NMR employs exploits high magnetic fields and radio frequency to probe NMR active nuclei such as  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  within a macromolecule as “spies” of protein structure and motion. Unlike X-ray crystallography (Section 1.2.1) and cryo-EM (Section 1.2.2), NMR offers the advantage of studying biomolecules under physiological conditions (59) or under favorable conditions in living cells (60).

NMR has played a significant role in protein folding research as the detection method for hydrogen/deuterium exchange measurement of pulsed-labeled kinetic protein folding intermediates (23). The development of three-dimensional triple-resonance experiments (61) and transverse relaxation optimized spectroscopy (TROSY) experiments (62) mitigated spectral overlap and enabled studying of large proteins. The advent of Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion (63) and chemical exchange dependent saturation transfer (CEST) (64) experiments have enabled detailed characterization of energetic of the excited states of protein that are weakly populated under native solution phase conditions. Combinations of these technologies with

advanced computational modeling produced atomically-resolved models of folding intermediates that have shaped our understanding of protein folding and misfolding (65, 66).

Wüthrich and coworkers (Nobel Prize Laureate 2002) developed the Nuclear Overhauser effect spectroscopy (NOESY) experiments that yield inter-nuclear distance constraints and 3D structures of proteins up to *ca.* 50 kDa (67). In addition, an extensive array of NMR experiments has been developed that quantify biomolecular dynamics over a wide range of timescales, from picoseconds to days.  $R_{1\rho}$ , ZZ exchange, and relaxation dispersion experiments are suited to probe the  $\mu$ s-ms regime.  $R_1$ ,  $R_2$  pulse sequences are used to investigate faster dynamic events. Studying the dynamics of large protein complexes was recently made possible via the application of methyl-TROSY experiments and isotopic labeling of methyl-containing side chains (68). High quality  $^{13}\text{C}$ - $^1\text{H}$  correlation maps for all six methyl-containing amino acids (Ile, Leu, Val, Ala, Thr, and Met) of systems with masses exceeding 1 MDa have been recorded (69). Notable examples include the interrogation of the archaeal proteasome in complex with various activators (70) and activation mechanism of the Hsp70 molecular chaperones (71).

Despite these advances, NMR still struggles with large and asymmetric multi-subunit complexes. Spectra from large asymmetric systems suffer from extensive overlap. *in vitro* reconstitution of the complex from separately expressed NMR active and NMR inactive subunits can be applied in a limited number of cases. First steps to remedy these shortcomings have already been taken (72). The ever-increasing magnetic field strengths, more elaborate sequential protein co-expression methodologies and NMR pulse sequences should push these limits further.

#### *1.2.4 Optical spectroscopy*

The vast majority of early protein folding experiments that elucidated the kinetics of the folding process were performed using optical methods. These include UV-Vis absorbance, far and near UV-Vis, circular dichroism (CD), fluorescence and Förster resonance energy transfer (FRET) spectroscopy, infrared (IR) and Raman spectroscopy, among other methods.

##### *1.2.4.1 Circular dichroism (CD) spectroscopy*

Far-UV (190-260 nm) and near-UV (250-350 nm) CD spectroscopy provide a measure of average secondary and tertiary structure content of proteins, respectively. The combination of this methodology with stopped-flow mixing and rapid-mixing devices has been a valuable source of information about the kinetics of protein folding (73).

Early stopped-flow data on model proteins showed that a substantial fraction of the overall change in the far-UV CD signal resulting from refolding falls within the instrumental dead time of approximately 1 ms. Meanwhile, the near-UV spectroscopic signature did not reach native levels until the later stages of the folding process (74). These results were generally interpreted as formation of intermediates with substantial secondary structure but lacking native-like tertiary structure packing. Later on the advent of faster turbulent- and continuous-mixing methods (75) and temperature-jump triggering helped resolve the kinetics down to the ns/ $\mu$ s time scale (76). However, the details of structural reorganization corresponding to the early folding events remained nebulous for many years and continue to be a hotly debated topic (77).

#### *1.2.4.2 Fluorescence*

The fluorescence emission of Trp, and to a lesser extent Tyr, side chains can serve as a valuable reporter of the local environment of these moieties (73). Although complex photophysical phenomena complicate data interpretation, generally a solvent-exposed tryptophan in the unfolded state of a protein shows a broad emission spectrum similar to that of free tryptophan or its derivative N-acetyl-L-tryptophanamide (NATA) ( $\lambda_{\max}$  350 nm). Shielding of the tryptophan side chain within the hydrophobic core of native proteins or in a compact folding intermediate typically results in a major blue shift of the emission maximum and enhanced fluorescence yield. Within the native state of proteins, amino acid side chains of Lys, Tyr, Gln, Asn, Glu, Asp, Cys, and His quench fluorescence by one of intersystem crossing, solvent quenching, excited-state proton and electron transfer mechanisms (78). The presence of multiple fluorophores can be a source of complication. Site-directed mutagenesis can be used to remove multiple fluorophores. A Trp-scanning strategy has been devised to characterize the structure of transient folding intermediates (79).

Similar to CD spectroscopy, monitoring fluorescence intensity as a function of folding or unfolding time can provide useful information on solvent accessibility and proximity to quenchers of an individual fluorescence probe. It serves as an exquisite method to study early folding events. Classic examples include misfolding of cytochrome *c* due to initial misligation of the heme iron by His33 (80), and implications of proline isomerization for the folding of ribonuclease A (81).

#### *1.2.4.3 Förster resonant energy transfer (FRET)*

In FRET a donor fluorescent dye is excited by a light source. This energy can subsequently be transferred nonradiatively to an acceptor dye. One of the requirements of FRET is that measurements should be performed between a pair of fluorophores where the emission of the donor dye overlaps with the absorption of the acceptor. As the two molecules come closer together, the donor emission decreases in intensity while the acceptor emission increases. FRET efficiency depends on the donor-to-acceptor distance with an inverse 6th power, and is thus referred to as "molecular ruler". It is most sensitive to measurement between 2-8 nm range (73).

Both fluorescence and FRET are valuable methods for characterizing the conformation of gas phase molecules (82, 83). Comparison of solution phase data to those obtained in the gas phase can reveal the nature of structural changes of gas phase proteins. The first steps towards this direction have been taken (84).

#### *1.2.4.4 Single-Molecule Methods*

Single-molecule experiments report on the motions of individual protein molecules and help to understand how this translates into an ensemble signal reported by other methods. The applications of single-molecule fluorescence spectroscopy and FRET have been particularly useful in this regard (85, 86). This approach was used to measure the length of protein folding barrier-crossing events (87). Perhaps one of the most mesmerizing applications of this methodology was direct observation of the operation of  $F_1$  ATPase rotary molecular machine. A fluorescent actin filament was attached to the  $\gamma$  subunit of the enzyme tethered to a surface. Supplying the enzyme with ATPase leads to

the rotation of the motion of “shaft” and the attached filament. This motion can be “seen” directly using a fluorescence microscope (88). FRET-based measurements with a donor on the  $\gamma$  subunit and an acceptor on the stationary peripheral stalk have revealed a great deal regarding the torque, speed, force generation, efficiency of this system (89). The rapid nature and specificity of these optical methods allow observation of individual turnover events. This technology has also enabled temporal mapping of protein energy landscape *in vivo* (90). Single-molecule force spectroscopy techniques such optical trap force spectroscopy and atomic force microscopy monitor the response of proteins to force (91). Recent examples utilize this technology to unravel the mechanism of action of chaperones on substrate proteins (92), and the importance of nascent protein folding (93).

### *1.2.5 Computational Methods*

Despite the success of experimental techniques in providing information on structure and dynamics of proteins, they are generally limited in their spatial and temporal resolution. With the exception of single-molecule measurements (Section 1.2.4), most methods report ensemble-average properties rather than individual molecules. An alternative is computational modeling of biomolecules.

All-atom molecular dynamics (MD) simulation have become the method of choice for simulating the motions of biomolecules (94). In MD simulations the positions and velocities of every atom in the system is evolved according to first-principles laws of physics. The forces acting upon particles representing atoms are calculated using a model known as a force field (95). Dramatic improvements in computation power and parallel computing pioneered by Wall Street “King Quant” turned self-taught computational

biochemist, David E. Shaw, have pushed the length of these simulation onto the millisecond time scale (96). The combination of these have shed new light on protein folding (97-99), drug binding (100, 101), protein design (102), and the conformational changes critical to enzyme function (103). The faithfulness of this methodology promises to increase as accuracy of force fields improves (104).

### 1.3 Fundamentals of Mass Spectrometry

The origins of mass spectrometry (MS) go back to Sir Joseph Thomson more than a century ago (105). Today MS is among the most widely used analytical techniques. An ion source and a mass analyzer are the two main elements of each mass spectrometer. A mass spectrometer reports the mass-to-charge ratio ( $m/z$ ) of gaseous ions produced by the ion source.

#### *1.3.1 Electrospray Ionization*

MS of biological macromolecules (e.g. proteins and DNA) was not possible prior to the late 1980s. Electron impact (EI) and other traditional ionization methods were ineffective for these analytes (106). Two ground-breaking inventions that culminated in a Nobel Prize completely changed this picture. John Fenn and Koichi Tanaka shared the 2002 Chemistry Nobel Prize for inventing electrospray ionization (ESI) (107) and matrix-assisted laser desorption/ionization (MALDI) (108), respectively. More than two decades later, a suite of “soft” ionization methods such as desorption electrospray ionization (DESI) (109) and matrix assisted ionization vacuum (MAIV) (110) have been developed. These allow for a wide variety of analytes to be transferred into the gas phase and ionized

without the breakage of any covalent bonds. ESI boasts the advantage of facile online coupling with liquid chromatography (LC) for separation of complex mixtures prior to MS. Because this thesis exclusively employed ESI, other ionization methods will not be discussed.

ESI generates multiply charged  $[M + zH]^{z+}$  ions. No practical size limit has been found for analytes compatible with ESI. The current record represents a 18 megadalton virus assembly (111) Multiple charging lowers the  $m/z$  of these large entities to a range where most mass spectrometers can operate. High charge states are also desirable for applications involving collision induced dissociation (CID) (112), and electron-mediated dissociation (113). The ESI entails infusion of the analyte solution through a capillary held at an electric potential of several kV. The Taylor cone that forms at the tip of this capillary emits charged droplets that undergo multiple rounds of solvent evaporation and jet fission. These yield nanodroplets that are charged close to the Rayleigh limit. The mechanism via which these nanodroplets generate gaseous ions has been a hot research topic since the inception of ESI. The current consensus (114, 115) is that low-molecular-weight analytes follow the ion evaporation model (IEM) (Figure 1.3, first column) whereby gaseous ions are ejected from the nanodroplet surface by field emission (116). By contrast, large globular species such as folded proteins follow the charge residue model (CRM) and are produced by evaporation of droplets to dryness (Figure 1.3, second column) (117). It has been proposed that disordered polymers and unfolded proteins are produced by the chain ejection model (CEM) (Figure 1.3, third column) (118).



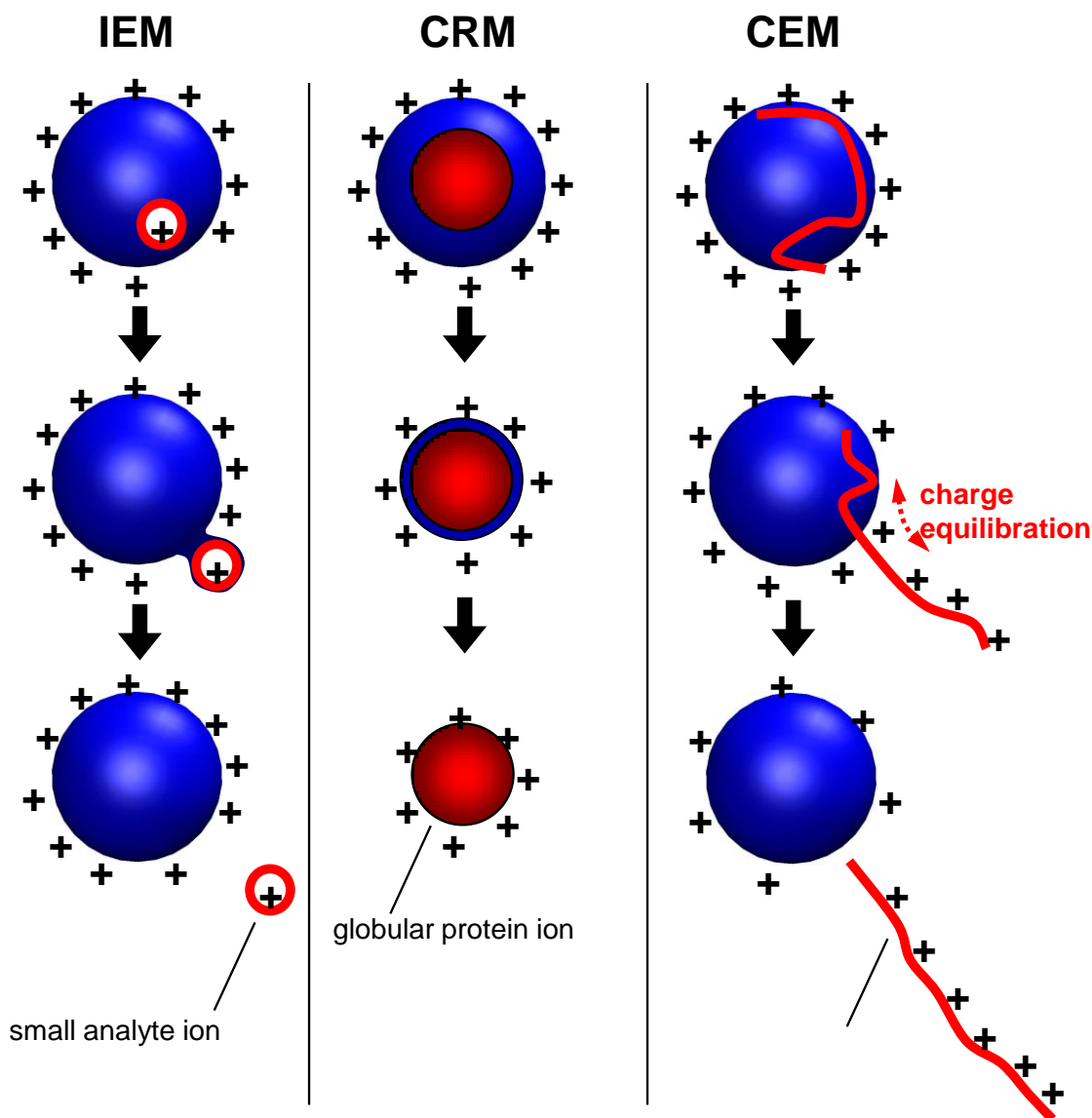


Figure 1.3 Summary of ESI mechanisms. IEM: Small ion ejection from a charged droplet. CRM: Release of a globular protein into the gas phase. CEM: Ejection of an unfolded protein chain. Charge equilibration is indicated by red arrows. Figure reproduced with permission from ©American Chemical Society (114).

### 1.3.2 Mass Analyzers

The mass analyzer is tasked with measuring the  $m/z$  and relative abundance of gaseous ions produced by the ion source. There are various types of mass analyzers including

quadrupole, time-of-flight (TOF), ion traps, Orbitraps, and Fourier transform ion cyclotron resonance (FT-ICR) mass analyzers. Because this work exclusively employed quadrupole-TOF instruments, only these two mass analyzers will be discussed in detail.

### *1.3.2.1 Quadrupole Mass Analyzer*

A quadrupole mass analyzer is composed of four parallel cylindrical rods. A radio frequency (RF) and a direct current (DC) are applied to opposing rod pairs. For a combination of RF and DC voltages, only ions with a certain  $m/z$  have the proper trajectory to traverse the quadrupole and reach the detector. By contrast, ions with unstable trajectories undergo collisions with the rods and get neutralized. The stability of an ion trajectory for any given RF and DC voltage is governed by the Mathieu equation. The RF and DC voltages can be scanned such that the desired  $m/z$  range is covered. In the absence of a DC voltage (RF-only mode), quadrupoles act as ion guides. Thus quadrupoles are routinely used as versatile ion transmission and selection devices for MS/MS applications (106).

Quadrupoles generally offer a limited resolution ( $R \approx 5000$ ) and limited  $m/z$  range. But modern devices can resolve up to 32,000  $m/z$  and transmit ions up to 100,000  $m/z$  in RF-only mode. Despite their limited resolution and slow scanning times, quadrupoles are very robust instruments and have remained a fixture in the pharmaceutical and environmental labs (119). Triple-quadrupole instruments excel at targeted quantitative workflows such as selected reaction monitoring (120).

### 1.3.2.2 Time of Flight (TOF) Mass Analyzer

A TOF mass analyzer separates ions based on the time it takes them to traverse a flight tube. An ion pusher accelerates an ion with a mass to charge ratio of  $m/z$  in a flight tube by application of a voltage,  $U$ . This leads to an ion velocity  $v$  according to

$$zU = \frac{1}{2}mv^2 \quad 1.1$$

Rearrangement yields

$$v = \left(\frac{2zU}{m}\right)^{1/2} \quad 1.2$$

The time  $t$  it takes an ion to traverse a flight tube with length  $l$  is given by

$$t = \frac{l}{v} = l\left(\frac{m}{2zU}\right)^{1/2} = l(2U)^{-1/2}(m/z)^{1/2} \quad 1.3$$

Equation 1.1 dictates that time of flight depends on the  $m/z$  of each ion. Ions with different  $m/z$  have different drift times and hence are separated in the flight tube. Ions with lower  $m/z$  will reach the detector first followed by their “heavier” counterparts. Modern mass spectrometers use a time-to-digital device to convert flight times to  $m/z$ . Reflectrons in orthogonal flight tubes are utilized to correct for energy dispersions in the original ion packet (119).

The reflectron uses a constant electrostatic field to reverse the trajectory of ion beam toward the detector. The more energetic ions penetrate deeper into the reflectron and dwell longer before their trajectory is reflected towards the detector. The ion of the same  $m/z$  and with lower initial energy will penetrate the reflectron later but will spend less time changing direction within reflectron. This mechanism corrects for differences in initial ion energy and ensures that ions with same  $m/z$  but slightly different in kinetic

energy reach the detector at the same time and record the same flight time. Without a reflectron, the TOF resolution is no better than  $R \approx 5000$ . Modern orthogonal TOF devices offer superb sensitivity at  $R \approx 40,000$  and sub ppm mass accuracy. The combination of the quadrupole and TOF yields one of the most popular instrument designs for biomolecular analysis known as Q-TOF (Figure 1.4).

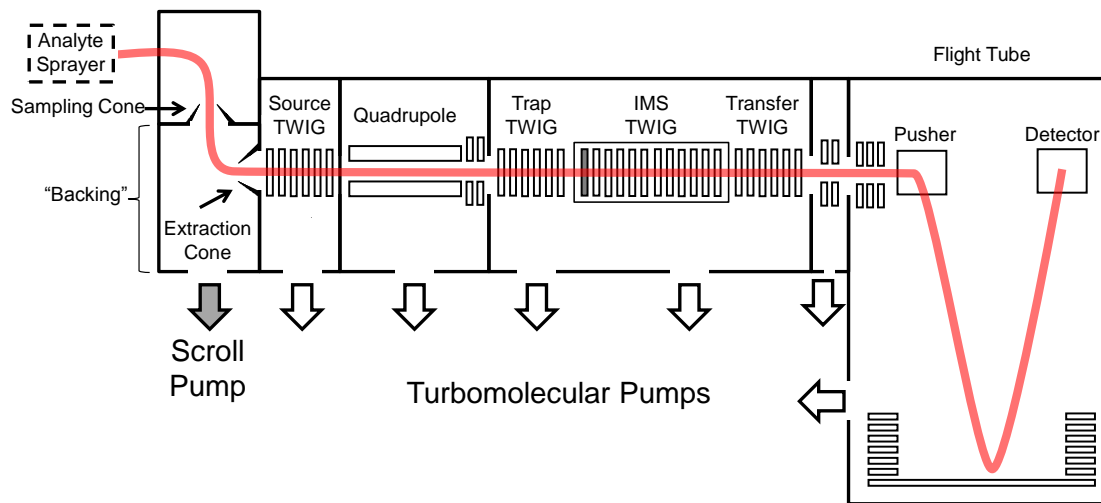


Figure 1.4 Schematic of the Q-TOF instrument used in this thesis with components marked. The red line indicates the ion trajectory.

## 1.4 Structural Mass Spectrometry of Proteins

At first glance the choice of MS as a tool for structural interrogation of proteins might seem peculiar. Protein structural changes generally do not result in any change their molecular weight, implying no change in the principle property measured by the mass spectrometer,  $m/z$ . In addition, the vacuum of the mass spectrometer seems like an unsuitable environment for proteins which have evolved to function in an aqueous environment. Unlike NMR, MS was not initially developed as a structural biology tool. The methodologies discussed in the following sections illustrate how the MS community

has collectively defied all odds in building a vast arsenal of techniques to report on protein structure, folding and dynamics (121-123).

#### *1.4.1 Hydrogen/Deuterium Exchange (HDX)*

The concept of HDX was introduced more than six decades ago by Linderstrøm-Lang (124). HDX exploits the incessant exchange of labile hydrogens within the protein with those of the solvent. Labile sites include the amide backbone hydrogen and hydrogens with sufficient acidity on the amino acid side chains. Disordered and solvent-exposed segments of the protein readily exchange their hydrogens upon exposure to D<sub>2</sub>O. By contrast, this process is orders of magnitude slower for regions involved in hydrogen-bonding networks (e.g.  $\alpha$ -helices and  $\beta$ -sheets) or those occluded from the solvent (125). Measurement in the early days used scintillation counting (126) and UV-spectrophotometry (127). HDX enjoyed great success when NMR-based methods were employed (128). MS-based measurement were first introduced in the early 1990s (129, 130). HDX is an indispensable tool for monitoring the structure and dynamics of proteins (131, 132).

Mass spectrometry-based measurements exploit the mass difference between hydrogen (H) and deuterium (D). In the typical bottom-up experiments, the exchange process is initiated by diluting the protein into a D<sub>2</sub>O-based buffer. HDX is then allowed to proceed for different incubation time periods (continuous-labeling). Subsequently, exchange is quenched by lowering the pH and temperature. This is followed by digestion by an acid-active protease (e.g. pepsin), and LC-MS analysis. HDX and other labeling methods discussed in Sections 1.4.2 and 1.4.3 are immune from complications that could

arise from analysis of proteins in the gas phase. MS simply serves as a sensitive tool for monitoring the amide deuteration pattern imprinted on the protein backbone in solution.

The elegance of HDX lies in the benign nature of the label and its relatively simple workflow. One of the biggest hurdles experimentalists have to deal with is “back exchange” during post-labeling sample handling. This spurious effect tends to degrade the structurally informative deuteration pattern (133). The half-life ( $t_{1/2}$ ) of an unstructured typical amide at pH 7.0, 25 °C is under 1 second. This imposes limitation on all down-stream sample handling protocols. LC-MS separation and digestions are required to be carried out near 0 °C and at pH ~2.7 where  $t_{1/2}$  is on the order of ~1 hour, thereby providing enough time to carry out the MS measurement. Short LC gradients on the order of 10-15 minutes are often employed. Despite these efforts back exchange is not completely suppressed. The side chain deuterium is entirely lost due to back exchange, and backbone amides’ deuterium content can be measured with a 30-40% back exchange. The use of anti-freeze compounds in LC solvents enables sub-zero LC and improved deuterium retention (134, 135) The application of rapid capillary electrophoresis separation is a new method to shorten the peptide separation step (136).

The spatial resolution of bottom-up HDX-MS depends on the length of the generated peptides, typically around 10 residues. Middle-down and top-down methodologies employing electron capture dissociation (ECD) or electron transfer dissociation (ETD) have been described that enable measurements down to the single-residue level (137). The application of ultraviolet-photodissociation for single-residue HDX measurements is an exciting but unexplored area (138). An alternative to the approaches described is subtractive method in cases where the proteolytic digestions

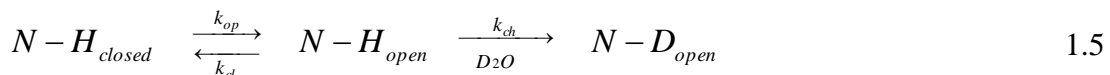
generates a great many overlapping peptides (139-142) However, occurrence of differential back exchange for overlapping peptides and error propagation complicate its application (143).

Backbone amide HDX can proceed via acid- or base-catalyzed mechanisms (131). The intrinsic rate of exchange ( $k_{ch}$ ) for a completely unstructured polypeptide can be described according to:

$$k_{ch} = k_{acid}[H^+] + k_{base}[OH^-] \quad 1.4$$

where  $k_{acid}$  and  $k_{base}$  are the rate constants for the acid- and base-catalyzed reactions, respectively. Base catalysis is dominant under physiological conditions suitable for studying proteins.

The currently accepted model for HDX divides amide hydrogens into two categories, “open” and “closed”, in dynamic equilibrium with each other with rates of interconversion  $k_{op}$  and  $k_{cl}$  according to:



This model posits that exchange can only occur for “open” hydrogen amides. These correspond to solvent-accessible amide hydrogens that have become disengaged from hydrogen bonding due to the conformational fluctuations described in Section 1.1.5. By contrast, “closed” amide hydrogens are exchange-incompetent because they are tightly hydrogen-bonded or not within the reach of the solvent. Once the “open” state is populated, deuteration proceeds with the intrinsic rate constant  $k_{ch}$ . Kinetic treatment of eq 1.5 yields:

$$[N - H] = \exp(-k_{HDX} t) \quad 1.6$$

where [N-H] refers to the concentration of a single undeuterated amide after exposure time  $t$  and

$$k_{HDX} = \frac{k_{op} k_{ch}}{k_{op} + k_{cl} + k_{ch}} \quad 1.7$$

Stably hydrogen-bonded amides with transient opening events typical of thermal fluctuations give rise to EX2 kinetic ( $k_{cl} \gg k_{ch}$ ). In this scenario eq 1.7 simplifies to

$$k_{HDX} = \frac{k_{op}}{k_{cl}} k_{ch} \quad 1.8$$

Under these conditions a large number of opening events is required for backbone amides to exchange. This results in isotopic distributions that shift to higher  $m/z$  over time as the protein undergoes more and more opening events. By contrast, long-lived global unfolding events lead to the less common EX1 kinetics ( $k_{cl} \ll k_{ch}$ ) that manifest as bimodal isotopic envelopes.

The model described above invokes a combination of hydrogen-bonding and solvent accessibility as the main determinants of exchange events (125). A few studies employing MD simulations have appeared in the literature that attempt to improve upon this rather simplistic model (144). Deeper insights into the exact nature of the conformational excursions that give rise to HDX should become available as the time scale accessible by MD simulations approaches experimental HDX timescale (145).

Although HDX does not routinely deliver the quantitative information that NMR provides, it is an excellent probe of protein structure and dynamics (131, 132). Comparative studies of protein-protein and protein-ligand interactions comprise the bulk



of studies found in the literature (146). Here the deuteration pattern of the protein of interest is compared in the bound and unbound forms. Binding often results a pronounced reduction in deuterium uptake in the vicinity of the binding site (131, 147), although other scenarios are also possible (148). This methodology has been applied to map the interaction of agonists with receptor proteins. Another area of significant interest is HDX of proteins implicated in neurodegeneration such as the amyloid  $\beta$  peptide (149),  $\alpha$ -synuclein (150), prion protein (151). HDX is also starting to gain a foothold in the pharmaceutical industry for assessing the structural integrity of monoclonal antibodies (152) and development of therapeutics (153, 154). HDX is also an exquisite tool for monitoring light-induced conformational changes (155, 156). Despite these advances, studies of membrane proteins continue to lag far behind soluble systems (157).

The examples described above correspond to studies performed under equilibrium conditions. From eq 1.4 it is obvious that pH affects the labeling rate (158, 159). This property is exploited in pulsed-labeling studies of kinetic events such as protein folding. After triggering folding in a rapid mixing device, the protein is *pulsed* for a brief time period (often a few ms at pD 8-10), followed by quenching and LC-MS. The discriminating labeling HDX rate under these conditions ( $10^3 \text{ s}^{-1}$  at pH 9) provides a binary readout of folded (“closed”) and unfolded (“open”) elements as the protein folds. The folding of  $\alpha_1$ -antitrypsin to its metastable structure (160) and ribonuclease H (161), and the effect of the GroEL/ES chaperone system on the folding mechanism of proteins (162) were explored in recent publications.

In addition to the solution phase investigations described above, gas phase HDX is a valuable tool for probing the conformation of peptides (163) and proteins (164) in the gas phase.

#### *1.4.2 Covalent Labeling*

In covalent labeling irreversible chemical labeling of protein is exploited to gain insights into protein structure. Solvent-exposed regions become modified with a reactive probe. By contrast, exclusion from solvent due to the location, binding interactions or conformational changes leads to protection (165). Covalent labeling is a complementary method to HDX. While HDX reports on a mix of backbone hydrogen-bonding, the structural information extracted from covalent labeling methods report purely on solvent accessibility of protein side chains (165) Covalent labeling techniques have been successfully applied for exploring the structure and interactions of proteins for many years.

Similar to bottom-up HDX, the spatial resolution depends on the size of the generated peptides. The first steps to improve the resolution to the individual amino acid level have been taken (166). These efforts are complicated by the fact that differentially and multiply labeled peptides do not co-elute during LC. A middle-down approach has shown to be effective for test peptides, but its applicability to proteins remain to be established (167). Also, labeling of enzymatic recognition sites may alter proteolytic cleavage patterns in bottom-up studies. However, this very property can be exploited for structural analysis (168). A clever strategy is specific photodissociation of peptides

labeled with UV-absorbing labeling agents (169). Top-down quantitation of the labeling location and extent represents an intriguing possibility (170).

#### *1.4.2.1 Amino acid-specific methods*

A large number of covalent labels with a wide range of chemical specificities have been described over the past 20 years (165). Amino acid-specific labels operate based on a chemistry that targets functional groups within one or only a few types of residues. For example, carboxyl group footprinting specifically targets acidic side chains (Glu and Asp) with the incorporation of glycine ethyl ester (GEE) (171, 172). This results in a mass shift of 85 Da, or 57 Da after ester hydrolysis, that can be localized and quantified using LC-MS. Labeling of free thiol groups of surface cysteine residues is another option (173). Using a large thiol label allows for the modification to be read by SDS-PAG gels, thereby eliminating the need for expensive mass spectrometry equipment (174). Similar experiments have been described for surface Arg, Lys, His, and Trp residues (165). Interpretation the data of specific labeling experiments is straightforward because the extent of modification at each site can be taken directly as solvent accessibility. On the other hand, structural information is limited to the few sites that are reactive for the labeling agent chosen.

#### *1.4.2.2 Amino Acid-Nonspecific Probes*

Amino acid-nonspecific methods have proven to be highly successful. These label agents have more promiscuous reactivity, thereby providing structural information from a

large fraction of the protein sequence. Among these, hydroxyl radical ( $\cdot\text{OH}$ ) labeling applications are the most abundant in the literature.  $\cdot\text{OH}$  is a nonspecific labeling agent that generates oxidative modifications (typical mass shifts 16 Da) in solvent-exposed side chains (175). The reactivity of side chains towards  $\cdot\text{OH}$  spans a wide range. Sulfur-containing residues (Met, Cys) react most readily, followed by aromatic and aliphatic side chains (176). 12/20 side chains of naturally occurring amino acids have been experimentally confirmed to become modified. Extraction of structural information from the data requires advanced methods that accounts for this large variation in reactivity.

Given the nonspecific nature of the label, each protein can pick up a large number of covalent labels. Prolonged labeling times can lead to labeling-induced structural changes. Double labeling of proteins is also a concern. Multiple labeling events may modify side chains that were not initially solvent exposed in the non-labeled protein, resulting in artifactual data. Thus, controlling the length of the labeling step is important.  $\cdot\text{OH}$  can be produced by various methods, including Fenton chemistry (177), electrochemistry (178), corona discharges (179), synchrotron radiolysis of water (175),  $\gamma$ -irradiation (180), and pulsed electron beams (181).

Among these methods, FPOP (fast photochemical oxidation of proteins) provides the fastest labeling times.  $\cdot\text{OH}$  is produced by photolysis of dilute  $\text{H}_2\text{O}_2$  inside a continuous-flow capillary tube using a pulsed 248 nm excimer laser. The lifetime  $\tau$  of  $\cdot\text{OH}$  generated in this way has been claimed to be on the order of 1  $\mu\text{s}$  (182, 183). This is faster than the rate of most conformational changes of proteins. Once laminar flow consideration are taken into account (184), tuning of the laser frequency can be used to ensure that every protein molecule has only been labeled once. This ensures faithful

representation of the protein surface topology in the labeling pattern. The fast time scale of FPOP paves the way for characterizing short-lived protein conformers. Using a rapid mixing device, this was first exploited by Stocks and coworkers to gain insights into folding and unfolding mechanism of a number of proteins (185-187). The experiments were later expanded to the submillisecond timescale by employing temperature-jump triggering (188). Over pulse-HDX methods, FPOP enables labeling of proteins on a much faster timescale under neutral pH conditions.

FPOP has shown great promise in the elucidation of protein regions that are changed upon interaction with other macromolecules, ligands. It has been applied to a growing number of biological problems. Similar to HDX, comparative experiments are the most popular. Probing protein-protein interactions (189), and protein-ligand interactions (190) are now routine. Studies of membrane proteins are rare, but benefit from stable labels and the opportunity to perform extensive downstream sample clean up (191).

The original FPOP toolbox has been expanded to include sulfate radical anion (precursor persulfate) (192) and iodine radical ion (precursor iodobenzoic acid) (170). Carbene-based labeling approaches have also been developed to take advantage of the high reactivity of this moiety (193, 194). Further advancement of covalent labeling methods awaits a more user-friendly workflow, better software to improve and facilitate data analysis and a better understanding of the chemistry of the labeling reactions.

### *1.4.3 Crosslinking*

A closely related technique to covalent labeling is covalent crosslinking (XL). Unlike labeling methods (Section 1.4.2), XL utilizes *bifunctional* reactive probes. A pair of

solvent-exposed functional groups on the protein side chains that are separated by a certain distance will be coupled by the XL agent. This is followed by mass spectrometric identification of the cross-linked peptides, thereby providing low resolution information on the spatial arrangement of reactive sites within the system. A large number of hetero- and homo-bifunctional crosslinkers have been developed over the years (191, 195). Various spacer lengths can be used to explore different distance regimes. The distance constraints from XL can be used in combination with data from other methods to map the three-dimensional structure of biological system. An impressive application of XL comes from the Aebersold group where they worked out the complete structure of the 55S mammalian mitochondrial ribosome (196).

The highly complex nature of XL (inter/intra protein, etc.) makes identification challenging. MS/MS spectra of crosslinked peptides are often of low quality and difficult to interpret because of low precursor abundance, and insufficient or complex fragmentation (197). There has been resurgence in the use of XL in the last few years in light of improvements in data analysis software (198).

#### *1.4.4 Native Electrospray Mass Spectrometry (Native ESI-MS)*

Contrary to the labeling methods discussed in the previous Sections, the validity of information gleaned from native ESI-MS rely on retention of solution phase structure in the gas phase (199). The experiments comprise of directly electrospraying proteins in 10-500 mM of a volatile salt (e.g. ammonium acetate). ESI-MS generates protonated ions with the  $[M + zH]^{z+}$  ion (Section 1.3.1). Compact structures give rise to narrow charge state distributions. By contrast unfolded conformers generate wide distributions of ions

that are much more highly charged. The first applications of native ESI-MS for structural analysis of proteins exploited this rather simple effect to study protein folding under kinetic and equilibrium conditions (200, 201) Those early studies also revealed that non-covalent protein complexes could be transferred into the vacuum and analyzed using the mass spectrometer (202, 203). To understand why these interactions that originate from the solution phase survive the gas phase environment (204), it is necessary to recall how the ions are generated in the ESI source the first place. ESI mechanism of different protein ions are reviewed in Section 1.3.1.

In recent years, native ESI-MS has proven to be valuable for qualitative and (in favourable cases) quantitative interrogation of protein-protein and protein-ligand interactions (205-207). Nonspecific binding can also affect ESI-MS studies on noncovalent protein–ligand complexes. A diverse set of interactions involving the binding of small molecules to proteins to analysis of MDa protein assemblies have been analyzed using ESI-ESI (111, 208). Sometimes CID of these complexes is used to complement and gain stoichiometric information (209). Some of the more recent applications involve methods that gauge the disorderedness of various proteins (210). Application has also been recently extended to studies of membrane proteins (211).

#### *1.4.5 Ion Mobility Spectrometry (IMS)*

IMS represents one of the most commonly used tools for structural characterization of proteins in the gas phase (212). IMS used to be limited to a handful of laboratories with instruments manufactured in-house. The popularity of the methodology is has been soaring in recent years due to the wide adoption of commercial Q-TOF

systems with travelling-wave ion mobility spectrometry capability (TWIMS)(213, 214). Wide acceptance by the MS community has spurred competition among vendors, and more sophisticated instruments are expected to be introduced in the coming years. It would not be surprising to see IMS become a standard feature on all instruments within the next decade.

In traditional drift tube IMS, ions traverse a tube with length  $L$  filled with a buffer gas under the influence of a weak electric field with strength  $E$ . The time an ion with charge  $z$  spends in this tube, the drift time  $t_d$ , depends on  $\Omega$  according to

$$\Omega = z t_d \frac{eE}{16NL} \left( \frac{18\pi}{\mu k_B T} \right)^{1/2} \frac{760 \text{ Torr}}{p} \frac{T}{273.2 \text{ K}} \quad 1.9$$

where  $\Omega$  is the rotationally averaged collision cross section (CCS) which is a measure of analyte size;  $k_B$  is Boltzmann's constant; and  $N$  is the number density of the buffer gas. Similar to native ESI, the information gained from IMS are only valid if the protein does not undergo major structural transitions as result of the ESI process and in the gas phase. Proteins have evolved to function in an aqueous environment. The role of water for biological function is well-documented (5, 7). It may seem unreasonable to expect the solution phase structure to survive in the gas phase. Indeed there are examples in the literature that provide evidence for these cases (83, 215-218) Nonetheless, the survival of noncovalent interactions as well as the overall size of the analytes in the gas phase can be taken as indication of the retention of some aspects of solution phase structure in the gas phase (219). Metastability of these gas phase structures and their trapping in local free energy minima has been proposed as a mechanism (220).



The conversion of the measured  $t_d$  to CCS values on traditional instruments equipped with drift tubes is performed using eq 1.9. By contrast extracting  $\Omega$  values from data acquired on TWIMS instruments is difficult. This is due to a lack of fundamental understanding of ion dynamics and their mobility within TW ion guides (213). To circumvent this problem, calibration methods based on reference CCSs values have been developed (221). The measured CCSs profiles are gauged against theoretical CCSs calculated for known crystal structures, or for structural models generated using MD simulations of gas phase ions. A number of algorithms have been developed for the purpose of calculating  $\Omega$  values from model structures (222). The exact hard sphere scattering algorithm appears to have the best mix of accuracy and computational economy. However, its application to high molecular weight systems is not feasible. The projection approximation method is by far the most cost effective method computationally. However it yields values that are systematically low (220, 223-225). The Robinson group and others employ a correction factor (226), but this approach is not undisputed among IMS practitioners.

Extracting high resolution structural information from IMS data is challenging because a given CCS may be compatible with a large number of candidate structures.(227) Protein ion-buffer gas interactions, among other factors, could complicate CCS measurements and subsequent efforts towards generating model structures. (228-230) Nonetheless several groups employ an integrative approach whereby a combination of CCS data, molecular modeling, and more recently constraints from cross linking experiments are used to build model of multi-subunit protein complexes.

Aside from structural modeling applications, one of the most useful features of IMS is its application as an orthogonal shape-selective separation method. IMS provides an extra dimension of separation, increases peak capacity and together with state of the art LC methods enables studying increasingly complex within a short amount of time (231).

## 1.5 Integrative Structural Biology

From the discussion in Sections 1.2 and 1.3, it clear that there is no single method that provides information on all aspects of protein structure and dynamics. For example, X-ray crystallography (Section 1.2.11.2.1) and cryo-EM (Section 1.2.2) yield valuable data on structure and overall architecture, but leave out important information about protein dynamics. NMR (Section 1.2.3) reports on dynamics on a wide time-scale, but struggles with large asymmetric assemblies and requires large amount of material. A number of MS-based methods (Section 1.4) are capable of dealing with complex systems, but deliver only low- and medium-resolution data. Computational methods (Section 1.2.5) often need experimental data to “steer” them towards the right direction in challenging simulations. In addition, proteins often are not amenable to one method or another, leaving gaps in our understanding. The emerging picture is that the “whole is greater than the sum of its parts”. The way forward appears to be performing a number of complementary experiments and amalgamating the obtained information to gain the complete picture (232). A number of examples have recently appeared in the literatures that demonstrate the potential of these *integrative* structural biology approaches (196,

233, 234). These point to a shifting landscape in structural biology: away from a single-perspective view of proteins and towards an integrative and comprehensive exploration.

## 1.6 Scope of Thesis

Chapter 2 and 3 of this dissertation build on previous work by Hambly (235) and Stocks (236) to study protein conformers under equilibrium and kinetic conditions. Chapter 2 examines the possibility of performing  $\cdot\text{OH}$  labeling experiments on semi-folded conformers populated under semi-denaturing pH values. In addition, the complementary nature of side chains solvent accessibility measurements to backbone hydrogen bonding status from HDX studies is demonstrated. In Chapter 3, submillisecond events in the folding of a model protein was followed using an easy-to-manufacture ultrarapid mixer and pulsed  $\cdot\text{OH}$  labeling. It is shown that the short radical lifetime afforded by the method is well-suited for taking snapshots of the fleeting intermediates that populate the folding pathway of proteins.

Chapter 4 addresses the possibility of studying the partially folded states of proteins employed in Chapter 2 using an complementary IMS methodology. Prior to performing the measurements, the impact of the source activation parameters is critically evaluated by comparing collisional unfolding profiles of protein ions. Once gentle experimental conditions are established, the gas phase conformational properties of semi-folded protein species are critically compared to the known trends derived from solution phase studies.

Finally Chapter 5 utilizes HDX-MS to monitor the conformational dynamics of a membrane-embedded molecular motor,  $F_0F_1$  ATP synthase. Using advanced peptide

mapping and multidimensional separation strategies, site-specific information regarding changes in the dynamics of the enzyme under engine on/engine off conditions was obtained. The experimental protocol developed here offers a general avenue for studying other membrane protein complexes in their natural membrane environment.

## 1.7 References

1. Voet DV, J. G., and Pratt, C.W. (2004) *Fundamentals of Biochemistry, 2nd Edition* (John Wiley & Sons).
2. Dill KA & MacCallum JL (2012) The Protein-Folding Problem, 50 Years On. *Science* 338: 1042-1046.
3. Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, Sonenberg N, Kay LE, & Forman-Kay JD (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* 519(7541): 106-U240.
4. van Holde K, Johnson W, & Shing Ho P (2006) *Principles of Physical Biochemistry* (Pearson Prentice Hall, Upper Saddle River, NJ) 2nd Ed.
5. Bakker HJ (2012) Water's response to the fear of water. *Nature* 491(7425): 533-535.
6. Baldwin RL (2014) Dynamic hydration shell restores Kauzmann's 1959 explanation of how the hydrophobic factor drives protein folding. *Proc. Natl. Acad. Sci. U. S. A.* 111(36): 13052-13056.
7. Lin MM & Zewail AH (2012) Hydrophobic forces and the length limit of foldable protein domains. *Proc. Natl. Acad. Sci. U.S.A.* 109: 9851-9856.
8. Pace CN, Shirley BA, McNutt M, & Gajiwala K (1996) Forces contributing to the conformational stability of proteins. *FASEB J.* 10: 75-83.

9. Shi Z, Krantz BA, Kallenbach N, & Sosnick TR (2002) Contribution of Hydrogen Bonding to Protein Stability Estimated from Isotope Effects. *Biochemistry* 41: 2120-2129.
10. Israelachvili JN (1974) Van der Waals forces in biological systems. *Q. Rev. Biophys.* 6: 341-387.
11. Kumar S & Nussinov R (1999) Salt bridge stability in monomeric proteins. *J. Mol. Biol.* 293(5): 1241-1255.
12. Hendsch ZS & Tidor B (1994) Do Salt Bridges Stabilize Proteins? A Continuum Electrostatic Analysis. *Protein Sci.* 3(2): 211-226.
13. Baxa MC, Haddadian EJ, Jumper JM, Freed KF, & Sosnick TR (2014) Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations. *Proc. Natl. Acad. Sci. U. S. A.* 111(43): 15396-15401.
14. Anfinsen CB (1973) Principles that Govern the Folding of Protein Chains. *Science* 181: 223-230.
15. Dill KA, Ozkan SB, Shell MS, & Weikl TR (2008) The Protein Folding Problem. *Annu. Rev. Biophys. Biomol. Struct.* 37: 289-316.
16. Fersht AR (2008) From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat. Rev. Mol. Cell Biol.* 9: 650-654.
17. Englander SW & Mayne L (2014) The nature of protein folding pathways. *Proc. Natl. Acad. Sci. U. S. A.* 111(45): 15873-15880.
18. Levinthal C (1969) Mossbauer Spectroscopy in Biological Systems. *Proceedings of a Meeting Held at Allerton House, Monticello, Illinois*, eds DeBrunner JTP & Munck E (University of Illinois Press, Chicago), pp 22-24.
19. Dill KA (1985) Theory for the Folding and Stability of Globular Proteins. *Biochemistry* 24: 1501-1509.

20. Bryngelson JD & Wolynes PG (1987) Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 84: 7524-7528.
21. Zwanzig R, Szabo A, & Bagchi B (1992) Levinthal's paradox. *Proc. Natl. Acad. Sci. U.S.A.* 89: 20-22.
22. Daggett V & Fersht A (2003) The present view of the mechanism of protein folding. *Nat. Rev. Mol. Cell Biol.* 4: 497-502.
23. Baldwin RL (2008) The Search for Folding Intermediates and the Mechanism of Protein Folding. *Annu. Rev. Biophys.* 37: 1-21.
24. Chamberlain AK & Marqusee S (2000) Comparison of Equilibrium and Kinetic Approaches for Determining Protein Folding Mechanisms. *Adv. Protein Chem.* 53: 283-328.
25. Jennings PA & Wright PE (1993) Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science* 262: 892-896.
26. Compiani M & Capriotti E (2013) Computational and Theoretical Methods for Protein Folding. *Biochemistry* 52(48): 8601-8624.
27. Karplus M & Weaver DL (1994) Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Sci.* 3: 650-668.
28. Sosnick TR & Barrick D (2011) The folding of single domain proteins -- have we reached a consensus? *Curr. Opin. Struct. Biol.* 21: 12-24.
29. Gianni S, Guydosh NR, Khan F, Caldas TD, Mayor U, White GWN, DeMarco ML, Daggett V, & Fersht AR (2003) Unifying features in protein-folding mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 100: 13286-13291.
30. Ferguson N & Fersht AR (2003) Early events in protein folding. *Curr. Opin. Struct. Biol.* 13: 75-81.
31. Haran G (2012) How, when and why proteins collapse: the relation to folding. *Curr. Opin. Struct. Biol.* 22: 14-20.

32. Henzler-Wildman K & Kern D (2007) Dynamic personalities of proteins. *Nature* 450(7172): 964-972.
33. Boehr DD, Nussinov R, & Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* 5: 789-796.
34. Stock D, Leslie AGW, & Walker JE (1999) Molecular Architecture of the Rotary Motor in ATP synthase. *Science* 286: 1700-1705.
35. Noskov SY, Bernèche S, & Roux B (2004) Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* 431: 830-834.
36. Tzeng S-R & Kalodimos CG (2013) Allosteric inhibition through suppression of transient conformational states. *Nat. Chem. Biol.* 9(7): 462-465.
37. Fang J, Nevin P, Kairys V, Venclovas Č, Engen JR, & Beuning PJ (2014) Conformational Analysis of Processivity Clamps in Solution Demonstrates that Tertiary Structure Does Not Correlate with Protein Dynamics. *Structure* 22(4): 572-581.
38. Fenwick RB, van den Bedem H, Fraser JS, & Wright PE (2014) Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proc. Natl. Acad. Sci. U. S. A.* 111(4): E445-E454.
39. Knowles TPJ, Vendruscolo M, & Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* 15(6): 384-396.
40. Gershenson A, Gierasch LM, Pastore A, & Radford SE (2014) Energy landscapes of functional proteins are inherently risky. *Nat. Chem. Biol.* 10(11): 884-891.
41. Jones N (2014) Atomic Secrets. *Nature* 505(7485): 602-602.
42. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, & Phillips DC (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* 181(4610): 662-666.

43. Parker MW (2003) Protein Structure from X-Ray Diffraction. *J. Biol. Phys.* 29: 341-362.
44. Keen DA & Goodwin AL (2015) The crystallography of correlated disorder. *Nature* 521(7552): 303-309.
45. Carugo O & Argos P (1997) Protein-protein crystal-packing contacts. *Protein Sci.* 6(10): 2261-2263.
46. Bogan MJ (2013) X-ray Free Electron Lasers Motivate Bioanalytical Characterization of Protein Nanocrystals: Serial Femtosecond Crystallography. *Anal. Chem.* 85(7): 3464-3471.
47. Waldrop MM (2014) X-ray science: The big guns. *Nature* 505: 604-606.
48. Suga M, *et al.* (2015) Native structure of photosystem II at 1.95 angstrom resolution viewed by femtosecond X-ray pulses. *Nature* 517(7532): 99-U265.
49. Kern J, *et al.* (2013) Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature. *Science* 340(6131): 491-495.
50. Milne JLS, *et al.* (2013) Cryo-electron microscopy - a primer for the non-microscopist. *Febs J.* 280(1): 28-45.
51. Cheng Y (2015) Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* 161(3): 450-457.
52. Zhao J, Benlekbir S, & Rubinstein JL (2015) Electron cryomicroscopy observation of rotational states in a eukaryotic V-ATPase. *Nature* 521(7551): 241-245.
53. Allegretti M, Klusch N, Mills DJ, Vonck J, Kuehlbrandt W, & Davies KM (2015) Horizontal membrane-intrinsic alpha-helices in the stator a-subunit of an F-type ATP synthase. *Nature* 521(7551): 237-240.
54. Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu R-M, Zhu P, & Li G (2014) Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. *Science* 344(6182): 376-380.



55. Beck F, *et al.* (2012) Near-atomic resolution structural model of the yeast 26S proteasome. *Proc. Natl. Acad. Sci. U.S.A.* 109(37): 14870-14875.
56. Lander GC, Estrin E, Matyskiela ME, Bashore C, Nogales E, & Martin A (2012) Complete subunit architecture of the proteasome regulatory particle. *Nature* 482(7384): 186-191.
57. Smith MTJ & Rubinstein JL (2014) Structural Biology Beyond blob-ology. *Science* 345(6197): 617-619.
58. Palmer AG (2004) NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* 104(8): 3623-3640.
59. Mittermaier A & Kay LE (2006) New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Science* 312: 224-228.
60. Haensel R, Luh LM, Corbeski I, Trantirek L, & Doetsch V (2014) In-Cell NMR and EPR Spectroscopy of Biomacromolecules. *Angew. Chem. Int. Ed.* 53(39): 10300-10314.
61. Kay LE, Ikura M, Tschudin R, & Bax A (1990) 3-dimensional triple-resonance NMR-spectroscopy of isotopically enriched proteins. *J. Mag. Res.* 89(3): 496-514.
62. Pervushin K, Riek R, Wider G, & Wuthrich K (1997) Attenuated T<sub>2</sub> relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.* 94: 12366-12371.
63. Loria JP, Rance M, & Palmer AG (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* 121(10): 2331-2332.
64. Vallurupalli P, Bouvignies G, & Kay LE (2012) Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *J. Am. Chem. Soc.* 134(19): 8148-8161.
65. Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM, & Kay LE (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* 430: 586-590.

66. Neudecker P, Robustelli P, Cavalli A, Walsh P, Lundström P, Zarrine-Afsar A, Sharpe S, Vendruscolo M, & Kay LE (2012) Structure of an Intermediate State in Protein Folding and Aggregation. *Science* 336: 362-366.
67. Wuthrich K (2003) NMR studies of structure and function of biological macromolecules (Nobel Lecture). *Angew. Chem. Int. Ed.* 42(29): 3340-3363.
68. Rosenzweig R & Kay LE (2014) Bringing Dynamic Molecular Machines into Focus by Methyl-TROSY NMR. *Annual Review of Biochemistry, Vol 83* 83: 291-315.
69. Mainz A, Religa TL, Sprangers R, Linser R, Kay LE, & Reif B (2013) NMR Spectroscopy of Soluble Protein Complexes at One Mega-Dalton and Beyond. *Angew. Chem. Int. Ed.:* 8746-8751.
70. Sprangers R & Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445: 618-622.
71. Zhuravleva A, Clerico EM, & Gierasch LM (2012) An Interdomain Energetic Tug-of-War Creates the Allosterically Active State in Hsp70 Molecular Chaperones. *Cell* 151(6): 1296-1307.
72. Mund M, Overbeck JH, Ullmann J, & Sprangers R (2013) LEGO-NMR Spectroscopy: A Method to Visualize Individual Subunits in Large Heteromeric Complexes. *Angew. Chem. Int. Ed.* 52: 11401-11405.
73. Roder H, Maki K, & Cheng H (2006) Early Events in Protein Folding Explored by Rapid Mixing Methods. *Chem. Rev.* 106: 1836-1861.
74. Sugawara T, Kuwajima K, & Sugai S (1991) Folding of Staphylococcal Nuclease A Studied by Equilibrium and Kinetic Circular Dichroism Spectra. *Biochemistry* 30: 2698-2706.
75. Uzawa T, Akiyama S, Kimura T, Takahashi S, Ishimori K, Morishima I, & Fujisawa T (2004) Collapse and search dynamics of apomyoglobin folding revealed by submillisecond observations of alpha-helical content and compactness. *Proc. Natl. Acad. Sci. U.S.A.* 101(5): 1171-1176.

76. Jones CM, Henry ER, Hu Y, Chan C, Luck SD, Bhuyan A, Roder H, Hofrichter J, & Eaton WA (1993) Fast events in protein folding initiated by nanosecond laser photolysis. *Proc. Natl. Acad. Sci. U.S.A.* 90: 11860-11864.
77. Camilloni C, Sutto L, Provasi D, Tiana G, & Broglia RA (2008) Early events in protein folding: Is there something more than hydrophobic collapse? *Protein Sci.* 17: 1424-1433.
78. Chen Y & Barkley MD (1998) Toward Understanding Tryptophan Fluorescence in Proteins. *Biochemistry* 37: 9976-9982.
79. Vallee-Belisle A & Michnick SW (2012) Visualizing transient protein-folding intermediates by tryptophan-scanning mutagenesis. *Nat. Struct. Mol. Biol.* 19(7): 731-736.
80. Elöve GA, Bhuyan AK, & Roder H (1994) Kinetic Mechanism of Cytochrome c Folding: Involvement of the Heme and Its Ligands. *Biochemistry* 33: 6925-6935.
81. Welker E, Maki K, Shastry MCR, Juminaga D, Bhat R, Scheraga HA, & Roder H (2004) Ultrarapid mixing experiments shed new light on the characteristics of the initial conformational ensemble during the folding of ribonuclease A. *Proc. Natl. Acad. Sci. U.S.A.* 101: 17681-17686.
82. Talbot FO, Rullo A, Yao H, & Jockusch RA (2010) Fluorescence Resonance Energy Transfer in Gaseous, Mass-Selected Polyproline Peptides. *J. Am. Chem. Soc.* 132: 16156-16164.
83. Frankevich V, Barylyuk K, Chingin K, Nieckarz R, & Zenobi R (2013) Native Biomolecules in the Gas Phase? The Case of Green Fluorescent Protein. *ChemPhysChem.* 14: 929-935.
84. Czar MF, Zosel F, König I, Nettels D, Wunderlich B, Schuler B, Zarrine-Afsar A, & Jockusch RA (2015) Gas-Phase FRET Efficiency Measurements To Probe the Conformation of Mass-Selected Proteins. *Anal. Chem.* 87(15): 7559-7565.
85. Cornish PV & Ha T (2007) A survey of single-molecule techniques in chemical biology. *ACS Chemical Biology* 2(1): 53-61.

86. Schuler B & Eaton WA (2008) Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* 18: 16-26.
87. Chung HS & Eaton WA (2013) Single-molecule fluorescence probes dynamics of barrier crossing. *Nature* 502(7473): 685-688.
88. Noji H, Yasuda R, Yoshida M, & Kinosita K (1997) Direct observation of the rotation of F-1-ATPase. *Nature* 386(6622): 299-302.
89. Diez M, *et al.* (2004) Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. *Nat. Struct. Mol. Biol.* 11(2): 135-141.
90. Wirth AJ, Platkov M, & Gruebele M (2013) Temporal Variation of a Protein Folding Energy Landscape in the Cell. *J. Am. Chem. Soc.* 135(51): 19215-19221.
91. Bujalowski PJ & Oberhauser AF (2013) Tracking unfolding and refolding reactions of single proteins using atomic force microscopy methods. *Methods* 60(2): 151-160.
92. Kellner R, Hofmann H, Barducci A, Wunderlich B, Nettels D, & Schuler B (2014) Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *Proc. Natl. Acad. Sci. U. S. A.* 111(37): 13355-13360.
93. Kaiser CM, Goldman DH, Chodera JD, Tinoco I, & Bustamante C (2011) The Ribosome Modulates Nascent Protein Folding. *Science* 334: 1723-1727.
94. Karplus M & McCammon JA (2002) Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 9: 646-652.
95. Dror RO, Dirks RM, Grossman JP, Xu H, & Shaw DE (2012) Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* 41: 429-452.
96. Borrell B (2008) Chemistry: Power play. *Nature* 451(7176): 240-243.
97. Piana S, Lindorff-Larsen K, & Shaw DE (2013) Atomic-level description of ubiquitin folding. *Proc. Natl. Acad. Sci. U.S.A.* 110(15): 5915-5920.

98. Henry ER, Best RB, & Eaton WA (2013) Comparing a simple theoretical model for protein folding with all-atom molecular dynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* 110(44): 17880-17885.
99. Dickson A & Brooks CL, III (2013) Native States of Fast-Folding Proteins Are Kinetic Traps. *J. Am. Chem. Soc.* 135(12): 4729-4734.
100. Shan Y, Kim ET, Eastwood MP, Dror RO, Seeliger MA, & Shaw DE (2011) How Does a Drug Molecule Find Its Target Binding Site? *J. Am. Chem. Soc.* 133: 9181-9183.
101. Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch EM, Wilson IA, & Baker D (2011) Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. *Science* 332(6031): 816-821.
102. Koga N, Tatsumi-Koga R, Liu GH, Xiao R, Acton TB, Montelione GT, & Baker D (2012) Principles for designing ideal protein structures. *Nature* 491(7423): 222-227.
103. Shan Y, Arkhipov A, Kim ET, Pan AC, & Shaw DE (2013) Transitions to catalytically inactive conformations in EGFR kinase. *Proc. Natl. Acad. Sci. U.S.A.* 110(18): 7270-7275.
104. Best RB (2012) Atomistic molecular simulations of protein folding. *Curr. Op. Struct. Biol.* 22: 52-61.
105. Griffiths J (2008) A Brief History of Mass Spectrometry. *Anal. Chem.* 80: 5678-5683.
106. Pavia DL (2009) Mass Spectrometry. *Introduction to Spectroscopy*, (Brooks/Cole, Cengage Learning, Belmont, CA), 4th Ed, pp 427, 431, 434.
107. Fenn JB (2003) Electrospray Wings for Molecular Elephants (Nobel Lecture). *Angew. Chem. Int. Ed.* 42: 3871-3894.
108. Karas M & Hillenkamp F (1988) Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10 000 Daltons. *Anal. Chem.* 60: 2299-2301.

109. Cooks RG, Ouyang Z, Takats Z, & Wiseman JM (2006) Ambient Mass Spectrometry. *Science* 311: 1566-1570.
110. Woodall DW, Wang B, Inutan ED, Narayan SB, & Trimpin S (2015) High-throughput characterization of small and large molecules using only a matrix and the vacuum of a mass spectrometer. *Anal. Chem.* 87(9): 4667-4674.
111. Snijder J, Rose RJ, Veessler D, Johnson JE, & Heck AJR (2013) Studying 18 MDa Virus Assemblies with Native Mass Spectrometry. *Angewandte Chemie International Edition* 52: 4020-4023.
112. Paizs B & Suhai S (2005) Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* 24: 508-548.
113. Zhurov KO, Fornelli L, Wodrich MD, Laskay UA, & Tsybin YO (2013) Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.* 42(12): 5014-5030.
114. Konermann L, Ahadi E, Rodriguez AD, & Vahidi S (2013) Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* 85: 2-9.
115. Ogorzalek Loo RR, Lakshmanan R, & Loo JA (2014) What Protein Charging (and Supercharging) Reveal about the Mechanism of Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* 25: 1675-1693.
116. Hogan CJ, Carroll JA, Rohrs HW, Biswas P, & Gross ML (2009) Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Anal. Chem.* 81: 369-377.
117. de la Mora FJ (2000) Electrospray Ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* 406: 93-104.
118. Konermann L, Rodriguez AD, & Liu J (2012) On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. *Anal. Chem.* 84: 6798-6804.
119. de Hoffmann E & Stroobant V (2007) *Mass Spectrometry: Principles and Applications* San Francisco, USA), Vol 18.

120. Picotti P & Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods* 9(6): 555-566.
121. Ben-Nissan G & Sharon M (2011) Capturing protein structural kinetics by mass spectrometry. *Chem. Soc. Rev.* 40: 3627-3637.
122. Kaltashov IA, Bobst CE, & Abzalimov RR (2013) Mass spectrometry-based methods to study protein architecture and dynamics. *Protein Sci.* 22(5): 530-544.
123. Konermann L, Vahidi S, & Sowole MA (2014) Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. *Anal. Chem.* 86: 213-232.
124. Hvidt A & Nielsen SO (1966) Hydrogen exchange in proteins. *Adv. Protein Chem.* 21: 287-386.
125. Skinner JJ, Lim WK, Bedard S, Black BE, & Englander SW (2012) Protein dynamics viewed by hydrogen exchange. *Protein Sci.* 21: 996-1005.
126. Englander SW & Poulsen FM (1969) Hydrogen-tritium exchange of the random chain peptide. *Biopolymers* 7: 379-393.
127. Englander JJ, Calhoun DB, & Englander SW (1979) Measurement and calibration of peptide group hydrogen-deuterium exchange by ultraviolet spectrophotometry. *Anal. Biochem.* 92: 517-524.
128. Roder H, Elöve GA, & Englander SW (1988) Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature* 335: 700-704.
129. Katta V & Chait BT (1991) Conformational Changes in Proteins Probed by Hydrogen-exchange Electrospray-ionisation Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 5: 214-217.
130. Zhang Z & Smith DL (1993) Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* 2: 522-531.

131. Konermann L, Pan J, & Liu Y (2011) Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem. Soc. Rev.* 40: 1224-1234.
132. Percy AJ, Rey M, Burns KM, & Schriemer DC (2012) Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry-A review. *Anal. Chim. Acta* 721: 7-21.
133. Engen JR & Wales TE (2015) Analytical Aspects of Hydrogen Exchange Mass Spectrometry. *Annu. Rev. Anal. Chem.* 8(1): 127-148.
134. Amon S, Trelle MB, Jensen ON, & Jorgensen TJD (2012) Spatially Resolved Protein Hydrogen Exchange Measured by Subzero-Cooled Chip-Based Nanoelectrospray Ionization Tandem Mass Spectrometry. *Anal. Chem.* 84: 4467-4473.
135. Pan J, Zhang S, Parker CE, & Borchers CH (2014) Subzero Temperature Chromatography and Top-Down Mass Spectrometry for Protein Higher-Order Structure Characterization: Method Validation and Application to Therapeutic Antibodies. *J. Am. Chem. Soc.* 136(37): 13065-13071.
136. Black WA, Stocks BB, Mellors JS, Engen JR, & Ramsey JM (2015) Utilizing Microchip Capillary Electrophoresis Electrospray Ionization for Hydrogen Exchange Mass Spectrometry. *Anal. Chem.* 87: 6280-6287.
137. Rand KD, Zehl M, & Jorgensen TJD (2014) Measuring the Hydrogen/Deuterium Exchange of Proteins at High Spatial Resolution by Mass Spectrometry: Overcoming Gas-Phase Hydrogen/Deuterium Scrambling. *Acc. Chem. Res.* 47(10): 3018-3027.
138. Brodbelt JS (2014) Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem. Soc. Rev.* 43(8): 2757-2783.
139. Hu WB, Walters BT, Kan ZY, Mayne L, Rosen LE, Marqusee S, & Englander SW (2013) Stepwise protein folding at near amino acid resolution by hydrogen exchange and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 110(19): 7684-7689.
140. Althaus E, Canzar S, Ehrler C, Emmett MR, Karrenbauer A, Marshall AG, Meyer-Base A, Tipton JD, & Zhang H-M (2010) Computing H/D-exchange rates



- of single residues from data of proteolytic fragments. *BMC Bioinformatics* 11: 424.
141. Kan ZY, Walters BT, Mayne L, & Englander SW (2013) Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. *Proc. Natl. Acad. Sci. U. S. A.* 110(41): 16438-16443.
  142. Fajer PG, Bou-Assaf GM, & Marshall AG (2012) Improved Sequence Resolution by Global Analysis of Overlapped Peptides in Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 23: 1202-1208.
  143. Sheff JG, Rey M, & Schriemer DC (2013) Peptide-Column Interactions and Their Influence on Back Exchange Rates in Hydrogen/Deuterium Exchange-MS. *J. Am. Soc. Mass Spectrom.* 24(7): 1006-1115.
  144. McAllister RG & Konermann L (2015) Challenges in the Interpretation of Protein H/D Exchange Data: A Molecular Dynamics Simulation Perspective. *Biochemistry* 54(16): 2683-2692.
  145. Noe F (2015) Beating the Millisecond Barrier in Molecular Dynamics Simulations. *Biophys. J.* 108: 228-229.
  146. Pirrone GF, Iacob RE, & Engen JR (2015) Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Anal. Chem.* 87(1): 99-118.
  147. Powell KD, Ghaemmaghami S, Wang MZ, Ma L, Oas TG, & Fitzgerald MC (2002) A General Mass Spectrometry-Based Assay for the Quantitation of Protein-Ligand Binding Interactions in Solution. *J. Am. Chem. Soc.* 124: 10256-10257.
  148. Konermann L, Rodriguez AD, & Sowole MA (2014) Type 1 and Type 2 scenarios in hydrogen exchange mass spectrometry studies on protein-ligand complexes. *Analyst* 139(23): 6078-6087.
  149. Pan JX, Han J, Borchers CH, & Konermann L (2012) Structure and Dynamics of Small Soluble A beta(1-40) Oligomers Studied by Top-Down Hydrogen Exchange Mass Spectrometry. *Biochemistry* 51(17): 3694-3703.

150. Mysling S, Betzer C, Jensen PH, & Jorgensen TJD (2013) Characterizing the Dynamics of  $\alpha$ -Synuclein Oligomers Using Hydrogen/Deuterium Exchange Monitored by Mass Spectrometry. *Biochemistry* 52(51): 9097-9103.
151. Skora L, *et al.* (2013) Burial of the Polymorphic Residue 129 in Amyloid Fibrils of Prion Stop Mutants. *J. Biol. Chem.* 288(5): 2994-3002.
152. Pan LY, Salas-Solano O, & Valliere-Douglass JF (2015) Antibody Structural Integrity of Site-Specific Antibody-Drug Conjugates Investigated by Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* 87(11): 5669-5676.
153. Iacob RE, *et al.* (2015) Hydrogen/deuterium exchange mass spectrometry applied to IL-23 interaction characteristics: potential impact for therapeutics. *Expert Review of Proteomics* 12(2): 159-169.
154. Marciano DP, Dharmarajan V, & Griffin PR (2014) HDX-MS guided drug discovery: small molecules and biopharmaceuticals. *Curr. Op. Struct. Biol.*: 105-111.
155. Pan Y, Brown L, & Konermann L (2011) Hydrogen Exchange Mass Spectrometry of Bacteriorhodopsin Reveals Light-Induced Changes in the Structural Dynamics of a Biomolecular Machine. *J. Am. Chem. Soc.* 133: 20237-20244.
156. Lee C-T, Malzahn E, Brunner M, & Mayer MP (2014) Light-Induced Differences in Conformational Dynamics of the Circadian Clock Regulator VIVID. *J. Mol. Biol.* 426(3): 601-610.
157. Whitelegge JP (2013) Integral Membrane Proteins and Bilayer Proteomics. *Anal. Chem.* 85(5): 2558-2568.
158. Goswami D, Devarakonda S, Chalmers MJ, Pascal BD, Spiegelman BM, & Griffin PR (2013) Time Window Expansion for HDX Analysis of an Intrinsically Disordered Protein. *J. Am. Soc. Mass Spectrom.* 24: 1584-1592.
159. Keppel TR & Weis DD (2015) Mapping Residual Structure in Intrinsically Disordered Proteins at Residue Resolution Using Millisecond Hydrogen/Deuterium Exchange and Residue Averaging. *J. Am. Soc. Mass Spectrom.* 26: 547-554.

160. Tsutsui Y, Dela Cruz RG, & Wintrode PL (2012) The folding mechanism of the metastable serpin  $\alpha_1$ -antitrypsin. *Proc. Natl. Acad. Sci. U.S.A.* 109: 4467-4472.
161. Walters BT, Mayne L, Hinshaw JR, Sosnick TR, & Englander SW (2013) Folding of a large protein at high structural resolution. *Proc. Natl. Acad. Sci. U.S.A.* 110(47): 18898-18903.
162. Georgescauld F, Popova K, Gupta AJ, Bracher A, Engen JR, Hayer-Hartl M, & Hartl FU (2014) GroEL/ES Chaperonin Modulates the Mechanism and Accelerates the Rate of TIM-Barrel Domain Folding. *Cell* 157(4): 922-934.
163. Pan J, Heath BL, Jockusch RA, & Konermann L (2012) Structural Interrogation of Electrosprayed Peptide Ions by Gas-Phase H/D Exchange and Electron Capture Dissociation Mass Spectrometry. *Anal. Chem.* 84: 373-378.
164. Rand KD, Pringle SD, Morris M, & Brown JM (2012) Site-Specific Analysis of Gas-Phase Hydrogen/Deuterium Exchange of Peptides and Proteins by Electron Transfer Dissociation. *Anal. Chem.* 84: 1931-1940.
165. Mendoza VL & Vachet RW (2009) Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* 28: 785-815.
166. Gau BC, Chen JW, & Gross ML (2013) Fast photochemical oxidation of proteins for comparing solvent-accessibility accompanying protein folding: Data processing and application to barstar. *BBA-Proteins Proteomics* 1834(6): 1230-1238.
167. Li X, Li Z, Xie B, & Sharp JS (2013) Improved Identification and Relative Quantification of Sites of Peptide and Protein Oxidation for Hydroxyl Radical Footprinting. *J. Am. Soc. Mass Spectrom.* 24(11): 1767-1776.
168. Luchini A, Espina V, & Liotta LA (2014) Protein painting reveals solvent-excluded drug targets hidden within native protein-protein interfaces. *Nat. Commun.* 5.
169. O'Brien JP, Needham BD, Henderson JC, Nowicki EM, Trent MS, & Brodbelt JS (2014) 193 nm Ultraviolet Photodissociation Mass Spectrometry for the Structural

- Elucidation of Lipid A Compounds in Complex Mixtures. *Anal. Chem.* 86(4): 2138-2145.
170. Chen J, Cui W, Giblin D, & Gross ML (2012) New Protein Footprinting: Fast Photochemical Iodination Combined with Top-Down and Bottom-Up Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 23: 1306-1318.
  171. Zhang H, Shen W, Rempel DL, Monsey J, Vidavsky I, Gross ML, & Bose R (2011) Carboxyl-group footprinting maps the dimerization interface and phosphorylation-induced conformational changes of a membrane-associated tyrosine kinase. *Mol. Cell. Proteomics* 10: M110.005678.
  172. Collier TS, Diraviyam K, Monsey J, Shen W, Sept D, & Bose R (2013) Carboxyl Group Footprinting Mass Spectrometry and Molecular Dynamics Identify Key Interactions in the HER2-HER3 Receptor Tyrosine Kinase Interface. *J. Biol. Chem.* 288(35): 25254-25264.
  173. Jha SK & Udgaonkar JB (2007) Exploring the Cooperativity of the Fast Folding Reaction of a Small Protein Using Pulsed Thiol Labeling and Mass Spectrometry. *J. Biol. Chem.* 282: 37479-37491.
  174. Krishnan B & Gierasch LM (2011) Dynamic local unfolding in the serpin  $\alpha$ -1 antitrypsin provides a mechanism for loop insertion and polymerization. *Nat. Struct. Mol. Biol.* 18: 222-226.
  175. Wang L & Chance MR (2011) Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Anal. Chem.* 83: 7234-7241.
  176. Xu G & Chance MR (2007) Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* 107: 3514-3543.
  177. Shcherbakova I, Mitra S, Beer RH, & Brenowitz M (2006) Fast Fenton footprinting: a laboratory-based method for the time-resolved analysis of DNA, RNA and proteins. *Nucleic Acids Res.* 34: e48.
  178. McClintock C, Kertesz V, & Hettich RL (2008) Development of an Electrochemical Oxidation Method for Probing Higher Order Protein Structure with Mass Spectrometry. *Anal. Chem.* 80: 3304-3317.

179. Maleknia SD & Downard KM (2014) Advances in radical probe mass spectrometry for protein footprinting in chemical biology applications. *Chem. Soc. Rev.* 43(10): 3244-3258.
180. Schorzman AN, Perera L, Cutalo-Patterson JM, Pedersen LC, Pedersen LG, Kunkel TA, & Tomer KB (2011) Modeling of the DNA-binding site of yeast Pms1 by mass spectrometry. *DNA Repair* 10: 454-465.
181. Watson C, Janik I, Zhuang T, Charvatova O, Woods RJ, & Sharp JS (2009) Pulsed Electron Beam Water Radiolysis for Submicrosecond Hydroxyl Radical Protein Footprinting. *Anal. Chem.* 81(7): 2496-2505.
182. Hambly DM & Gross ML (2005) Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* 16: 2057-2063.
183. Gau BC, Sharp JS, Rempel DL, & Gross ML (2009) Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* 81: 6563-6571.
184. Konermann L, Stocks BB, & Czarny T (2010) Laminar Flow Effects During Laser-Induced Oxidative Labeling For Protein Structural Studies by Mass Spectrometry. *Anal. Chem.* 82: 6667-6674.
185. Stocks BB & Konermann L (2009) Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* 81: 20-27.
186. Stocks BB & Konermann L (2010) Time-Dependent Changes in Side Chain Solvent Accessibility During Cytochrome c Folding Probed By Pulsed Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* 398: 362-373.
187. Stocks BB, Sarkar A, Wintrode PL, & Konermann L (2012) Early Hydrophobic Collapse of alpha(1)-Antitrypsin Facilitates Formation of a Metastable State: Insights from Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* 423(5): 789-799.
188. Chen J, Rempel DL, Gau B, & Gross ML (2012) Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level. *J. Am. Chem. Soc.* 134: 18724-18731.

189. Jones LM, Sperry JB, Carroll JA, & Gross ML (2011) Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Anal. Chem.* 83: 7657-7661.
190. Zhang H, Gau BC, Jones LM, Vidavsky I, & Gross ML (2011) Fast Photochemical Oxidation of Proteins for Comparing Structures of Protein-Ligand Complexes: The Calmodulin-Peptide Model System. *Anal. Chem.* 83: 311-318.
191. Serpa JJ, Parker CE, Petrotchenko EV, Han J, Pan J, & Borchers CH (2012) Mass spectrometry-based structural proteomics. *Eur. J. Mass Spectrom.* 18: 251-267.
192. Gau BC, Chen H, Zhang Y, & Gross ML (2010) Sulfate Radical Anion as a New Reagent for Fast Photochemical Oxidation of Proteins. *Anal. Chem.* 82: 7821-7827.
193. Jumper CC, Bomgarden R, Rogers J, Etienne C, & Schriemer DC (2012) High-Resolution Mapping of Carbene-Based Protein Footprints. *Anal. Chem.* 84: 4411-4418.
194. Jumper CC & Schriemer DC (2011) Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Anal. Chem.* 83(8): 2913–2920.
195. Walzthoeni T, Leitner A, Stengel F, & Aebersold R (2013) Mass spectrometry supported determination of protein complex structure. *Curr. Op. Struct. Biol.* 23(2): 252-260.
196. Greber BJ, Bieri P, Leibundgut M, Leitner A, Aebersold R, Boehringer D, & Ban N (2015) The complete structure of the 55S mammalian mitochondrial ribosome. *Science* 348(6232): 303-308.
197. Walzthoeni T, Claassen M, Leitner A, Herzog F, Bohn S, Forster F, Beck M, & Aebersold R (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat. Methods* 9(9): 901-903.
198. Yang B, *et al.* (2012) Identification of cross-linked peptides from complex samples. *Nat. Methods* 9(9): 904-906.
199. Marcoux J & Robinson CV (2013) Twenty Years of Gas Phase Structural Biology. *Structure* 21(9): 1541-1550.

200. Konermann L, Collings BA, & Douglas DJ (1997) Cytochrome *c* Folding Kinetics Studied by Time-Resolved Electrospray Ionization Mass Spectrometry. *Biochemistry* 36: 5554-5559.
201. Konermann L & Douglas DJ (1998) Equilibrium Unfolding of Proteins Monitored by Electrospray Ionization Mass Spectrometry: Distinguishing Two-State from Multi-State Transitions. *Rapid Commun. Mass Spectrom.* 12: 435-442.
202. Chowdhury SK, Katta V, & Chait BT (1990) Probing Conformational Changes in Proteins by Mass Spectrometry. *J. Am. Chem. Soc.* 112: 9012-9013.
203. Ganem B, Li Y-T, & Henion JD (1991) Observation of Noncovalent Enzyme-Substrate and Enzyme-Product Complexes by Ion-Spray Mass Spectrometry. *J. Am. Chem. Soc.* 113: 7818-7819.
204. Benesch JLP & Robinson CV (2009) Dehydrated but unharmed. *Nature* 462: 576-577.
205. Loo JA (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* 16: 1-23.
206. Benesch JLP & Ruotolo BT (2011) Mass spectrometry: come of age for structural and dynamical biology. *Curr. Op. Struct. Biol.* 21: 641-649.
207. Mehmood S, Allison TM, & Robinson CV (2015) Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annual Review of Physical Chemistry, Vol 66* 66: 453-474.
208. Hopper JTS & Robinson CV (2014) Mass Spectrometry Quantifies Protein Interactions-From Molecular Chaperones to Membrane Porins. *Angew. Chem. Int. Ed.* 53(51): 14002-14015.
209. Hall Z, Hernández H, Marsh JA, Teichmann SA, & Robinson CV (2013) The Role of Salt Bridges, Charge Density, and Subunit Flexibility in Determining Disassembly Routes of Protein Complexes. *Structure* 21(8): 1325-1337.
210. Beveridge R, Covill S, Pacholarz KJ, Kalapothakis JMD, MacPhee CE, & Barran PE (2014) A Mass-Spectrometry-Based Framework To Define the Extent of Disorder in Proteins. *Anal. Chem.* 86(22): 10979-10991.

211. Laganowsky A, Reading E, Hopper JTS, & Robinson CV (2013) Mass spectrometry of intact membrane protein complexes. *Nat. Protocols* 8(4): 639-651.
212. Wyttenbach T, Pierson NA, Clemmer DE, & Bowers MT (2014) Ion Mobility Analysis of Molecular Dynamics. *Annu. Rev. Phys. Chem.* 65: 175–196.
213. Shvartsburg AA & Smith RD (2008) Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* 80: 9689–9699.
214. Giles K, Williams JP, & Campuzano I (2011) Enhancements in travelling wave ion mobility resolution. *Rapid Commun. Mass Spectrom.* 25: 1559-1566.
215. Warnke S, von Helden G, & Pagel K (2013) Protein Structure in the Gas Phase: The Influence of Side-Chain Microsolvation. *J. Am. Chem. Soc.* 135(4): 1177-1180.
216. Ly T & Julian RR (2010) Elucidating the Tertiary Structure of Protein Ions in Vacuo with Site Specific Photoinitiated Radical Reactions. *J. Am. Chem. Soc.* 132: 8602-8609.
217. Hall Z, Politis A, Bush MF, Smith LJ, & Robinson CV (2012) Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics. *J. Am. Chem. Soc.* 134: 3429-3438.
218. Balthasart F, Plavec J, & Gabelica V (2013) Ammonium Ion Binding to DNA G-Quadruplexes: Do Electrospray Mass Spectra Faithfully Reflect the Solution-Phase Species? *J. Am. Soc. Mass Spectrom.* 24: 1-8.
219. Ruotolo BT & Robinson CV (2006) Aspects of native proteins are retained in vacuum. *Curr. Op. Chem. Biol.* 10: 402-408.
220. Wyttenbach T & Bowers MT (2011) Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment. *J. Phys. Chem. B* 115: 12266-12275.



221. Ruotolo BT, Benesch JLP, Sandercock AM, Hyung S-J, & Robinson CV (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* 3: 1139-1152.
222. Shvartsburg AA & Jarrold MF (1996) An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* 261(1-2): 86-91.
223. Scarff CA, Patel VJ, Thalassinos K, & Scrivens JH (2009) Probing Hemoglobin Structure by Means of Traveling-Wave Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 20: 625-631.
224. Jurneczko E & Barran PE (2011) 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. *Analyst* 136: 20-28.
225. Mao Y, Woenckhaus J, Kolafa J, Ratner MA, & Jarrold MF (1999) Thermal unfolding of unsolvated cytochrome c: Experiment and molecular dynamics simulations. *J. Am. Chem. Soc.* 121(12): 2712-2721.
226. Marklund EG, Degiacomi MT, Robinson CV, Baldwin AJ, & Benesch JLP (2015) Collision Cross Sections for Structural Proteomics. *Structure* 23(4): 791-799.
227. Rabuck JN, Hyung S-J, Ko KS, Fox CC, Soellner MB, & Ruotolo BT (2013) Activation State-Selective Kinase Inhibitor Assay Based on Ion Mobility-Mass Spectrometry. *Anal. Chem.* 85(15): 6995-7002.
228. Hogan CJ, Ruotolo BT, Robinson CV, & de la Mora JF (2011) Tandem Differential Mobility Analysis-Mass Spectrometry Reveals Partial Gas-Phase Collapse of the GroEL Complex. *J. Phys. Chem. B* 115(13): 3614-3621.
229. Wyttenbach T, Bleiholder C, & Bowers MT (2013) Factors contributing to the collision cross section of polyatomic ions in the kilodalton to gigadalton range: application to ion mobility measurements. *Anal. Chem.* 85(4): 2191-2199.
230. Chen Y-L, Collings BA, & Douglas DJ (1997) Collision Cross Sections of Myoglobin and Cytochrome c ions with Ne, Ar, and Kr. *J. Am. Soc. Mass Spectrom.* 8: 681-687.

231. Valentine SJ, Kulchania M, Barnes CAS, & Clemmer DE (2001) Multidimensional separations of complex peptide mixtures: a combined high-performance liquid chromatography/ion mobility/time-of-flight mass spectrometry approach. *Int. J. Mass Spectrom.* 212(1-3): 97-109.
232. van den Bedem H & Fraser JS (2015) Integrative, dynamic structural biology at atomic resolution-it's about time. *Nature Methods* 12(4): 307-318.
233. Politis A, Stengel F, Hall Z, Hernandez H, Leitner A, Walzthoeni T, Robinson CV, & Aebersold R (2014) A mass spectrometry-based hybrid method for structural modeling of protein complexes. *Nat. Methods* 11(4): 403-406.
234. Nyon MP, *et al.* (2015) An integrative approach combining ion mobility mass spectrometry, X-ray crystallography and NMR spectroscopy to study the conformational dynamics of  $\alpha$ 1-antitrypsin upon ligand binding. *Protein Science* In Press: 1301-1312.
235. Hambly DM & Gross ML (2007) Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin. *Int. J. Mass Spectrom.* 259: 124-129.
236. Stocks BB, Rezvanpour A, Shaw GS, & Konermann L (2011) Temporal Development of Protein Structure During S100A11 Folding and Dimerization Probed by Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* 409: 669-679.

## **Chapter 2: Mapping pH-Induced Protein Structural Changes under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry**

### **2.1 Introduction**

Protein conformations are governed by the solvent environment. The highly ordered native state is usually favored in neutral aqueous solution. Chemical denaturants or extremes of pH often induce extensive unfolding. Partially disordered intermediates can be populated under mildly denaturing equilibrium conditions (1, 2). These intermediates are of particular interest because they provide insights into folding and aggregation mechanisms (3-6). Unfortunately, intermediates do not crystallize, thereby precluding X-ray structure determinations (7). Some NMR techniques are better suited for this purpose (8), but their application is far from routine. Optical tools only report on global features (5). The structural characterization of equilibrium intermediates, therefore, remains difficult.

An array of mass spectrometry (MS) techniques is available for protein conformational studies (9-14). Amide hydrogen/deuterium exchange (HDX)-MS measures backbone structure and dynamics (15-19). Covalent tagging methods (20, 21) such as  $\cdot\text{OH}$  labeling (22, 23) are complementary to HDX-MS (24).  $\cdot\text{OH}$  induces oxidative modifications at solvent accessible side chains, whereas buried residues are protected. A convenient way to generate  $\cdot\text{OH}$  is the photolysis of  $\text{H}_2\text{O}_2$  by a pulsed UV laser for the "fast photochemical oxidation of proteins" (FPOP) (25). Both HDX and FPOP are promising tools for characterizing partially disordered intermediates. For a

better understanding of the following sections it is beneficial to briefly summarize some basic concepts.

D<sub>2</sub>O exposure of an amide X-H induces conversion to X-D. In the EX2 regime the concentration of X-H after labeling is

$$[X-H] = [X-H]_0 \exp(-g_{HDX}) \quad 2.1$$

where

$$g_{HDX} = K_{op} \times f_{HDX} \quad 2.2$$

$K_{op}$  is the equilibrium constant of opening/closing thermal fluctuations, and

$$f_{HDX} = (k_{acid}[D^+] + k_{base}[OD^-] + k_{D_2O})t \quad 2.3$$

The rate constants  $k_{acid}$  and  $k_{base}$  are for HDX catalysis by D<sup>+</sup> and OD<sup>-</sup>, respectively, and  $k_{D_2O}$  is for solvent catalysis. These three parameters depend on the nature of the neighboring side chains (26, 27). The labeling time is denoted as  $t$ .

A similar framework can be proposed for FPOP (23). ·OH exposure of an unmodified side chain X<sub>u</sub> induces conversion to the oxidized form X<sub>ox</sub>. The concentration of X<sub>u</sub> after labeling is

$$[X_u] = [X_u]_0 \exp(-g_{ox}) \quad 2.4$$

where

$$-g_{ox} = \alpha \times f_{ox} \quad 2.5$$

The dimensionless factor  $\alpha$  represents the solvent accessibility of X<sub>u</sub>. Completely inaccessible side chains have  $\alpha = 0$ , whereas for full exposure  $\alpha$  is equal to unity.  $f_{ox}$  can be broken down according to (23)

$$f_{ox} = k_{int}[\cdot\text{OH}]_0 \tau \quad 2.6$$

where the rate constant  $k_{int}$  reflects the intrinsic reactivity of  $X_u$  (22).  $[\cdot\text{OH}]_0$  is the initial radical concentration of the labeling pulse, and  $\tau$  is the  $\cdot\text{OH}$  lifetime.

Equations 2.2 and 2.5 reveal that the HDX and FPOP levels are both governed by two terms: The first term ( $K_{op}$  or  $\alpha$ ) is of primary interest because it reflects structural properties. The second term ( $f_{HDX}$  or  $f_{ox}$ ) depends on the intrinsic chemistry and on the experimental conditions. Equilibrium unfolding studies have to deal with the complication that *both* terms can be affected by the solvent environment. Changes in  $f_{HDX}$  and  $f_{ox}$  will give rise to "secondary" effects, *i.e.*, alterations in labeling behavior that are not related to conformational changes (28). Evidently, a thorough understanding of these secondary effects is essential for a proper interpretation of experimental data.

We now focus on the characterization of equilibrium intermediates that are formed at low pH. Such experiments are most informative when comparing the properties of an intermediate to those of the native state. For NMR-based HDX investigations the known pH dependence of  $f_{HDX}$  allows the elimination of secondary effects at the single residue level (Equation 2.3) (2, 26, 29). HDX/MS yields peptide-averaged data that complicate the application of such corrections, but in principle the same strategy can be used (27, 30).

The separation of conformational and secondary effects in FPOP has received almost no attention in the literature, and the pH behavior of  $f_{ox}$  (Equations 2.5 and 2.6) remains largely unexplored. Radiolysis (31, 32) and Fenton studies (33) suggest that  $k_{int}$  can be affected by pH, although the extent of these changes is compound-dependent (28). Different oxidation sites will exhibit distinct  $k_{int}$  vs. pH profiles, thus complicating the use

of dosimeter probes (34). The  $\cdot\text{OH}$  lifetime  $\tau$  can be affected by pH-dependent scavenging effects, *e.g.*, in the presence of titratable organic anions such as formate or acetate (28, 32). The pH dependence of  $[\cdot\text{OH}]_0$  is unknown. In  $\gamma$ -ray mediated  $\cdot\text{OH}$  labeling experiments it has been demonstrated that secondary effects can completely mask conformational changes (28).

Here we explore the use of FPOP for monitoring pH-induced conformational changes at equilibrium, using myoglobin (Mb) as model system. This protein has long served as a paradigm for structural MS (12, 19, 25, 35-37). Native holo-myoglobin (hMb) adopts a compact fold with 8  $\alpha$ -helices that accommodate a heme group in a binding pocket (38). Heme removal yields apo-myoglobin (aMb). The structures of hMb and "native" aMb at pH 7 are similar, with exception of the region around helix F (19, 39). Acid-induced equilibrium denaturation of aMb proceeds through a pH 4 intermediate (2, 39). We employ FPOP for studies on hMb and aMb at pH 7, the intermediate at pH 4, and acid-unfolded aMb at pH 2. Although Mb has previously been examined by  $\cdot\text{OH}$  labeling (23, 25, 36, 37), pH-dependent equilibrium investigations involving partially folded intermediates have not been reported yet. We demonstrate that interference from secondary effects is minimal under proper conditions, allowing the use of FPOP for the in-depth structural characterization of partially disordered proteins.

## 2.2 Experimental Section

### 2.2.1 Materials

Horse ferri-hMb, Gln, Met, LeuEnk (YGGFL), [Met<sup>5</sup>] enkephalin (YGGFM), mouse leptin fragment 116-130 amide (SCSLPQTSGLQKPES), human-rat adrenocorticotrophic

hormone fragment 1-10 (SYSMEHFRWG), human [Glu<sup>1</sup>]-fibrinopeptide B (EGVNDNEEGFFSAR), Met-Lys-[Ser<sup>2</sup>, Arg<sup>3</sup>, Pro<sup>5</sup>, Arg<sup>8</sup>]-bradykinin (MKRSRGPSRR), bovine  $\alpha$ -chymotrypsin and bovine liver catalase were purchased from Sigma (St. Louis, MO). Bradykinin (RPPGFSPFR) and human angiotensin I (DRVYIHPFHL) were supplied by Bachem Bioscience (King of Prussia, PA). Sequencing-grade trypsin was obtained from Promega (Madison, WI). aMb was prepared using butanone extraction (40). Subsequently, the solution was dialyzed against aqueous HCl (pH 2), yielding 300  $\mu$ M stock solutions. For all experiments the protein was diluted to a final concentration of 10  $\mu$ M in 10 mM sodium phosphate buffer, and the pH was adjusted using HCl. Prior to labeling, the protein solutions were centrifuged to remove small amounts of precipitate that might have formed. Concentrations were verified optically using  $\epsilon_{409} = 188,000 \text{ M}^{-1} \text{ cm}^{-1}$  for hMb and  $\epsilon_{280} = 13,980 \text{ M}^{-1} \text{ cm}^{-1}$  for aMb (41). All experiments were conducted at room temperature, unless noted otherwise.

### 2.2.2 *Optical Spectroscopy*

Fluorescence emission spectra were recorded on a PTI instrument (Birmingham, NJ) with  $\lambda_{\text{ex}} = 280 \text{ nm}$ . Circular dichroism (CD) data were acquired on a Jasco J-810 instrument (Easton, MD). Raw data were converted to molar ellipticity  $[\theta]$  (42). All solutions were identical to those used for FPOP, except for the presence of H<sub>2</sub>O<sub>2</sub> and Gln.

### 2.2.3 *Oxidative Labeling*

FPOP occurs faster than most conformational changes. Hence, this approach is free of labeling-induced structural artifacts (43). Experiments were conducted as described

previously (23), with minor modifications. Two syringes were employed to mix protein and H<sub>2</sub>O<sub>2</sub> immediately prior to laser exposure. All other solvent conditions remained constant before and after mixing. This strategy minimizes protein contact with peroxide, thereby suppressing background oxidation. Each syringe was advanced at 20  $\mu\text{L min}^{-1}$ . The final mixture contained 0.02% (~6 mM) H<sub>2</sub>O<sub>2</sub>, 15 mM Gln, and 10  $\mu\text{M}$  protein at pH 7, 4, or 2. For experiments on test peptides the concentration of each analyte was 10  $\mu\text{M}$ . An excimer laser (GAM EX 50, Orlando, FL) generating 18 ns laser pulses at 64 Hz, 248 nm, and 37.5 mJ was used for  $\cdot\text{OH}$  production 40 cm downstream of the mixer (standard laser safety procedures have to be followed). Gln is a radical scavenger that reduces the radical lifetime to ca. 1  $\mu\text{s}$  (43). The irradiation window width was ~1 mm. 160  $\mu\text{L}$  capillary outflow aliquots were collected 1 s after labeling in microcentrifuge tubes containing 80  $\mu\text{L}$  of 100 mM phosphate buffer, 5  $\mu\text{M}$  Met and 1  $\mu\text{M}$  catalase. To achieve uniform digestion efficiency, care was taken that all samples had the same solvent composition. To this end, after FPOP the pH 4 and pH 7 samples were acidified to pH 2 prior to flash freezing and lyophilization. The samples were resuspended in 50  $\mu\text{L}$  ammonium bicarbonate (65 mM, pH 8.5) with 35% acetonitrile. Trypsin digestion was conducted using a 1:20 (w/w) enzyme:protein ratio at 37 °C for 24 h. For chymotrypsin digestion a 1:60 enzyme:protein ratio was used at 30 °C. Tryptic peptides formed the basis of quantitative oxidation measurements (see below). The additional use of chymotrypsin was helpful for detecting and verifying oxidation sites. Uniform digestion efficiency was verified by comparing digests for non-irradiated controls.



#### 2.2.4 LC/ESI-MS

All experiments were conducted on a Synapt mass spectrometer (Waters, Milford, MA) equipped with a LockSpray dual electrospray ionization (ESI) source operated at a capillary voltage of +2.8 kV. When using MS/MS mode the lock mass channel was sampled every 60 seconds and contained 5  $\mu$ M leucine enkephalin ( $m/z$  556.2771) in 50:50 methanol:water and 1% acetic acid. The mass spectrometer was coupled to a UPLC (Waters Acquity, Milford, MA) employing either a 1.7  $\mu$ m C4 (BEH300) 2.1 mm  $\times$  50 mm reversed-phase column for intact protein analyses, or a 1.7  $\mu$ m C18 (BEH130) 2.1 mm  $\times$  100 mm column for peptide measurements. Elution was conducted using a water/acetonitrile gradient in the presence of 0.1% formic acid at 100  $\mu$ L min<sup>-1</sup>. The identity of tryptic peptides was confirmed by MS/MS. Signal intensities of unmodified Mb peptides were measured relative to a bradykinin internal standard, which was spiked into each sample prior to LC/MS at a concentration of 5  $\mu$ M. LeuEnk was used as internal standard for test peptide analyses. To find the oxidation sites, tryptic and chymotryptic digests of labeled pH 2 aMb were subjected to data-dependent MS/MS with iterative exclusion (44). Fragmentation was conducted by collision-induced dissociation (CID) with Argon by adjusting the trap collision energy, *i.e.*, the potential offset between entrance ion guide and the trap T-wave. PEAKS 5.3 (45) was used for data analysis. In addition, the digests were analyzed using MS<sup>E</sup>. Collision energy ramps covering the range 10-34 V in 4 V steps were utilized in multiple runs to achieve optimal fragmentation for all peptides. All oxidation site assignments were verified manually.

### 2.2.5 Data Analysis

The extent of oxidation at individual protein segments (tryptic peptides) was tracked by employing background-corrected "fraction unmodified" ( $F_u$ ) values (23, 34). These values were determined by expressing the measured LC/MS signal intensities relative to the internal standard. The ·OH labeling behavior of test peptides was quantified analogously. Partially oxidized peptides are characterized by  $0 < F_u < 1$ , whereas a lack of labeling corresponds to an  $F_u$  value of unity. Tryptic peptide data were further converted to "normalized oxidation levels" (*NOLs*) according to (46)

$$NOL = \frac{\ln F_u - \ln F_u(\text{native})}{\ln F_u(\text{unfolded}) - \ln F_u(\text{native})} \quad 2.7$$

In this expression  $F_u(\text{unfolded})$  represents data obtained for aMb after labeling at pH 2 (unfolded reference), whereas  $F_u(\text{native})$  corresponds to hMb at pH 7 (native reference).

## 2.3 Results and Discussion

### 2.3.1 Optical Spectroscopy

Acid-induced denaturation of aMb usually involves at least three species: native aMb at pH 7, an intermediate at pH 4, and the unfolded state at pH 2 (2, 39). The extent to which the intermediate becomes populated can be modulated by solvent and temperature conditions (2, 39, 47, 48). Also, many older studies were conducted on sperm whale aMb, whereas this work (like most other MS investigations) (12, 19, 25, 35, 36) uses the horse variant. Optical tests were therefore conducted to ensure that our samples showed the expected behavior.

Secondary structure changes were followed by CD spectroscopy. A plateau in the unfolding curve at pH 4 confirms formation of the intermediate (Figure 2.1A) (2). The CD spectrum of native aMb (Figure 2.1B) has minima at 209 and 222 nm, attesting to a high helical content. Acid-unfolded aMb exhibits a minimum around 201 nm. The pH 4 intermediate has 35% lower helicity than native aMb, as judged by the ellipticity at 222 nm (Figure 2.1B). Formation of the intermediate causes a maximum in the Trp fluorescence profile around pH 4 (Figure 2.1C) (47). The emission maxima at pH 7 and pH 4 are close to 330 nm, whereas a shift to ~342 nm has occurred at pH 2 (Figure 2.1D). These data reveal that the Trp7 and Trp14 remain in a hydrophobic environment at pH 7 and 4. The red-shift at pH 2 indicates exposure to water (49). Overall, Figure 2.1 confirms formation of the pH 4 intermediate under the conditions of our work.

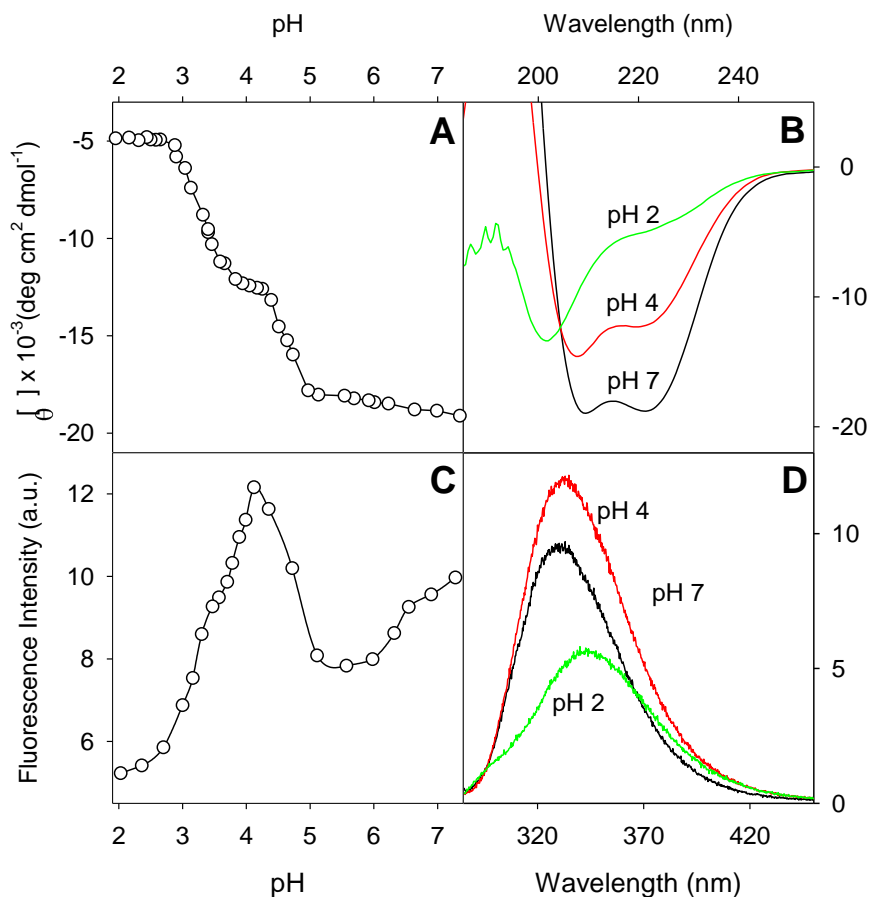


Figure 2.1 Acid-induced unfolding of aMb monitored by CD and fluorescence spectroscopy. (A) mean residue ellipticity monitored at 222 nm. (B) far-UV CD spectra at pH 2, 4, and 7. (C) Trp fluorescence emission intensity recorded at 335 nm. (D) Trp fluorescence emission spectra at pH 2, 4, and 7.

### 2.3.2 Validation of FPOP Strategy Using Test Peptides

Peptides have previously been used for probing secondary effects in  $\cdot\text{OH}$  labeling, albeit only for radiolysis studies (28). The premise of such experiments is that all side chains in a short peptide should remain solvent accessible, regardless of pH. Major changes in oxidation behavior must therefore be attributed to secondary effects. Here we apply this approach to FPOP, using the same experimental conditions as for subsequent Mb studies.

Six peptides were chosen that contain all twenty residues. Free amino acids were not used because their reactivities differ from those in a protein backbone context (22, 34).

Figure 2.2 (filled symbols) displays  $F_u$  values for the six peptides. Major oxidation sites, identified by MS/MS, are indicated in bold. The labeling levels of the peptides are different, reflecting their intrinsic reactivities (22). Importantly,  $F_u$  of all six species remains almost unchanged across the pH range. Free Gln was used in our work to control the  $\cdot\text{OH}$  lifetime  $\tau$  (25). It is fortuitous that the scavenging activity of Gln appears to be fairly uniform, such that  $\tau$  does not undergo significant changes between pH 7 and 2.

Earlier studies showed that thioether conversion to sulfoxide can be affected by protonation of an adjacent carboxylate (31). To verify this effect we examined an additional peptide that contains a C-terminal Met (YGGFM).  $F_u$  for this peptide is hardly affected when lowering the pH from 7 to 4, but it drops significantly at pH 2 (Figure 2.2, 1 open circles). This transition likely reflects protonation of the C-terminus (50). The observed effect is therefore consistent with the known thioether/carboxylate interaction (31). Notably, there is no discernible pH effect when Met is located elsewhere in the sequence (as in MKRSRGPSRR and SYSMEHFRWG, Figure 2.2). The latter peptide is interesting because it has Met situated next to Glu. Glu gets protonated around pH 4, but apparently the carboxylate is too far removed to have an effect on the Met behavior.

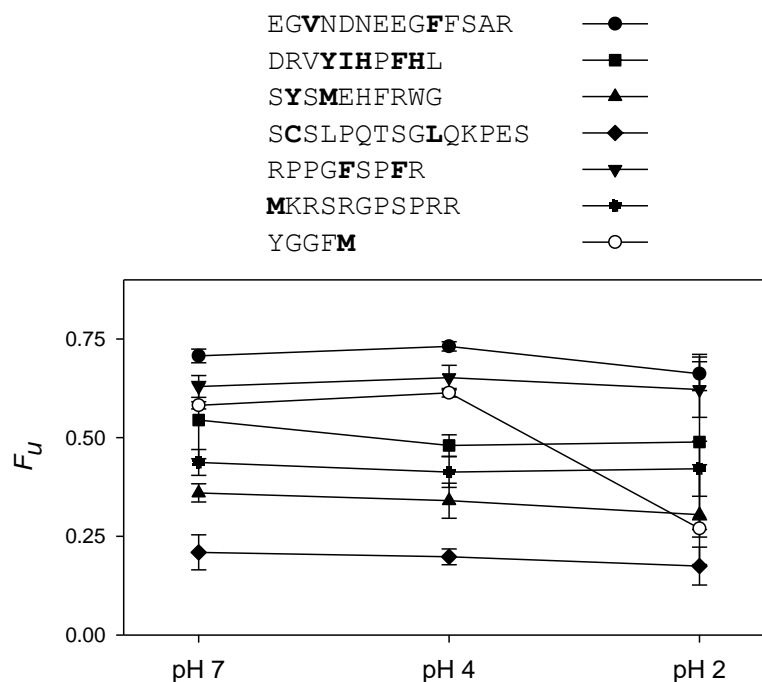


Figure 2.2 Fraction unmodified ( $F_u$ ) after FPOP for various test peptides at different pH. Peptide sequences are indicated; major oxidation sites identified by MS/MS are highlighted in bold. The labeling conditions were identical to those used for subsequent protein experiments.

In conclusion, the data of Figure 2.2 reveal that pH-dependent secondary effects are very minor under the FPOP conditions employed here. It will be seen that protein conformational changes affect the oxidation behavior to a much larger extent. Sequences with a C-terminal Met represent a special case, but such cases are rare. For all other sequences, pH-dependent changes of the intrinsic factor  $f_{ox}$  are small. The FPOP behavior at any given site is therefore governed by the solvent accessibility parameter  $\alpha$  (Equation 2.5). This finding greatly simplifies the application of FPOP for studies on pH-induced conformational changes. Unlike for HDX measurements (Equation 2.3) FPOP data do not have to be corrected for pH effects under the conditions employed here.

### 2.3.3 Intact Protein Analyses

X-ray structural data are available for native hMb (51), but not for aMb. As a point of reference we therefore included native hMb in our FPOP experiments. It is instructive to examine the global labeling pattern before proceeding to spatially-resolved data. Figure 2.3A shows deconvoluted spectra of native hMb and aMb after labeling at pH 7. In addition to the unmodified protein peak at 16,952 Da there is a series of +16 Da covalent adducts that represent oxidative modifications. An earlier FPOP investigation reported intact protein data for native hMb and aMb that were almost indistinguishable (36). In contrast, Figure 2.3A reveals that aMb clearly undergoes more oxidation ( $F_u = 0.52 \pm 0.02$ ) than hMb ( $F_u = 0.60 \pm 0.02$ , Figure 2.3D), consistent with the more open structure of the former (19, 39). Also shown in Figure 2.3A is the mass profile of a non-irradiated control, revealing that background oxidation is minimal. The pH 4 aMb intermediate undergoes more extensive labeling than the native state ( $F_u = 0.40 \pm 0.01$ , Figure 2.3B). Even more oxidation is seen for acid-unfolded aMb at pH 2 ( $F_u = 0.13 \pm 0.01$ , Figure 2.3C). Overall, these FPOP data reveal a successive increase in solvent accessible surface area as the protein progresses from native hMb to the unfolded state. The observed behavior is consistent with a stepwise breakdown in structure that is seen optically (Figure 2.1) and that has been reported in previous NMR investigations (2, 39).

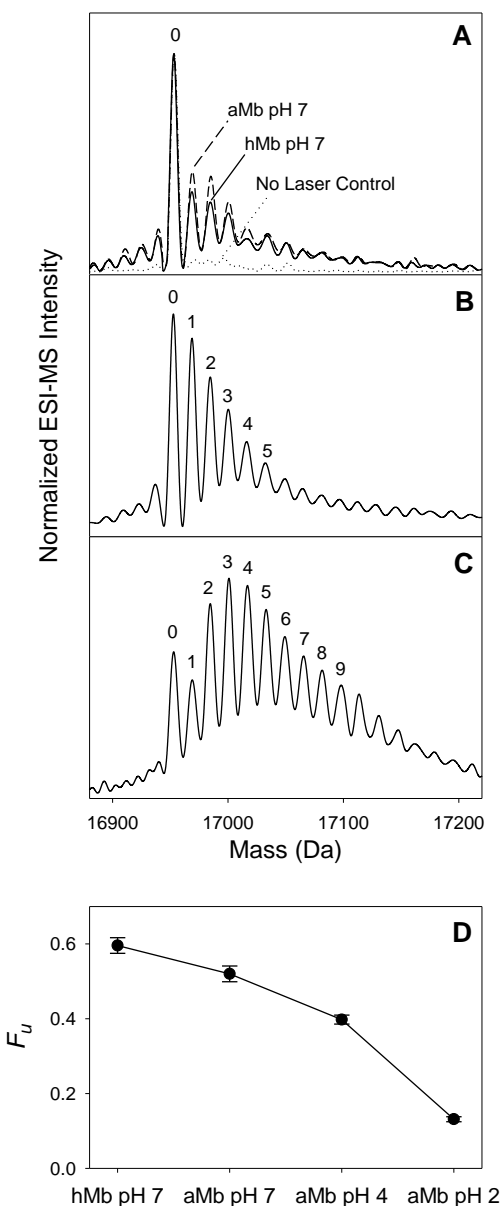


Figure 2.3 Deconvoluted intact myoglobin mass spectra after FPOP under various conditions. (A) hMb (solid line) and aMb (dashed line) at pH 7. Also shown is the spectrum of a non-irradiated hMb control that was otherwise treated identically to the FPOP samples. The denaturing LC mobile phase induces heme loss prior to analysis, such that all samples appear as aMb, regardless of heme binding state during labeling. (B) aMb at pH 4. (C) aMb at pH 2. Numbers (0, 1, 2, ...) indicate how many oxygen atoms (+ 16 Da adducts) are incorporated as a result of  $\cdot\text{OH}$  labeling. (D) Background-corrected "fraction unmodified" ( $F_u$ ) values for the various myoglobin conformers, obtained from intact protein mass distributions. All points represent an average of three independent measurements. Error bars indicate standard deviations.



#### 2.3.4 Solvent Accessibility Measurements Using Normalized Oxidation Levels

Previous FPOP investigations compared hMb and aMb at pH 7, classifying various residues as "not labeled" or "labeled" (25, 36). The strategy used here goes beyond this binary readout by employing *NOLs* for monitoring conformational changes in a quantitative fashion (Equation 2.7). *NOLs* relate conformational changes to the solvent accessibility of two reference structures, *i.e.*, native hMb at neutral pH (51), and unfolded aMb at pH 2 (2, 39, 47).  $NOL = 0$  represents a protein segment (represented by a tryptic peptide) that has an accessibility equal to that in native hMb. Conversely,  $NOL = 1$  describes peptides with an accessibility equal to that in the acid-unfolded state. *NOLs* between zero and unity reflect intermediate degrees of solvent access.

Tryptic digestion of Mb yields 12 detectable peptides (T1, T2, etc., Figure 2.4), with a sequence coverage of ~90%. Primary oxidation sites were detected by MS/MS of labeled unfolded aMb (Figure 2.5). This procedure identified 39 oxidized residues, comprising M, W, F, H, V, I, L, E, P, K, and R (highlighted in bold, Figure 2.4). All of these have previously been found to be oxidizable in studies on other proteins (34). Modified cysteine is missing from our list because Mb does not contain this type of residue.

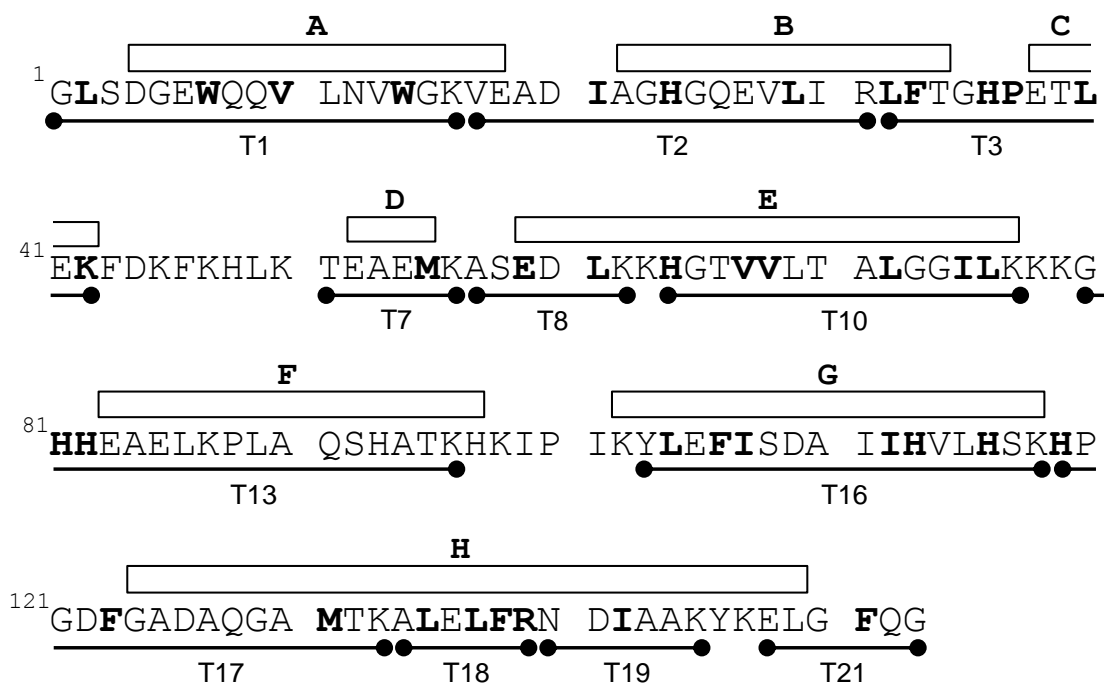


Figure 2.4 Amino acid sequence of myoglobin (pdb code: 1WLA).  $\alpha$ -helices A-H are denoted by square boxes. Tryptic peptides (T1-T21) are indicated by continuous lines. Detected oxidation sites are displayed in boldface.

T1, T2, T16, T17, and T18 show relatively minor *NOL* changes down to pH 4, whereas a major transition occurs upon acidification to pH 2 (Figure 2.6A). In contrast, T3, T7, and T21 retain native-like accessibilities only after heme loss, while becoming quite accessible in the pH 4 intermediate (Figure 2.6B). T8, T10, T13, and T19 lose a fairly substantial amount of protection in each of the three steps. T13 (helix F) has the highest accessibility of all aMb peptides at pH 7 and at pH 4 (Figure 2.6C).

To facilitate a structural interpretation the measured *NOLs* were mapped onto the crystal structure of native hMb (Figure 2.7). A five-level color scheme was used, where blue represents a solvent accessibility close to that of native hMb ( $0 < NOL < 0.2$ ). Red is

used for segments with an accessibility resembling acid-denatured aMb ( $0.8 < NOL < 1$ ). As dictated by the normalization procedure (Equation 2.7) native hMb appears entirely blue (Figure 2.7A), whereas acid-denatured aMb is completely red (Figure 2.7D).

### 2.3.5 *Structure of Native aMb*

Heme removal at pH 7 results in elevated solvent accessibility for helices E and F, and for the C-terminal part of helix H (Figure 2.7B, green). These helices constitute the main elements of the heme binding pocket (51). Our data, therefore, support the utility of FPOP for mapping protein-ligand contacts (52). The relatively low oxidation level of T18 ( $NOL = 0.17$ , blue in Figure 2.7B) might indicate that parts of the binding pocket undergo partial closure in native aMb, in line with data reported by Hambly and Gross (25).

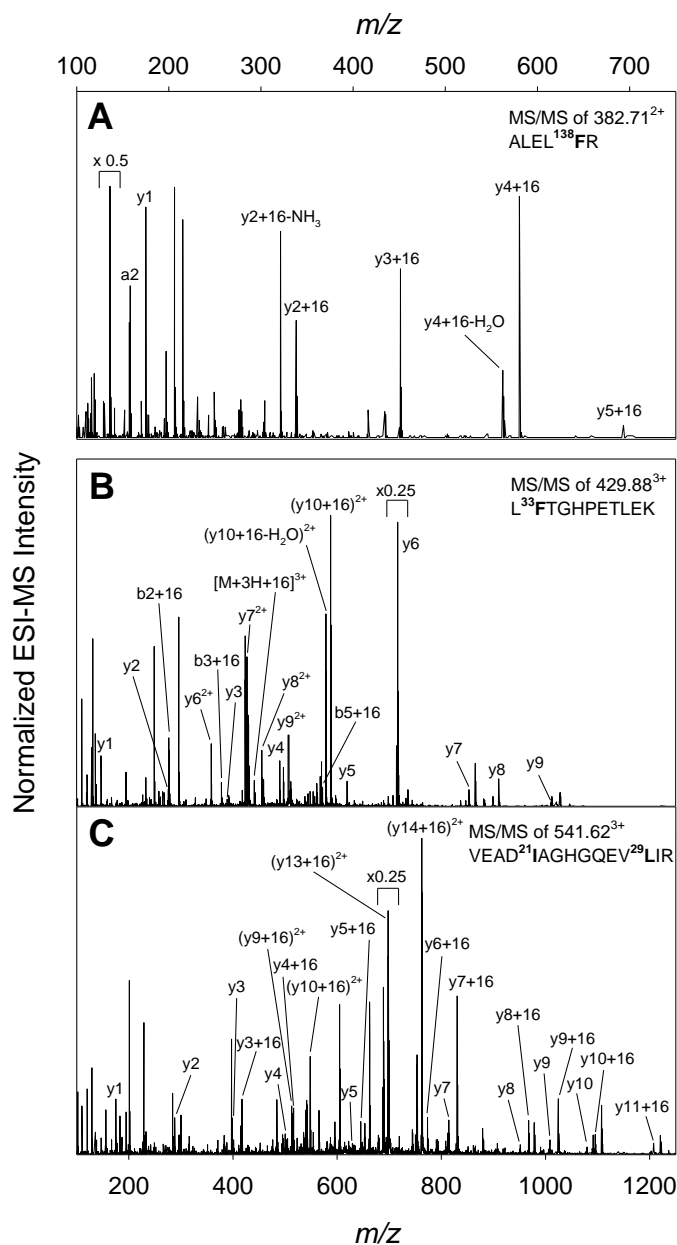


Figure 2.5 Representative tandem mass spectra, obtained by CID of tryptic peptides after FPOP at pH 2. Detected oxidation sites are highlighted in bold. (A) CID of [T18+16] results in y ions that are all shifted by 16 Da, except for y1. This pinpoints F138 as the labeling site. (B) MS/MS data for [T3+16] can be treated similarly. Unmodified y ions are observed up to y9, along with y10+16. Thus F33 is oxidized. Panel (C) illustrates a more complex case, obtained by CID of [T2+16] which represents two isobaric ions. The spectrum contains unmodified y1 and y2. Starting from y3, pairs of unmodified and oxidized y ions appear, identifying L29 as the first oxidation site. Starting with y11 no more unmodified fragments are seen, such that I21 must be the second oxidation site.

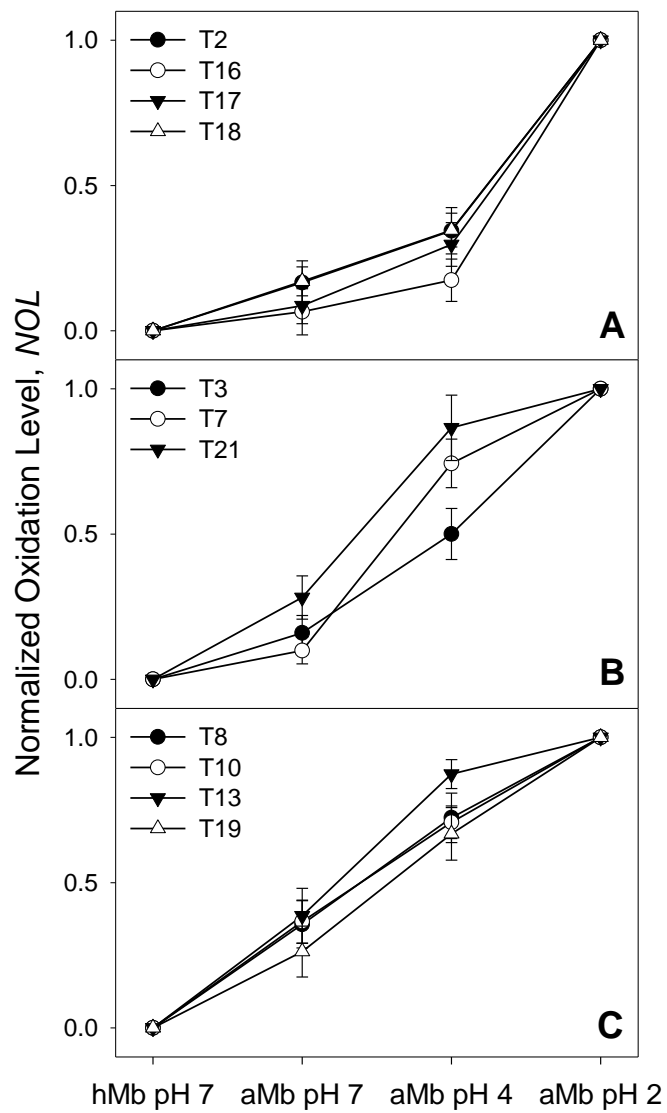


Figure 2.6 Normalized oxidation levels (*NOLs*) of protein segments (tryptic peptides T1, T2, ...) at different pH. The data were split into panels to prevent clutter. Error bars were determined as in Figure 2.3.

NMR measurements suggest a major reduction in the stability of helix F after heme loss (39). HDX/MS reveals that helix F in native aMb exchanges faster than in native hMb (19). Nonetheless, it has remained unclear whether helix F is completely unfolded in aMb, or whether structure is partially retained (2, 19, 39). The FPOP data of Figure 2.7B confirm that helix F is affected by heme loss. However, the *NOL* values of only up to ~0.4 are quite moderate when compared to the unfolded reference. Our data therefore show that helix F does not become completely unstructured in native aMb.

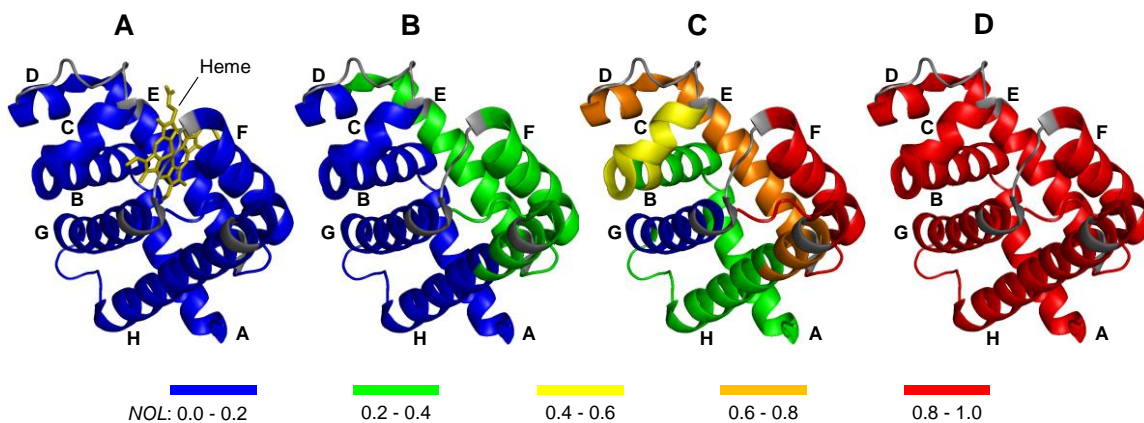


Figure 2.7 Mapping of *NOLs* (from Figure 2.6) to the crystal structure of native hMb (1WLA)(51) using a five-level color code as indicated along the bottom. The four panels represent (A) hMb at pH 7, (B) aMb at pH 7, (C) aMb at pH 4, and (D) aMb at pH 2. Regions not covered by tryptic peptide mapping are shown in gray.

It has previously been noted that HDX and FPOP represent complementary, rather than redundant structural probes (24). The current work supports this assessment, as seen from a comparison of results obtained by both techniques. Figure 2.8A displays FPOP results for native aMb, using an orientation that provides a head-on view of the binding pocket. Heme is included for visualization purposes. Recent HDX data (19) are displayed

in Figure 2.8B; amides that undergo a major deuteration increase upon heme removal are highlighted. It is seen that both techniques detect structural changes in helices F and in the C-terminal part of helix H. In contrast, the backbone hydrogen bonding of helix E is not strongly affected by heme loss (19) (Figure 2.8B), whereas FPOP clearly shows an increased side chain solvent accessibility in this region (Figure 2.8A). The type of information obtained from the two techniques is therefore clearly different.

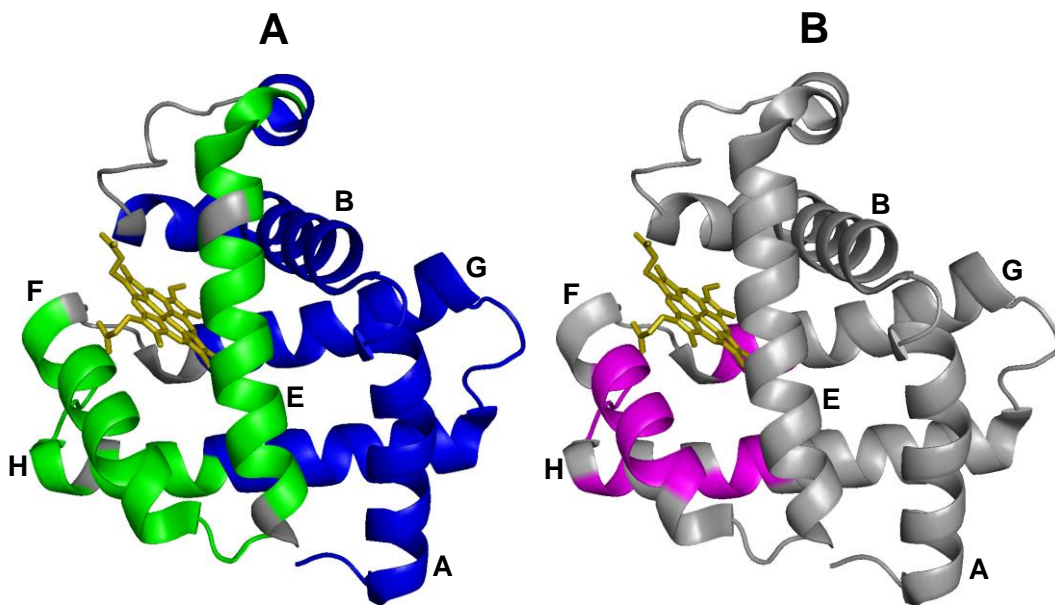


Figure 2.8 Visualization of regions most affected by heme removal at pH 7 as detected by (A) FPOP and (B) HDX. In panel A, regions with elevated *NOLs* (0.2 - 0.4) are highlighted in green, segments that retain native-like solvent accessibility are depicted in blue. In panel B, residues that display the largest loss in HDX protection (amide deuteration difference > 0.5, relative to native hMb)(19) are colored magenta.

### 2.3.6 *The pH 4 Intermediate*

The aMb pH 4 intermediate epitomizes a "molten globule", characterized by substantial secondary structure but with relatively little tertiary packing (53, 54). The exact conformational properties of this intermediate remain under investigation (55, 56). FPOP

reveals that helix F is extensively disordered at pH 4 (Figure 2.7C, red). A substantial loss in solvent protection is also seen for helices E and D (orange), and for helix C (yellow). Helices A, as well as parts of helices B and H retain partial protection (green). Remarkably, the solvent accessibility of helix G at pH 4 is comparable to that in native hMb (Figure 2.7C, blue). This protection points to a central role of helix G as core element at pH 4. Our data therefore suggest a semi-folded ABGH structure for the intermediate, where A, B and H protect G from three sides (as in native hMb) (51).

Our findings complement earlier NMR spin relaxation and HDX measurements that also point to a pH 4 state with partially intact but fluctuating AGH helices, and with participation of helix B (A[B]GH) (2, 39, 56). Participation of helix B in this partially structured state is clearly evident from our data (Figure 2.7C, green). NMR suggests that helix A represents a key stabilizing element because it appears to have the most secondary structure at pH 4 (39). Our data result in a somewhat different picture, emphasizing the role of helix G as central nucleus that allows hydrophobic docking of the surrounding semi-folded helices A, B, and H. Partial hydration of the pH 4 core was seen in spin labeling experiments, consistent with moderate solvent access to A, B, and H (Figure 2.7C, green). Finally, we note that the *equilibrium* ABGH unfolding intermediate characterized here is very different from the BEF[G] species that becomes populated during hMb unfolding in *kinetic* experiments (23).

## 2.4 Conclusions

Oxidative labeling methods such as FPOP are being increasingly used for mapping changes in solvent accessibility that occur as a result of conformational transitions or



binding events. A potential problem is the fact that comparative experiments involve data acquired under different conditions, requiring a separation of conformational and secondary effects. Denaturants such as formic and acetic acid, urea, and guanidinium salts act as scavengers that can suppress labeling (28). Similarly, changes in protein concentration (e.g., after addition of putative binding partners) can reduce oxidation nonspecifically (57).

This work examined the use of FPOP for monitoring acid-induced conformational changes. To equalize  $\cdot\text{OH}$  labeling conditions all measurements were conducted at the same protein concentration (57). The solutions contained phosphate buffer that was pH-adjusted by HCl. Organic acids and bases were avoided, since the scavenging activity of those species depends on their titration state (28, 32). We also ensured that proteolysis was conducted using the same solvent environment for samples that had been labeled at different pH. Under these optimized conditions secondary effects were found to be minimal in the range that is commonly used in unfolding experiments, *i.e.*, pH 7 to 2. Thus, changes in FPOP behavior can be directly interpreted in the context of protein conformational transitions. The situation is more complicated for pH-dependent HDX/MS studies. The intrinsic pH profile of each individual amide is different (26) and therefore the required correction strategies can become tedious when applied at the peptide level. In comparison, FPOP experiments are more straightforward due to the near-absence of intrinsic pH effects.

Structural insights obtained here by FPOP for the Mb model system are complementary to earlier data from traditional bioanalytical tools. Our findings pave the way towards  $\cdot\text{OH}$  labeling experiments on pH-induced transitions of other proteins.

## 2.5 References

1. Bai Y, Sosnick TR, Mayne L, & Englander SW (1995) Protein Folding Intermediates: Native State Hydrogen Exchange. *Science* 269: 192-197.
2. Hughson FM, Wright PE, & Baldwin RL (1990) Structural Characterisation of a Partly Folded Apomyoglobin Intermediate. *Science* 249: 1544-1548.
3. Chiti F & Dobson CM (2006) Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* 75: 333-366.
4. Detoma AS, Salamekh S, Ramamoorthy A, & Lim MH (2012) Misfolded proteins in Alzheimer's disease and type II diabetes. *Chem. Soc. Rev.* 41: 608-621.
5. Bartlett AI & Radford SE (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* 16: 582-588.
6. Chamberlain AK & Marqusee S (2000) Comparison of Equilibrium and Kinetic Approaches for Determining Protein Folding Mechanisms. *Adv. Protein Chem.* 53: 283-328.
7. Parker MW (2003) Protein Structure from X-Ray Diffraction. *J. Biol. Phys.* 29: 341-362.
8. Mittermaier A & Kay LE (2006) New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Science* 312: 224-228.
9. Kaltashov IA & Eyles SJ (2012) *Mass Spectrometry in Structural Biology and Biophysics: Architecture, Dynamics, and Interactions of Biomolecules* (Wiley, Hoboken, NJ) 2nd Ed.
10. Sharon M (2010) How Far Can We Go with Structural Mass Spectrometry of Protein Complexes? *J. Am. Soc. Mass Spectrom.* 21: 487-500.
11. Ashcroft AE (2010) Mass Spectrometry and the Amyloid Problem - How Far Can We Go in the Gas Phase? *J. Am. Soc. Mass Spectrom.* 21: 1087-1096.

12. Dobo A & Kaltashov IA (2001) Detection of Multiple Protein Conformational Ensembles in Solution via Deconvolution of Charge-State Distributions in ESI MS. *Anal. Chem.* 73: 4763-4773.
13. Ly T, Liu ZJ, Pujanauski BG, Sarpong R, & Julian RR (2008) Surveying ubiquitin structure by noncovalent attachment of distance constrained bis(crown) ethers. *Anal. Chem.* 80: 5059-5064.
14. Dimova K, *et al.* (2009) Structural insights into the calmodulin-Munc13 interaction obtained by cross-linking and mass spectrometry. *Biochemistry* 48: 5908-5921.
15. Englander SW (2006) Hydrogen exchange and mass spectrometry: A historical perspective. *J. Am. Soc. Mass Spectrom.* 17(11): 1481-1489.
16. Percy AJ, Rey M, Burns KM, & Schriemer DC (2012) Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry-A review. *Anal. Chim. Acta* 721: 7-21.
17. Jacob RE & Engen JR (2012) Hydrogen Exchange Mass Spectrometry: Are We Out of the Quicksand? *J. Am. Soc. Mass Spectrom.* 23: 1003-1010.
18. Rand KD, Zehl M, Jensen ON, & Jørgensen TJD (2009) Protein Hydrogen Exchange Measured at Single-Residue Resolution by Electron Transfer Dissociation Mass Spectrometry. *Anal. Chem.* 81: 5577-5584.
19. Pan J, Han J, Borchers CH, & Konermann L (2009) Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* 131(35): 12801-12808.
20. Mendoza VL & Vachet RW (2009) Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* 28: 785-815.
21. Jumper CC, Bomgarden R, Rogers J, Etienne C, & Schriemer DC (2012) High-Resolution Mapping of Carbene-Based Protein Footprints. *Anal. Chem.* 84: 4411-4418.

22. Wang L & Chance MR (2011) Structural Mass Spectrometry of Proteins Using Hydroxyl Radical Based Protein Footprinting. *Anal. Chem.* 83: 7234-7241.
23. Stocks BB & Konermann L (2009) Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* 81: 20-27.
24. Zheng X, Wintrode PL, & Chance MR (2008) Complementary Structural Mass Spectrometry Techniques Reveal Local Dynamics in Functionally Important Regions of a Metastable Serpin. *Structure* 16: 38-51.
25. Hambly DM & Gross ML (2005) Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* 16: 2057-2063.
26. Bai Y, Milne JS, Mayne L, & Englander SW (1993) Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins: Struct., Funct., Genet.* 17: 75-86.
27. Coales SJ, E SY, Lee JE, Ma A, Morrow JA, & Hamuro Y (2010) Expansion of time window for mass spectrometric measurement of amide hydrogen/deuterium exchange reactions. *Rapid Commun. Mass Spectrom.* 24: 3585-3592.
28. Tong X, Wren JC, & Konermann L (2008)  $\gamma$ -Ray-Mediated Oxidative Labeling for Detecting Protein Conformational Changes by Electrospray Mass Spectrometry. *Anal. Chem.* 80: 2222-2231.
29. Pan Y & Briggs MS (1992) Hydrogen Exchange in Native and Alcohol Forms of Ubiquitin. *Biochemistry* 31: 11405-11412.
30. Keppel TR, Howard BA, & Weis DD (2011) Mapping Unstructured Regions and Synergistic Folding in Intrinsically Disordered Proteins with Amide H/D Exchange Mass Spectrometry. *Biochemistry* 50(40): 8722-8732.
31. Hiller KO, Masloch B, Gobl M, & Asmus KD (1981) Mechanism of the OH radical induced oxidation of methionine in aqueous solution. *J. Am. Chem. Soc.* 103: 2734-2743.
32. Buxton GV, Greenstock CL, Helman WP, & Ross AB (1988) Critical Review of Rate Constants for Reactions of Hydrated Electrons, Hydrogen Atoms and

- Hydroxyl Radicals (.OH/.O) in Aqueous Solution. *J. Phys. Chem. Ref. Data* 17: 513-886.
33. Schoneich C & Yang J (1996) Oxidation of methionine peptide by fenton systems: the importance of peptide sequence, neighbouring groups and EDTA. *J. Chem. Soc., Perkin Trans. 2*: 915-924.
  34. Xu G & Chance MR (2007) Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* 107: 3514-3543.
  35. Johnson RS & Walsh KA (1994) Mass spectrometric measurement of protein amide hydrogen exchange rates of apo- and holo-myoglobin. *Protein Sci.* 3: 2411-2418.
  36. Hambly DM & Gross ML (2007) Laser flash photochemical oxidation to locate heme binding and conformational changes in myoglobin. *Int. J. Mass Spectrom.* 259: 124-129.
  37. Maleknia SD & Downard KM (2001) Unfolding of apomyoglobin helices by synchrotron radiolysis and mass spectrometry. *Eur. J. Biochem.* 268: 5578-5588.
  38. Evans SV & Brayer GD (1990) High-resolution Study of the Three-dimensional Structure of Horse Heart Metmyoglobin. *J. Mol. Biol.* 213: 885-897.
  39. Eliezer D, Yao J, Dyson HJ, & Wright PE (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* 5: 148-155.
  40. Teale FWJ (1959) Cleavage of the heme-protein link by acid methylethylketone. *Biochim. Biophys. Acta* 35: 543.
  41. Antonini E & Brunori M (1971) *Hemoglobin and Myoglobin in Their Reactions With Ligands* (North-Holland Publishing Company, Amsterdam, London).
  42. Kelly SW, Jess TJ, & Price NC (2005) How to Study Protein by Circular Dichroism. *Biochim. Biophys. Acta* 1751: 119-139.

43. Gau BC, Sharp JS, Rempel DL, & Gross ML (2009) Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* 81: 6563–6571.
44. Bendall SC, *et al.* (2009) An Enhanced Mass Spectrometry Approach Reveals Human Embryonic Stem Cell Growth Factors in Culture. *Mol. Cell. Proteomics* 8: 421-432.
45. Lacerda CMR, Xin L, Rogers I, & Reardon KF (2008) Analysis of iTRAQ data using Mascot and Peaks quantification algorithms. *Briefings Funct. Genomics Proteomics* 7: 119-126.
46. Stocks BB, Rezvanpour A, Shaw GS, & Konermann L (2011) Temporal Development of Protein Structure During S100A11 Folding and Dimerization Probed by Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* 409: 669-679.
47. Goto Y & Fink AL (1990) Phase Diagram for Acidic Conformational States of Apomyoglobin. *J. Mol. Biol.* 214: 803-805.
48. Eliezer D, Jennings PA, Dyson HJ, & Wright PE (1997) Populating the equilibrium molten globule state of apomyoglobin under conditions suitable for structural characterization by NMR. *FEBS Lett.* 417: 92-96.
49. van Holde K, Johnson W, & Shing Ho P (2006) *Principles of Physical Biochemistry* (Pearson Prentice Hall, Upper Saddle River, NJ) 2nd Ed.
50. Creighton TE (1993) *Proteins* (W. H. Freeman & Co, New York).
51. Maurus R, Overall CM, Bogumil R, Luo Y, Mauk AG, Smith M, & Brayer GD (1997) A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket. *Biochim. Biophys. Acta* 1341: 1-13.
52. Zhang H, Gau BC, Jones LM, Vidavsky I, & Gross ML (2011) Fast Photochemical Oxidation of Proteins for Comparing Structures of Protein-Ligand Complexes: The Calmodulin-Peptide Model System. *Anal. Chem.* 83: 311-318.
53. Ptitsyn OB (1987) Protein Folding: Hypotheses and Experiments. *J. Protein Chem.* 6: 273-293.

54. Barrick D & Badwin RL (1993) The molten globule intermediate of apomyoglobin and the process of protein folding. *Protein Sci.* 2: 869-876.
55. Armstrong BD, Choi J, Lopez C, Wesener DA, Hubbell W, Cavagnero S, & Han S (2011) Site-Specific Hydration Dynamics in the Nonpolar Core of a Molten Globule by Dynamic Nuclear Polarization of Water. *J. Am. Chem. Soc.* 133: 5987-5995.
56. Bertagna AM & Barrick D (2004) Nonspecific hydrophobic interactions stabilize an equilibrium intermediate of apomyoglobin at a key position within the AGH region. *Proc. Natl. Acad. Sci. U.S.A.* 101: 12514-12519.
57. Hambly DM & Gross ML (2009) Microsecond Time-Scale Hydroxyl Radical Profiling of Solvent-Accessible Protein Residues. *Protein Mass Spectrometry*, ed Whitelegge JP (Elsevier, Amsterdam), Vol 52, pp 151-177.

## **Chapter 3: Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling**

### **3.1 Introduction**

Forty years after Anfinsen's Nobel Prize (1), protein folding continues to be a mesmerizing research area (2). During the folding process polypeptide chains undergo conformational motions on a multidimensional energy landscape, from an initial disordered state to the compact native conformation. Folding is driven by free energy gradients. The loss of conformational entropy is compensated by the formation of hydrophobic contacts, hydrogen bonds, and electrostatic interactions (3). Kinetic investigations play a central role for deciphering folding mechanisms. Many of these studies focus on the structural characterization of transient intermediates (4-6). Unfortunately, this task is complicated by the rapid time scale of folding. Depending on the protein, the process goes to completion within microseconds to minutes (7, 8).

Kinetic folding experiments require the combination of a suitable trigger with a technique that can probe the ensuing conformational changes. Mixing-based triggers (for denaturant dilution or pH jumps) are most widely used (9). Stopped-flow mixers have a dead time of ca. 1 ms (10). Turbulent (11-13) and laminar (14-16) continuous-flow mixers provide access to times as short as ~10  $\mu$ s. Even faster triggering is possible via pressure jumps (17), or by employing pulsed lasers for temperature (T) jump (18, 19), electron injection (20), or photodissociation events (21). An ideal detection technique for



these rapid time-resolved studies would report on the protein conformation at a level of detail comparable to X-ray crystallographic or NMR-based structure determination methods. Devising such a technique is difficult, which is one of the main reasons why the folding mechanisms of many proteins remain poorly understood.

Examining the submillisecond regime of folding poses the greatest experimental challenges. Yet, these early events hold the key for understanding fundamental issues such as the interplay between hydrophobic collapse and secondary structure formation (22-25), and the issue of downhill (26) vs. barrier-limited folding (27). Detection techniques used for submillisecond studies include circular dichroism (CD) and fluorescence spectroscopy, energy transfer measurements, and small-angle X-ray scattering (SAXS) (4, 11, 28-30). While these tools provide valuable data, the information obtained is mostly limited to global structural features. The use of site-specific reporter groups is an interesting approach (31), but the introduction of such moieties may disrupt the protein structure. Pulsed hydrogen-deuterium exchange (HDX) monitors hydrogen bond formation in a spatially-resolved fashion. The readout of HDX experiments can be performed by NMR (9) or mass spectrometry (MS) (32) for time points down to 0.4 ms (9). A potential drawback of pulsed HDX is that the alkaline conditions required for labeling (9) can interfere with folding (33). The development of novel techniques for monitoring protein structure on rapid time scales, therefore, remains an important goal (34).

Side chain covalent labeling represents a promising approach for folding investigations. Numerous reagents are available that modify solvent accessible sites, whereas burial provides protection (35). The locations of these modifications and the

extent of labeling can be determined by MS-based peptide mapping (35-37). Hydroxyl radical ( $\cdot\text{OH}$ ) (36, 38-40) represents a modifier that can be generated via photolysis of  $\text{H}_2\text{O}_2$  (41), causing oxidative modifications primarily at sulfur containing, aromatic, and aliphatic residues (36). The acronym FPOP (“fast photochemical oxidation of proteins”) has been coined for experiments where a pulsed laser is used for  $\text{H}_2\text{O}_2$  photolysis (42). The  $\cdot\text{OH}$  lifetime under typical conditions has been estimated to be  $\sim 1 \mu\text{s}$  (43). FPOP can be performed at neutral pH, thereby circumventing one of the issues encountered with pulsed HDX (33). Most importantly, FPOP holds the potential to provide detailed structural insights on a time scale that has thus far been dominated by low resolution (global) conformational probes (34).

Our laboratory was the first to employ FPOP for kinetic folding experiments (44), but the continuous-flow mixer employed in those early studies was only suitable for time points down to 10 ms (45). Yet, we pointed out that FPOP offers an avenue for submillisecond protein folding experiments, when used in conjunction with improved mixers or optical triggering (44). Gross et al. subsequently implemented a T jump strategy for submillisecond FPOP folding studies (46, 47).

Despite the demonstrated feasibility of T jump/FPOP measurements (46, 47), there is still a need for additional developments in this area. The scope of T jump experiments tends to be somewhat limited for several reasons. (i) Only proteins that undergo cold unfolding at a temperature above freezing are amenable to T jump refolding (48). (ii) It is a complication that often either the initial state is incompletely denatured (16), or the refolded state is destabilized by denaturants (19). (iii) The length of the accessible time window is restricted by temperature drifts (49). In comparison, rapid

mixing represents a trigger that is more widely applicable because most proteins can be refolded via pH jumps or by denaturant dilution (11, 16). Mixing provides access to much wider time windows for capturing entire folding pathways, although this sometimes requires the use of different mixers for different time regimes (11). Very recently, Wu and Lapidus (50) combined FPOP with submillisecond mixing. That work was the first to demonstrate the feasibility of such an approach, but the experiments were limited to global (intact protein) measurements.

Here, we employ FPOP with rapid mixing and peptide mapping for spatially-resolved folding experiments in the range of submilliseconds to seconds. We focus on the 153 residue protein apo-myoglobin (aMb). Similar to its heme-bound form (holomyoglobin, hMb), aMb adopts a compact globular fold at neutral pH. The seven helices in native aMb (N) are designated as A-E, G, and H. Helix F is partially disordered (51). The acid-denatured state (U) at pH 2 is extensively unfolded (51). HDX suggests that folding proceeds via two sequential obligatory on-pathway intermediates ( $I_1$  and  $I_2$ ) that are highly dynamic (9, 52). After 0.4 ms hydrogen bonding is detectable in helices A, G, and H ( $I_1$ ). Hydrogen bonding then spreads to B, E, and C, generating  $I_2$  after 6 ms. This is followed by formation of the ABCDE(F)GH native state (9, 53).

Although aMb has long served as paradigm for folding investigations (5, 49, 51), its submillisecond behavior has not been fully elucidated yet. Recent fluorescence data indicate the formation of several early intermediates (54), calling for an expansion of the classical (52)  $U \rightarrow I_1 \rightarrow I_2 \rightarrow N$  scenario. SAXS revealed a rapid collapse that is much larger than expected for an AGH core (30), indicating the early involvement of additional helices. Specifically, SAXS suggests that helix E acts as a hydrophobic nucleus that

drives submillisecond folding (30), whereas HDX indicates that E is formed much later (9, 55). The rapid mixing/FPOP strategy employed here is well suited for addressing unresolved questions related to the aMb folding mechanism. By monitoring the solvent accessibility of individual protein segments we track the temporal evolution of structure from a hitherto unutilized vantage point.

## 3.2 Experimental Section

### 3.2.1 Materials

As in earlier submillisecond investigations (16, 30, 49) horse skeletal muscle aMb was used for this study. aMb was produced by heme removal from hMb (Sigma, St. Louis, MO) using butanone extraction (56). The resulting solution was dialyzed against dilute aqueous HCl (pH 2) at 4 °C for ketone removal. Several batches were pooled and lyophilized. Resuspension of aMb in aqueous HCl (pH 2) provided 2 mM stock solutions that were centrifuged at 13000 *g* for 15 min to remove trace amounts of insoluble material. Concentrations were verified optically using  $\epsilon_{280} = 13,980 \text{ M}^{-1} \text{ cm}^{-1}$  (57). Glutamine, methionine, catalase, and dextran-coupled fluorescein (DCF, average MW 20,000 Da) were from Sigma. Bradykinin was supplied by Bachem (King of Prussia, PA). RapiGest acid labile surfactant came from Waters (Milford, MA,) and sequencing-grade trypsin was from Promega (Madison, WI). All experiments were conducted at room temperature and all solutions were at pH 7, unless noted otherwise.

### 3.2.2 Submillisecond Mixer Design

Time points between 0.2 ms and 6 ms were studied using a coaxial laminar flow mixer (Figure 3.1A). A similar design has previously been described for optical investigations (14). Two-dimensional hydrodynamic focusing causes the compression of a central liquid stream into a thin filament by a coaxial sheath flow, ensuring submillisecond diffusive mixing (14-16). The device was constructed in-house using fused-silica capillaries (Polymicro Technologies, Phoenix, AZ) and a PEEK tee that was fitted with 1/16" ferrule/nut connectors (Upchurch, Oak Harbor, WA). An inner capillary (i.d. 20  $\mu\text{m}$ , o.d. 90  $\mu\text{m}$ , TSP020090) with a round cross section was passed through the mixing tee for delivering the central stream. This capillary was connected to a gas-tight 1 mL syringe (SGE, Melbourne, Australia) containing unfolded aMb (or DCF for test experiments). The plunger of the syringe was advanced at 6  $\mu\text{L min}^{-1}$  by a syringe pump (Harvard Apparatus, Boston, MA). The outlet of the inner capillary was cut at a 90° angle to ensure a symmetrical flow pattern. This inner capillary was inserted into an outer capillary with a square cross section (i.d. 100  $\mu\text{m} \times 100 \mu\text{m}$ , o.d. 363  $\mu\text{m}$ , WWP100375). The outer capillary delivers the sheath flow (refolding buffer, or KI quencher for fluorescence experiments), introduced via the second inlet of the tee. Starting from the exit of the inner capillary, the polyimide coating of the outer capillary was removed using a butane torch to create a transparent window. The sheath flow was introduced using a Waters 1525  $\mu\text{HPLC}$  pump at 594  $\mu\text{L min}^{-1}$ , for a total flow rate of  $F = 600 \mu\text{L min}^{-1}$  and a nominal mixing ratio of 1:99. Laminar flow past the inner capillary outlet exhibits a parabolic velocity profile, where the velocity  $v_0$  of the central filament is  $v_0 = 2.096 \times v_{avg}$  (58). The average flow velocity  $v_{avg}$  is given by  $F/A$ , where  $A$  is the cross sectional area of the outer

capillary. For the flow rates and mixer geometry used here,  $v_{avg} = 1 \text{ m s}^{-1}$ , and hence  $v_0 = 2.096 \text{ m s}^{-1}$ . The reaction time  $t$  in the central filament is given by  $t = d / v_0$ , where  $d$  is the distance downstream of the inner capillary outlet. A reaction time of 0.2 ms corresponds to  $d = 0.42 \text{ mm}$ . This distance-time correlation does not consider the time interval  $\Delta t_{mix}$  required for the actual mixing step. It will be seen that this approximation is justified because  $t \gg \Delta t_{mix}$  for our folding experiments.

### 3.2.3 Mixer Characterization

For verifying system performance the laminar flow mixer was mounted on the sample stage of an Axiovert 25 fluorescence microscope (Zeiss, Jena, Germany). Test experiments were conducted by pumping 1 mM DCF in 10 mM phosphate buffer through the inner capillary. The 20 kDa molecular weight of DCF ensures Stokes diffusion properties similar to those of a typical protein. Fluorescence excitation was performed using a mercury lamp. Emission was monitored via a green reflector and a CCD camera (QImaging, Surrey, BC, Canada). Northern Eclipse software was used for data acquisition and analysis. Preliminary tests (not shown) revealed that DCF fluorescence was quenched (59) to  $< 10\%$  when exposed to 2.5 M KI. Mixing tests were thus conducted with a sheath flow consisting of 2.5 M KI in phosphate buffer.  $\text{I}^-$  diffusion into the central filament causes fluorescence quenching down to a constant background level within  $\sim 0.2 \text{ mm}$  past the inner capillary outlet (Figure 3.1B). A normalized fluorescence intensity  $F_{norm}$  was calculated as a function of distance  $d$  according to (11)

$$F_{norm}(d) = \frac{I(d) - I_q(d)}{I_0(d) - I_q(d)} \quad 3.1$$

where  $I(d)$  is the fluorescence intensity integrated across the channel cross section from Figure 3.1B, and  $I_0(d)$  is the corresponding profile acquired in the absence of quencher (not shown).  $I_q(d)$  represents the background signal, measured on a pre-quenched sample with 2.5 M KI in both the inner and the outer capillary (not shown). The  $F_{norm}(d)$  plot generated from Equation 3.1 reveals that diffusive mixing is complete within 0.1 ms (Figure 3.1C). Additional control experiments to verify proper mixer performance under FPOP conditions are discussed below.

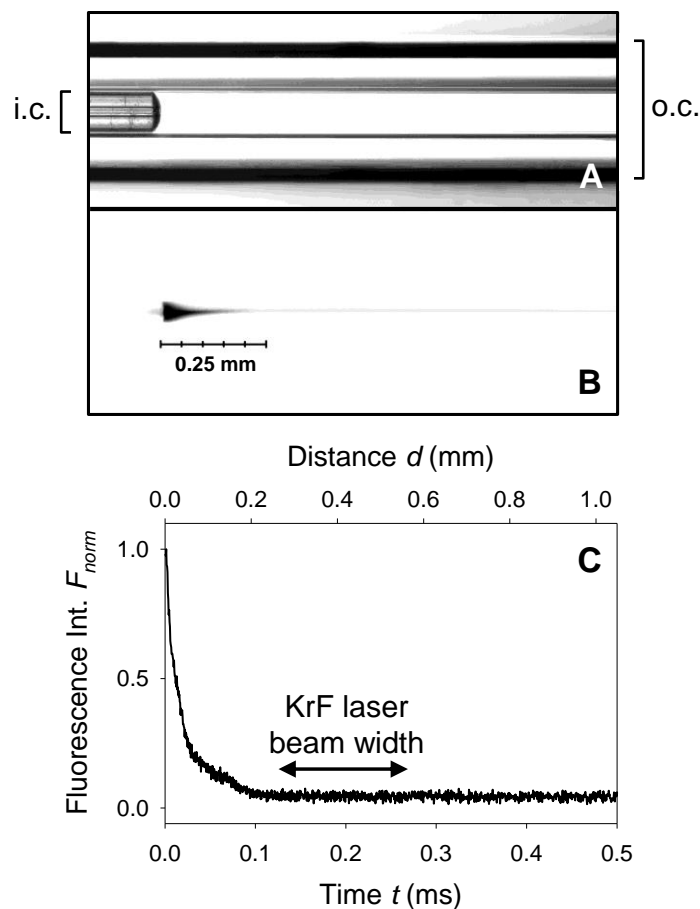


Figure 3.1 Design and characterization of the submillisecond laminar flow mixer. (A) Bright field image; i.c., inner capillary; o.c., outer capillary. (B) Fluorescence image of the hydrodynamically focused flow. The inner stream contains dextran-conjugated fluorescein. High fluorescence levels appear in black. The sheath flow contains 2.5 M KI quencher. Quenching is evident from the fading of the central filament within 0.1 ms. (C) Normalized fluorescence intensity (eq 1). The double-headed arrow represents the approximate width of the FPOP laser beam.

### 3.2.4 Oxidative Labeling

Reaction times up to  $t = 6$  ms were examined by FPOP using the laminar flow mixer of Figure 3.1. The device was mounted on a custom-built xyz micromanipulator. Different time points were studied by moving the device relative to the beam of a pulsed KrF excimer laser (GAM EX 50, Orlando, FL) that produced  $\cdot\text{OH}$  by  $\text{H}_2\text{O}_2$  photolysis. The



laser generated 18 ns pulses at 700 Hz, 248 nm, and 35 mJ. The laser was focused onto the capillary using a 500 mm convex lens, with a lens-capillary distance of 54.5 cm. Knife-edge profiling revealed a beam width of  $\sim 0.3$  mm at the irradiation spot. Geometric considerations imply that the earliest time point accessible under these conditions is  $t \approx 0.2$  ms (Figure 3.1C).

As an additional test for the fidelity of the laminar flow mixer we conducted FPOP on native aMb, using pH 7 in the inner capillary and in the sheath flow. Due to the absence of a pH jump for these tests the protein conformation remains unchanged as it travels through the mixer. Any inhomogeneities in  $\text{H}_2\text{O}_2$  or protein concentration will manifest themselves as differences in the oxidation pattern (44). Figure 3.2 shows that spectra acquired for time points between 0.2 ms and 6 ms are indistinguishable, thereby confirming effective and reproducible mixing.

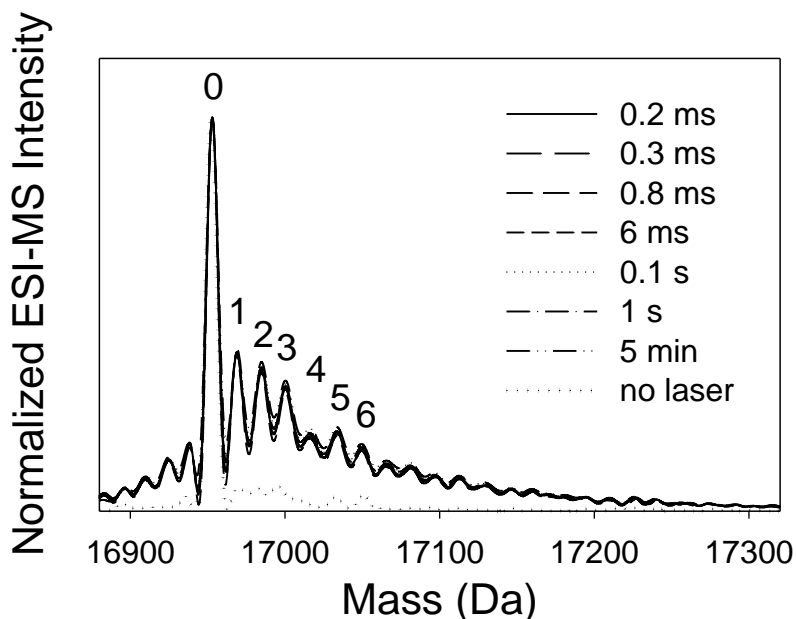


Figure 3.2 Deconvoluted mass spectra of native aMb mixed with buffer and  $\text{H}_2\text{O}_2$  at pH 7 after FPOP at various time points. Data up to  $t = 6$  ms were recorded using the laminar flow mixer of

Figure 3.1. These spectra verify complete and consistent mixing. Spectra for  $t = 0.1$  s and longer were acquired using a conventional mixer, as discussed in Section 3.2. Numbers indicate how many oxygen atoms (+16 Da adducts) were incorporated. Also shown are data for a sample exposed to the same conditions as in FPOP experiments, but without laser exposure, demonstrating that background oxidation is minimal.

Folding of aMb was triggered by a pH jump from 2 to 7 (9). The solution of denatured protein emanating from the inner capillary contained aqueous HCl (pH 2), 15 mM Gln (as radical quencher) (43), 150 mM NaCl, and 1 mM aMb. Protein concentrations on the order of 1 mM and higher are routinely used in NMR spectroscopy (60). The sheath flow consisted of 10 mM sodium phosphate buffer at pH 7, 150 mM NaCl, 15 mM Gln, and 0.1% (v/v) (~30 mM) H<sub>2</sub>O<sub>2</sub>. The solution collected at the mixer outlet had a protein concentration of 10  $\mu$ M, taking into account the 1:99 mixing ratio.

Folding times beyond 6 ms were studied using a conventional two-syringe continuous-flow mixer (45). Syringe 1 (20  $\mu$ L min<sup>-1</sup>) contained 20  $\mu$ M aMb in aqueous HCl at pH 2, 15 mM Gln, 150 mM NaCl. Syringe 2 (20  $\mu$ L min<sup>-1</sup>) contained 10 mM phosphate buffer at pH 7.8, 15 mM Gln, 150 mM NaCl along with 0.04% (v/v) (~12 mM) H<sub>2</sub>O<sub>2</sub>, resulting in pH 7 after mixing. FPOP was performed for reaction times of 0.1 s and 1 s. For these conventional mixing experiments the conditions were adjusted to ensure that the FPOP efficiency was consistent with the laminar flow mixing data. The oxidation pattern of native aMb (without pH jump) served as benchmark for this purpose. A laser frequency of 48 Hz at 37.5 mJ and a lens-capillary distance of 47 cm (beam width ~0.8 mm) provided results that were indistinguishable from those of the laminar flow mixer (Figure 3.2). Refolded ( $t = 5$  min, pH 7) and unfolded ( $t = 0$ , pH 2) controls were prepared using manual mixing (45). “Secondary” effects are minimal, such that

alterations in FPOP patterns before and after the pH jump directly reflect protein structural changes (Chapter 2).

100  $\mu\text{L}$  capillary outflow aliquots were collected in microcentrifuge tubes containing 20  $\mu\text{L}$  10 mM phosphate buffer, 1  $\mu\text{M}$  catalase, 5  $\mu\text{M}$  Met. The collection tubes for  $t = 0$  (pH 2) samples contained an equivalent amount of NaOH, resulting in a pH value that ensures catalase activity for peroxide removal. 50  $\mu\text{L}$  of the collected samples were consumed for intact protein analyses. The remaining 50  $\mu\text{L}$  were digested with trypsin using a 1:20 (w/w) enzyme:protein ratio at 37  $^{\circ}\text{C}$  for 24 h in the presence of 0.1% (w/v) RapiGest. Digestion was terminated by acidification using 100 mM HCl to a final protein concentration of 5  $\mu\text{M}$ . Samples were flash frozen in liquid nitrogen and stored at -80  $^{\circ}\text{C}$ .

### *3.2.5 LC/ESI-MS and Data Analysis*

MS data were acquired on a Waters Q-TOF Ultima API mass spectrometer. Chromatographic separation was achieved by interfacing the instrument to a Waters Acquity UPLC system, employing a C4 (BEH300) 2.1 mm  $\times$  50 mm column for intact protein analyses, and a C18 (BEH130) 2.1 mm  $\times$  100 mm column for peptide measurements. Elution was carried out using a water/acetonitrile gradient with 0.1% formic acid at 100  $\mu\text{L min}^{-1}$ . Tryptic digestion of aMb resulted in 12 detectable peptides. Peptide numbering and sequence range are as follows: T1(1-16), T2(17-31), T3(32-42), T7(51-56), T8(57-62), T10(64-77), T13(80-96), T16(103-118), T17(119-133), T18(134-139), T19(140-145), T21(148-153) (Chapter 2), for a sequence coverage of ~90%. MS/MS revealed 39 oxidized residues in unfolded aMb after FPOP, comprising M, W, F,

H, V, I, L, E, P, K, and R as discussed in (Chapter 2). Quantitation of protein oxidation was conducted at the peptide level on the basis of background corrected “fraction unmodified” ( $F_u$ ) values. A detailed description of this procedure has been provided previously (44, 45) (Chapter 2). Briefly, the strategy relies on intensity measurements of unoxidized peptides relative to a bradykinin internal standard. By focusing on the signals of unoxidized species, the method avoids complications associated with the detection of heterogeneous oxidation products and differential ionization efficiencies. 5  $\mu$ M bradykinin was spiked into each sample prior to LC-MS. The normalized oxidation level ( $NOL$ ) of each peptide was determined as a function of folding time  $t$  according to

$$NOL(t) = \frac{\ln F_u(t) - \ln F_u(\text{native aMb})}{\ln F_u(0) - \ln F_u(\text{native aMb})} \quad 3.2$$

By using acid-unfolded ( $t = 0$ ) and native aMb as reference points, Equation 3.2 ensures that  $NOL$  values fall between zero and unity.  $NOL = 1$  reflects the case where a peptide has the same average solvent accessibility as in the acid-unfolded state. Conversely,  $NOL = 0$  applies to peptides that exhibit the same accessibility as in native aMb. The logarithmic normalization used in Equation 3.2 ensures that differences in intrinsic reactivity of the various peptides cancel out, such the  $NOLs$  of different protein segments are directly comparable (44). Equation 3.2 can also be used to quantify conformational changes at the intact protein level. All  $NOLs$  represent an average of three independent measurements. Error bars correspond to standard deviations.

### 3.3 Results and Discussion

Changes in solvent accessibility during aMb refolding were probed by exposing the protein to FPOP at various time points following a pH jump. We monitored reaction

times between 0.2 ms and 6 ms using the laminar flow mixer of Figure 3.1. Conventional mixing and manual techniques were employed for longer reaction times, as outlined in the Section 3.2.

### 3.3.1 Global Solvent Accessibility Changes

Prior to discussing spatially-resolved data, it is instructive to consider solvent accessibility changes at the intact protein level. FPOP of acid-denatured aMb yields a mass distribution that exhibits a series of satellite peaks (Figure 3.3A,  $t = 0$ ). The 16 Da spacing of these signals reflects side chain oxygen incorporation (36). The prevalence of these modifications reflects the high solvent accessibility of the acid-denatured state (Chapter 2). A significant change in the mass envelope is observable after 0.2 ms (Figure 3.3B). Labeling progressively decreases with increasing folding time, and after 100 ms (Figure 3.3C) the protein shows only slightly more oxidation than in the native state (Figure 3.2D,  $t = 5$  min).

For expressing changes in global solvent accessibility in a quantitative fashion, Equation 3.2 was applied to the intact protein data. The *NOL* profile obtained in this way (Figure 3.4A) drops to 0.75 within 0.8 ms, confirming the formation of partially protected elements on the submillisecond time scale. A further decrease to  $NOL = 0.16$  has taken place after 100 ms, which indicates that most of the  $U \rightarrow N$  conformational changes are complete by this time. This ~100 ms time scale is consistent with the results of CD and SAXS experiments (30).

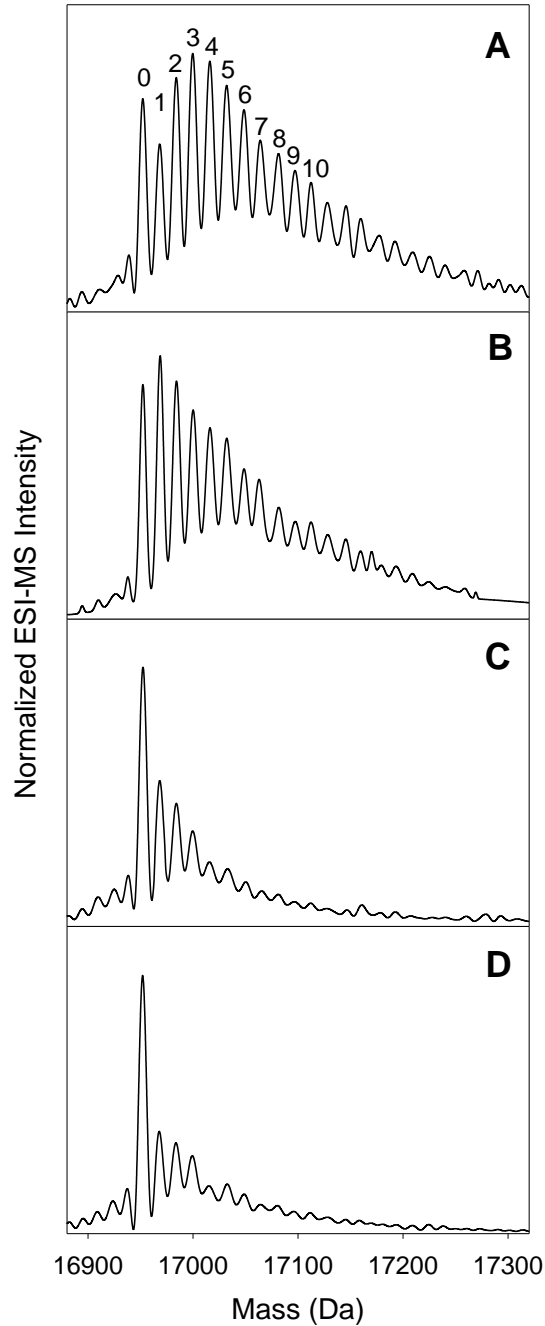


Figure 3.3 Deconvoluted mass distribution obtained by exposing aMb to FPOP at different time points during folding: (A) acid denatured ( $t = 0$ ); (B) 0.2 ms; (C) 0.1 s; and (D) 5 min. Numbers in (A) indicate how many oxygens (+16 Da adducts) were incorporated into protein side chains.

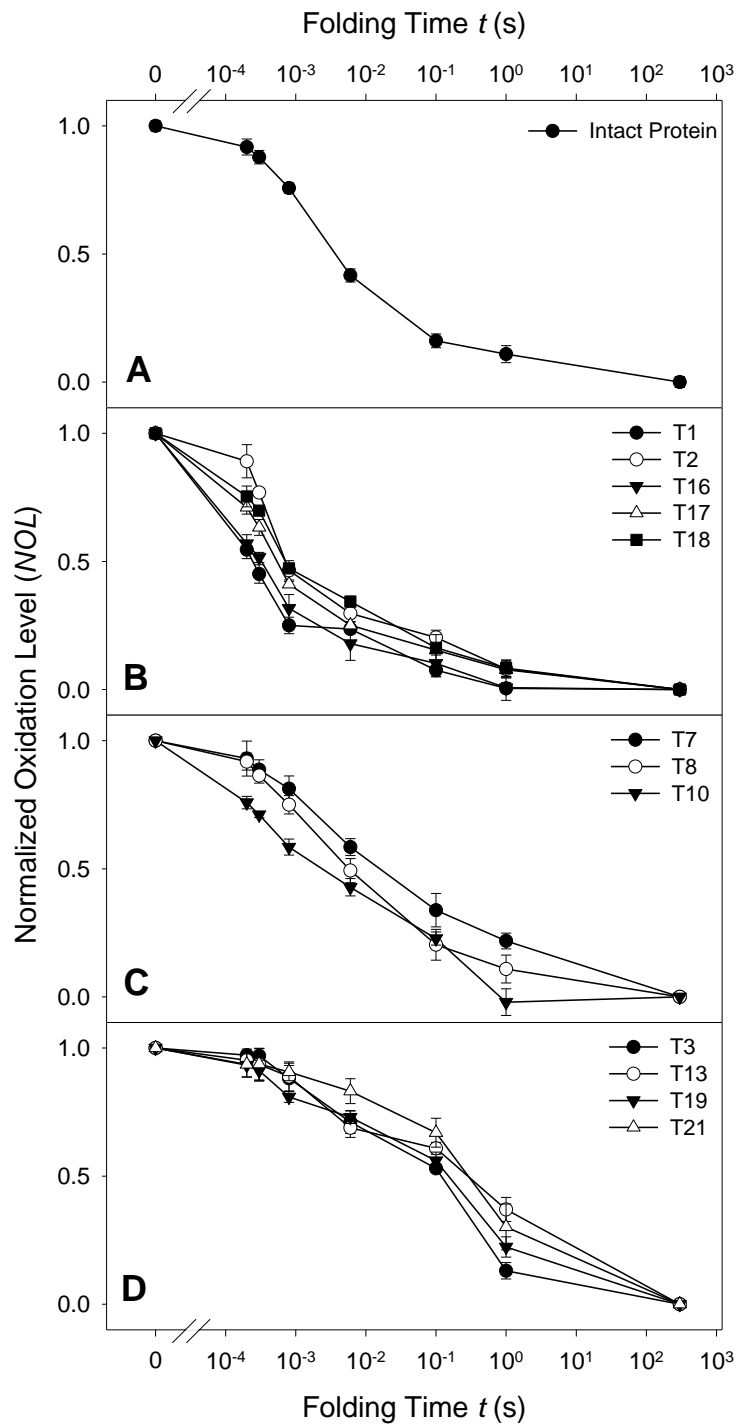


Figure 3.4 Normalized oxidation level (*NOL*) of (A) the intact protein, and (B-D) individual segments (tryptic peptides T1, T2,...) at different folding times. Peptide data were separated to prevent cluttering. The time axis has a log scale, where “0” corresponds to the denatured state.

### 3.3.2 Solvent Accessibility of Individual Protein Segments

For examining the folding mechanism in a spatially-resolved fashion the aMb samples were proteolyzed after FPOP. The *NOL* approach used here quantifies the average solvent accessibility at the level of individual peptides (Chapter 2). T1, T2, T16, T17, and T18 gain significant protection early on, and have dropped to  $NOL < 0.5$  after only 6 ms (Figure 3.4B). Other segments (T3, T13, T19, and T21, Figure 3.4D) exhibit slower kinetics and do not reach the  $NOL = 0.5$  mark until  $\sim 1$  s after the onset of folding. The remaining peptides (T7, T8, and T10, Figure 3.4C) exhibit a behavior in-between the two cases. Among these, T10 stands out in that it gets protected more rapidly than the other segments.

Using a five-level color code, the measured *NOLs* were mapped onto the crystal structure of hMb to facilitate a structural interpretation of the FPOP data (Figure 3.5). Segments exhibiting a solvent accessibility close to that in unfolded aMb ( $1 > NOL > 0.8$ ) are highlighted in red, whereas native-like solvent accessibilities ( $0.2 > NOL > 0$ ) are colored blue. The normalization scheme dictates that U (Figure 3.5A) and N (Figure 3.5H) appear entirely red and blue, respectively.

### 3.3.3 Submillisecond Events

The 0.2 ms time point reveals significantly reduced solvent accessibility in helices A and G (yellow, Figure 3.5B). This suggests docking of the two helices as one of the first folding steps, consistent with early N/C-terminal coupling seen for other proteins (61). Formation of such sequence-remote contacts reduces the  $\phi\Psi$  space available to the chain, thereby accelerating the conformational search towards the native state (62). X-ray data



imply that this docking in aMb is mediated by hydrophobic side chain interactions involving I111, V114, L115 (helix G), and V13, W14, V17 (helix A) (63). Weak protection at 0.2 ms is also seen for helix E and for part of H (orange in Figure 3.5B). All other segments still remain fully accessible (red). The observation of these early events is consistent with recent aMb fluorescence data that were interpreted in terms of collapse and docking (16). The structural changes seen in Figure 3.5B evidently do not affect the entire protein. Hence, our data do not support the view (24) that the first folding step is a nonspecific overall collapse that produces an unstructured globule.

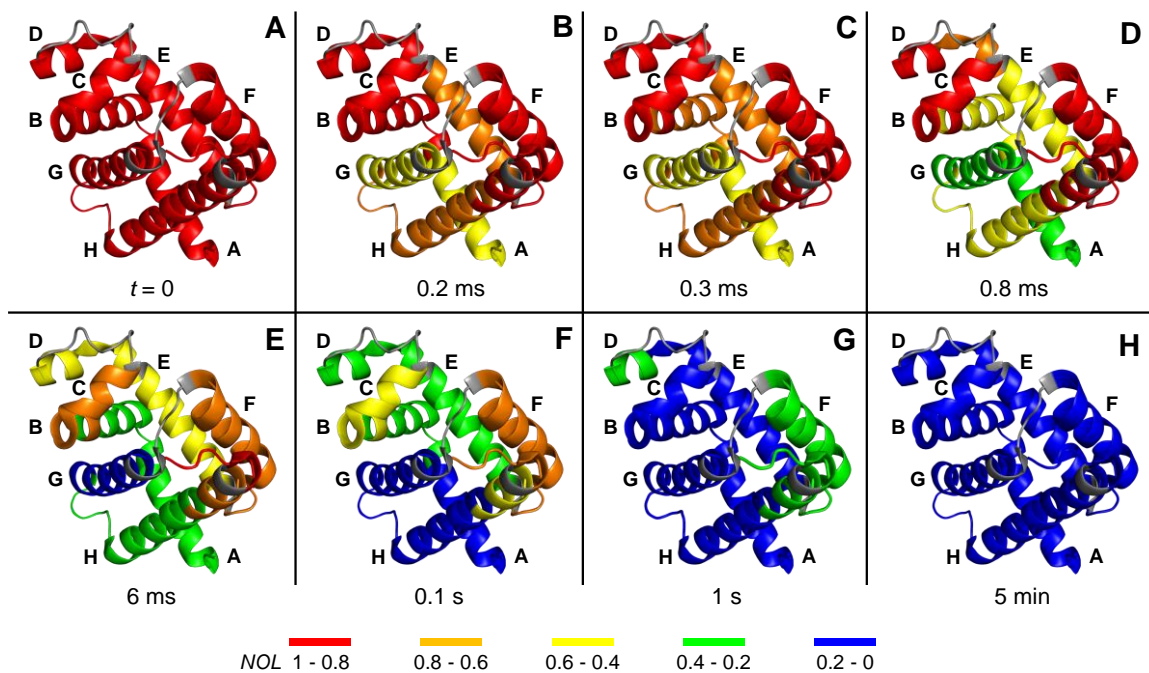


Figure 3.5 Structural changes during aMb folding as measured by FPOP. *NOL* data (from Figure 3.4) were mapped onto the crystal structure of native hMb (pdb code 1WLA)(63) using a five-level color code. Regions for which no structural data are available appear in gray.

The initial A(E)G(H) conformation seen in Figure 3.5B is somewhat different from the submillisecond AGH structure deduced from HDX data (9). FPOP reveals

partial protection in helix E after 0.2 ms, whereas HDX shows only marginal early stabilization of this segment (9, 55). This behavior illustrates the complementary nature of FPOP and HDX, *i.e.*, the ability to monitor side chain solvent accessibility and backbone hydrogen bonding, respectively (64). The involvement of helix E in the initial stages of aMb folding had been proposed on the basis of mutational (65) and spectroscopic (30) experiments, and the observations of this work lend support to this scenario. In light of existing HDX data (9) it can be concluded that partial clustering and docking of the helix E side chains precedes the formation of backbone hydrogen bonds in this segment, reminiscent of reports for other proteins (22, 24, 66). Helix B starts to show protection at a slightly later stage (orange, Figure 3.5C). After 0.8 ms FPOP reveals the presence of a fairly well developed ABEGH structure, where A and G have reached significant protection with *NOLs* around 0.3 (green in Figure 3.5D).

It is a long-standing question whether submillisecond structure formation is primarily driven by hydrophobic side chain interactions or backbone hydrogen bonding (22-25). FPOP is uniquely positioned to address this issue. Hydrophobic residues are characterized by positive Kyte/Doolittle scores, whereas hydrophilic elements sites exhibit negative values (Figure 3.6).(67) A hydrophobically-driven scenario will show early protection from  $\cdot\text{OH}$  attack for nonpolar residues. We will focus on  $t = 0.8$  ms, because this time point represents the culmination of the preceding submillisecond events. Figure 3.7 shows a plot of *NOL* vs. hydrophobicity for  $t = 0.8$  ms. There is an obvious trend for the points in Figure 3.7 to segregate into two groups; bottom right (hydrophobic & protected) and top left (hydrophilic & solvent accessible). In other words, there is a clear tendency where submillisecond structure formation is promoted by

hydrophobic side chains. Conversely, hydrophilic regions tend to be excluded from early folding events.

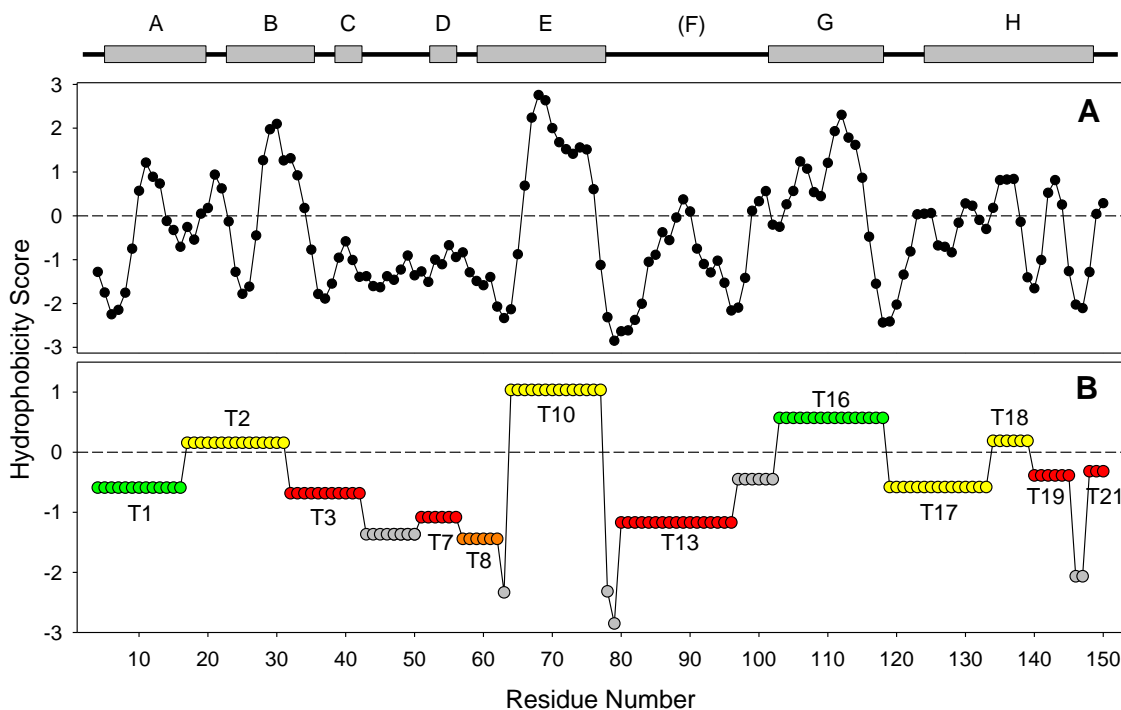


Figure 3.6 (A) Kyte/Doolittle hydrophobicity plot of aMb generated by ExPASy ([expasy.org/protscale](http://expasy.org/protscale)). A window size of 7 and a triangular weight variation model was applied. (B) Average Kyte/Doolittle hydrophobicity of the 12 aMb tryptic peptides. Positive values represent hydrophobic segments, negative values correspond to hydrophilic regions. The coloring reflects *NOLs* for a folding time of 0.8 ms (from Figure 3.4D).

T1 and T17 are apparent outliers to the observed trend (Figure 3.7, open symbols). Despite its low *average* hydrophobicity, T1 contains several nonpolar residues (Figure 3.6). It seems likely that the early behavior of T1 is governed by these hydrophobic sites. Specifically, docking of T1 (helix A) against T16 (helix G) is mediated by sequence-remote nonpolar contacts, as discussed in detail above. A somewhat different scenario applies to T17 (GH loop and half of helix H). T17 is situated

between two segments that are hydrophobic, and that show pronounced early protection (T16 and T18, Figure 3.6). The T16-T18 region comprises much of the G-turn-H motif, which is stabilized by nonpolar contacts between L103, I111, L115, F123, M131, L135, and F138 (63, 68). Only two of these residues (F123 and M131) are part of T17, but it appears that their participation in the helix-turn-helix motif is sufficient to induce early stabilization of T17.

In summary, our data imply that submillisecond events during aMb folding are driven primarily by the formation of local and sequence-remote hydrophobic side chain contacts. The alternative, *i.e.*, a scenario driven by backbone hydrogen bonding would *not* be expected to show a correlation between early protection and hydrophobicity. The undeniable presence of such a correlation (Figure 3.7) is the hallmark of a folding mechanism that is based on hydrophobicity. This view is consistent with recent HDX work where hydrogen bonding was interpreted as a consequence, rather than the cause, of hydrophobic contact formation (9). Our conclusions are further corroborated by experiments on isolated aMb fragments, none of which show a strong tendency to independently form helices (with possible exception of H) (68).

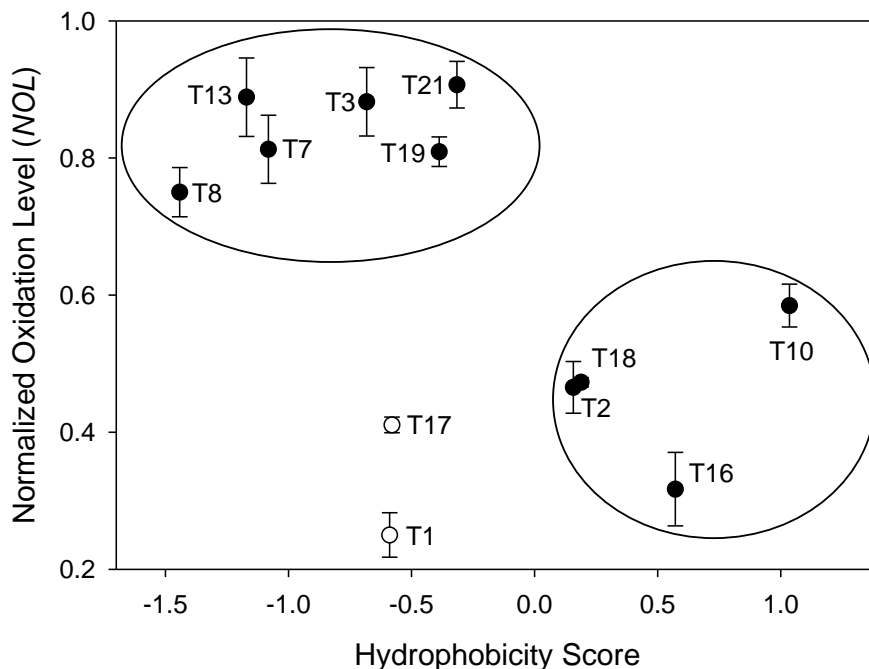


Figure 3.7 Plot of *NOL* vs. average Kyte/Doolittle hydrophobicity score for the aMb segments at  $t = 0.8$  ms. The two ellipsoids highlight the trend of the data points to cluster into two groups. The behavior of T1 and T17 (open symbols) is discussed in the text.

### 3.3.4 Folding Events Beyond 1 ms

After 6 ms helix G exhibits a native-like solvent accessibility (blue, Figure 3.5E). G is framed by A, B, and H, all of which show significant protection as well, albeit at a somewhat lower level (green). Figure 3.5E also reveals the onset of structure formation for D (yellow) and F (orange). This  $t = 6$  ms labeling behavior is remarkably similar to that observed in a recent equilibrium FPOP study at pH 4 (Chapter 2), consistent with the view that the pH 4 molten globule of aMb resembles a kinetic intermediate.<sup>(9, 52, 53)</sup>

The time-dependent changes in solvent accessibility seen in Figure 3.5E-H clearly highlight the role of helix G as central nucleus that promotes structure formation of adjacent regions. As noted, G is the first to reach native-like protection ( $t = 6$  ms, Figure 3.5E). After 0.1 s helices A and H are packed against G, with native *NOLs* for

almost the entire AGH cluster (blue, Figure 3.5 F). At  $t = 1$  s native packing interactions have spread to B, C, and E (blue, Figure 3.5G). The last elements that still show small deviations from the native aMb structure after 1 s are parts of helices D and F, as well as the terminus of H (green). This stepwise structural consolidation, evident from the accumulation of “blue” adjacent elements in Figure 3.5E-H, is reminiscent of nucleation processes reported for a range of other proteins (69).

### 3.4 Conclusions

Early stopped-flow investigations revealed the presence of a burst-phase during folding of many proteins, corresponding to early conformational changes that were beyond the millisecond time resolution of the experiment (53). The subsequent development of improved mixing technology and laser-based triggering allowed the *kinetics* of these submillisecond events to be resolved by monitoring changes in spectroscopic signatures (11-21). For many of those studies, however, the *exact structural nature* of the initial folding events remained open to interpretation due to the global nature of the spectroscopic probes used.

The current work demonstrates that the combination of FPOP and laminar flow mixing provides detailed structural insights all the way down to the submillisecond range. By monitoring time-dependent changes in solvent accessibility it is possible to track which regions of a protein fold first, and how the native structure assembles in a stepwise fashion. The level of structural detail provided in this way significantly exceeds the information obtainable by traditional spectroscopy. The use of different modes of mixing can extend the time range accessible to FPOP as far as required for the protein under

investigation. In the present case, the earliest and the latest time points studied encompass six orders of magnitude. Clearly, the highest impact of FPOP will be in the submillisecond range which has thus far suffered from a scarcity of detailed structural data. A particularly exciting prospect is a side-by-side comparison of FPOP data with computer simulations that have recently begun to overlap with the time range that is now accessible to experimental investigations (70).

Our kinetic analyses were restricted to structural mapping at the peptide level, using a procedure that addresses the different intrinsic reactivities of oxidation sites (44). The development of improved MS/MS and data analysis strategies may allow an extension towards the single amino acid level. First steps in this direction have recently been described (71).

The aMb data of this work reveal a clear trend, where submillisecond folding events are driven by hydrophobic side chain interactions. Instead of causing a nonspecific collapse as sometimes suggested, the docking of nonpolar residues follows a well organized sequence – akin to a three-dimensional jigsaw puzzle that gradually proceeds towards the native state. Earlier HDX investigations inferred the formation of side chain contacts on the basis of backbone hydrogen bonding (9), whereas FPOP *directly* reports on side chain solvent accessibilities. In future experiments it will be interesting to see if other proteins exhibit submillisecond events resembling those seen here for aMb. Laser-activated labeling probes other than  $\cdot\text{OH}$  might provide an additional dimension to those studies (72).

### 3.5 References

1. Anfinsen CB (1973) Principles that Govern the Folding of Protein Chains. *Science* 181: 223-230.
2. Service RF (2008) Problem Solved\* (\*sort of). *Science* 321: 784-786.
3. Dill KA & Chan HS (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4: 10-19.
4. Yamada S, Bouley Ford ND, Keller GE, Ford WC, Gray HB, & Winkler JR (2013) Snapshots of a protein folding intermediate. *Proc. Natl. Acad. Sci. U.S.A.* 110(5): 1606-1610.
5. Baldwin RL (2008) The Search for Folding Intermediates and the Mechanism of Protein Folding. *Annu. Rev. Biophys.* 37: 1-21.
6. Brockwell DJ & Radford SE (2007) Intermediates: ubiquitous species on folding energy landscapes? *Curr. Op. Struct. Biol.* 17: 30-37.
7. Davis CM, Xiao S, Raeigh DP, & Dyer RB (2012) Raising the Speed Limit for  $\beta$ -Hairpin Formation. *J. Am. Chem. Soc.* 134: 14476-14482.
8. Tsutsui Y, Dela Cruz RG, & Wintrode PL (2012) The folding mechanism of the metastable serpin  $\alpha_1$ -antitrypsin. *Proc. Natl. Acad. Sci. U.S.A.* 109: 4467-4472.
9. Uzawa T, Nishimura C, Akiyama S, Ishimori K, Takahashi S, Dyson HJ, & Wright PE (2008) Hierarchical folding mechanism of apomyoglobin revealed by ultra-fast H/D exchange coupled with 2D NMR. *Proc. Natl. Acad. Sci. U.S.A.* 105: 13859-13864.
10. Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, & Hofrichter J (2000) Fast Kinetics and Mechanisms in Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* 29: 327-359.
11. Roder H, Maki K, & Cheng H (2006) Early Events in Protein Folding Explored by Rapid Mixing Methods. *Chem. Rev.* 106: 1836-1861.



12. Wu Y, Kondrashkina E, Kayatekin C, Matthews CR, & Bilsel O (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc. Natl. Acad. Sci. U.S.A.* 105: 13367-13372.
13. Akiyama S, Takashi S, Ishimori K, & Morishima I (2000) Stepwise formation of  $\alpha$ -helices during cytochrome c folding. *Nat. Struct. Biol.* 6: 514-520.
14. Pabit SA & Hagen SJ (2002) Laminar-Flow Fluid mixer for Fast Fluorescence Kinetics Studies. *Biophys. J.* 83: 2872-2878.
15. Knight JB, Vishwanath A, Brody JP, & Austin RH (1998) Hydrodynamic Focusing on a Silicon Chip: Mixing Nanoliters in Microseconds. *Phys. Rev. Lett.* 80: 3863-3866.
16. Lapidus LJ, Yao S, McGarrity KS, Hertzog DE, Tubman E, & Bakajin O (2007) Protein hydrophobic collapse and early folding steps observed in a microfluidic mixer. *Biophys. J.* 93(1): 218-224.
17. Dumont C, Emilsson T, & Gruebele M (2009) Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. *Nat. Methods* 6: 515-519.
18. Yang WY & Gruebele M (2003) Folding at the speed limit. *Nature* 423: 193-197.
19. Nölting B, Golbik R, & Fersht AR (1995) Submillisecond events in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 92: 10668-10672.
20. Wittung-Stafshede P, Gray HB, & Winkler JR (1997) Rapid Formation of a Four-Helix Bundle. Cytochrome b562 Folding Triggered by Electron Transfer. *J. Am. Chem. Soc.* 119: 9562-9563.
21. Tucker MJ, Courter JR, Chen J, Atasoylu O, Smith AB, III, & Hochstrasser RM (2010) Tetrazine Phototriggers: Probes for Peptide Dynamics. *Angew. Chem. Int. Ed.* 49(21): 3612-3616.
22. Haran G (2012) How, when and why proteins collapse: the relation to folding. *Curr. Opin. Struct. Biol.* 22: 14-20.

23. Udgaonkar JB (2013) Polypeptide chain collapse and protein folding. *Arch. Biochem. Biophys.* 531: 24-33.
24. Sadqi M, Lapidus LJ, & Munoz V (2003) How fast is protein hydrophobic collapse? *Proc. Natl. Acad. Sci. U.S.A.* 100: 12117-12122.
25. Rose GD, Fleming PJ, Banavar JR, & Maritan A (2006) A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 45: 16623-16633.
26. Mayor U, *et al.* (2003) The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* 421: 863-867.
27. Huang F, Sato S, Sharpe TD, Ying L, & Fersht AR (2007) Distinguishing between cooperative and unimodal downhill protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 104: 123-127.
28. Bartlett AI & Radford SE (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* 16: 582-588.
29. Fierz B, Satzger H, Root C, Gilch P, Zinth W, & Kiefhaber T (2007) Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proc. Natl. Acad. Sci. U.S.A.* 104: 2163-2168.
30. Uzawa T, Akiyama S, Kimura T, Takahashi S, Ishimori K, Morishima I, & Fujisawa T (2004) Collapse and search dynamics of apomyoglobin folding revealed by submillisecond observations of alpha-helical content and compactness. *Proc. Natl. Acad. Sci. U.S.A.* 101(5): 1171-1176.
31. Nagarajan S, Taskent-Sezgin H, Parul D, Carrico I, Raleigh DP, & Dyer RB (2011) Differential ordering of the protein backbone and side chains during protein folding revealed by site-specific recombinant infrared probes. *J. Am. Chem. Soc.* 133: 20335-20340.
32. Gu Z, Rao MK, Forsyth WR, Finke JM, & Matthews CR (2007) Structural analysis of kinetic folding intermediates for a TIM barrel protein, indole-3-glycerol phosphate synthase, by hydrogen exchange mass spectrometry and Go-model simulation. *J. Mol. Biol.* 374: 528-546.

33. Bieri O & Kiefhaber T (2001) Origin of apparent fast and non-exponential kinetics of lysozyme folding measured in pulsed hydrogen exchange experiments. *J. Mol. Biol.* 310(4): 919-935.
34. Gruebele M (2010) Weighing Up Protein Folding. *Nature* 468: 640-641.
35. Mendoza VL & Vachet RW (2009) Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* 28: 785-815.
36. Xu G & Chance MR (2007) Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* 107: 3514-3543.
37. Jha SK & Udgaonkar JB (2007) Exploring the Cooperativity of the Fast Folding Reaction of a Small Protein Using Pulsed Thiol Labeling and Mass Spectrometry. *J. Biol. Chem.* 282: 37479-37491.
38. Smedley JG, Sharp JS, Kuhn JF, & Tomer KB (2008) Probing the pH-dependent prepore to pore transition of Bacillus anthracis protective antigen with differential oxidative protein footprinting. *Biochemistry* 47: 10694-10704.
39. Sharp JS, Becker JM, & Hettich RL (2004) Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectrometry. *Anal. Chem.* 76: 672-683.
40. Maleknia SD, Wong JWH, & Downard KM (2004) Photochemical and electrophysical production of radicals on millisecond timescales to probe the structure, dynamics, and interactions of proteins. *Photochem. Photobiol. Sci.* 3: 741-748.
41. Aye TT, Low TY, & Sze SK (2005) Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry. *Anal. Chem.* 77: 5814-5822.
42. Gau BC, Sharp JS, Rempel DL, & Gross ML (2009) Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* 81: 6563-6571.

43. Hambly DM & Gross ML (2005) Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* 16: 2057-2063.
44. Stocks BB & Konermann L (2009) Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* 81: 20-27.
45. Stocks BB, Rezvanpour A, Shaw GS, & Konermann L (2011) Temporal Development of Protein Structure During S100A11 Folding and Dimerization Probed by Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* 409: 669-679.
46. Chen J, Rempel DL, & Gross ML (2010) Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* 132: 15502-15504.
47. Chen J, Cui W, Giblin D, & Gross ML (2012) New Protein Footprinting: Fast Photochemical Iodination Combined with Top-Down and Bottom-Up Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 23: 1306-1318.
48. Creighton TE (1993) *Proteins* (W. H. Freeman & Co, New York).
49. Ballew RM, Sabelko J, & Gruebele M (1996) Direct observation of fast protein folding: The initial collapse of apomyoglobin. *Proc. Natl. Acad. Sci. U.S.A.* 93: 5759-5764.
50. Wu L & Lapidus LJ (2013) Combining ultrarapid mixing with photochemical oxidation to probe protein folding. *Anal. Chem.* 85: 4920-4924.
51. Eliezer D, Yao J, Dyson HJ, & Wright PE (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* 5: 148-155.
52. Jamin M & Baldwin RL (1998) Two Forms of the pH4 Folding Intermediate of Apomyoglobin. *J. Mol. Biol.* 276: 491-504.
53. Jennings PA & Wright PE (1993) Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science* 262: 892-896.

54. Xu M, Beresneva O, Rosario R, & Roder H (2012) Microsecond Folding Dynamics of Apomyoglobin at Acidic pH. *J. Phys. Chem. B* 116: 7014-7025.
55. Dyson HJ, Wright PE, & Scheraga HA (2006) The role of hydrophobic interactions in initiation and propagation of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* 103: 13057-13061.
56. Teale FWJ (1959) Cleavage of the heme-protein link by acid methylethylketone. *Biochim. Biophys. Acta* 35: 543.
57. Antonini E & Brunori M (1971) *Hemoglobin and Myoglobin in Their Reactions With Ligands* (North-Holland Publishing Company, Amsterdam, London).
58. Shah RK & London AL (1978) Rectangular Ducts. *Advances in Heat Transfer*, eds Irvine TF & Hartnett JP (Academic Press, New York, San Francisco, London), pp 196-222.
59. Chan C-K, Hu Y, Takahashi S, Rousseau DL, Eaton WA, & Hofrichter J (1997) Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl. Acad. Sci. U.S.A.* 94: 1779-1784.
60. Vallurupalli P, Bouvignies G, & Kay LE (2012) Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *J. Am. Chem. Soc.* 134(19): 8148-8161.
61. Krishna MMG & Englander SW (2005) The N-terminal to C-terminal motif in protein folding and function. *Proc. Natl. Acad. Sci. U.S.A.* 102: 1053-1058.
62. Lin MM & Zewail AH (2012) Hydrophobic forces and the length limit of foldable protein domains. *Proc. Natl. Acad. Sci. U.S.A.* 109: 9851-9856.
63. Maurus R, Overall CM, Bogumil R, Luo Y, Mauk AG, Smith M, & Brayer GD (1997) A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket. *Biochim. Biophys. Acta* 1341: 1-13.
64. Zheng X, Wintrode PL, & Chance MR (2008) Complementary Structural Mass Spectrometry Techniques Reveal Local Dynamics in Functionally Important Regions of a Metastable Serpin. *Structure* 16: 38-51.

65. Nishimura C, Dyson HJ, & Wright PE (2006) Identification of Native and Non-Native Structure in Kinetic Folding Intermediates of Apomyoglobin. *J. Mol. Biol.* 355: 139-156.
66. Stocks BB, Sarkar A, Wintrode PL, & Konermann L (2012) Early Hydrophobic Collapse of alpha(1)-Antitrypsin Facilitates Formation of a Metastable State: Insights from Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* 423(5): 789-799.
67. Kyte J & Doolittle R (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157: 105-132.
68. Reymond MT, Merutka G, Dyson HJ, & Wright PE (1997) Folding propensities of peptide fragments of myoglobin. *Protein Sci.* 6: 706-716.
69. Nölting B & Agard DA (2008) How general is the nucleation-condensation mechanism? *Proteins* 73: 754-764.
70. Lindorff-Larsen K, Piana S, Dror RO, & Shaw DE (2011) How Fast-Folding Proteins Fold. *Science* 334: 517-520.
71. Chen J, Rempel DL, Gau B, & Gross ML (2012) Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level. *J. Am. Chem. Soc.* 134: 18724-18731.
72. Jumper CC, Bomgarden R, Rogers J, Etienne C, & Schriemer DC (2012) High-Resolution Mapping of Carbene-Based Protein Footprints. *Anal. Chem.* 84: 4411-4418.

## Chapter 4: Partially Disordered Proteins Studied by Ion Mobility-Mass Spectrometry: Implications for the Preservation of Solution Phase Structure in the Gas Phase

### 4.1 Introduction

Electrospray ionization (ESI) mass spectrometry (MS) (1) plays a key role for investigating the structure and dynamics of proteins (2-6). Although ESI-MS involves the detection of analyte ions *in the gas phase*, most studies employing this technique aim to characterize the behavior of proteins *in solution*. Depending on the experimental design, this phase transition can have implications for the results obtained. The issue is not of great concern for studies that simply require a mass readout following solution phase covalent modifications, such as backbone H/D exchange (4), side chain labeling (5), or cross-linking (6).

The situation is more intricate for experiments that rely on the preservation of structure and noncovalent interactions in the gas phase. Ion mobility (IM) MS represents an example of this second type of ESI-MS-based experiment (7-9). In IM-MS, electrosprayed ions are passed through a buffer gas under the influence of a weak electric field. The drift time  $t_d$  of analyte ions depends on  $\Omega/z$ , where  $z$  is the charge state and  $\Omega$  is the rotationally averaged collision cross section (CCS). The latter represents a measure of analyte size (7, 10, 11). Unfolded protein ions have larger  $\Omega$  values than tightly folded conformers, such that IM-MS provides structural information on gaseous biomolecular systems (10, 11). In recent years, IM-MS has found widespread use due to the availability

of commercial time-of-flight systems that allow IM separations via travelling-wave ion guides (TWIGs) (12, 13).

The question to what extent gaseous protein ions retain memory of their solution phase structures remains controversial (14). Native globular proteins in aqueous solution possess a hydrophobic core that is shielded from the solvent. Hydrophilic side chains on the surface are solvated by water (15). ESI causes desolvation as well as major changes in the protonation state of titratable groups (16, 17). It has been proposed that “inside out” conformations represent the most stable gaseous state, characterized by a hydrophilic core and a hydrophobic exterior (18-20). Some investigations support the occurrence of major structural changes upon transfer into the gas phase (21-25), while others suggest that proteins survive the ESI process largely unharmed (26-28). Evidence for the retention of native-like solution phase structures comes from the observation that specific protein-protein (29-32) and protein-ligand interactions (30, 33) can be preserved in the gas phase. In several cases it was demonstrated that the CCSs of gaseous protein complexes are consistent with solution phase conformations (34, 35). However, there are also examples of systems where solution phase interactions are lost during ESI (36, 37).

The picture emerging from these findings seems to be that desolvated biomolecular ions can get trapped in local free energy minima. Under optimized conditions these metastable states can retain aspects of their native solution phase structure, at least on the millisecond time scale of IM-MS experiments (14, 26, 38, 39). Extended ion storage, collisional activation, and gas phase proton transfer promote the loss of solution phase memory (24, 40, 41).



The preceding considerations focused on proteins that adopt a compact folded structure in solution prior to ESI. In addition, there is growing interest in the characterization of partially disordered solution phase species (42-47). Such conformers can be populated by exposing proteins to mildly acidic pH and/or organic cosolvents (43, 44). Non-native proteins play a key role as folding intermediates (48), and as precursors of cytotoxic aggregates (49). Many proteins that participate in cellular regulation processes are intrinsically disordered (50). The biophysical characterization of all these partially disordered species by traditional analytical methods remains difficult (51). Considering the widespread use of IM-MS for investigating native proteins and protein complexes, it is important to delineate whether this technique is an equally powerful tool for the characterization of partially disordered conformers (26, 45-47).

Several challenges can be envisioned when studying partially disordered proteins in the gas phase. (1) The relatively low number of intramolecular contacts (50) suggests a high vulnerability to structural changes upon transfer into the gas phase. (2) The extensive structural heterogeneity complicates comparisons of solution phase and gas phase structures. (3) Ions generated from partially disordered chains tend to have higher ESI protonation states than those generated from tightly folded conformers (52, 53). Coulombic repulsion in these highly charged ions favors the transition to elongated conformations, thereby possibly disrupting residual native-like elements (7, 26, 41, 54). (4) Repulsive charge-charge interactions notwithstanding, it has been reported that some disordered polypeptide chains can collapse into compact structures after ESI (55).

On the basis of the above considerations it seems uncertain whether IM-MS of partially disordered proteins can provide structural information that is relevant in a solution phase

context. The current work examines this issue using myoglobin (Mb) as model system. Mb is uniquely suited for these investigations, because its solution conformation can be manipulated in a highly controlled fashion. The stepwise unfolding of Mb has been studied by NMR spectroscopy (56, 57), H-exchange (58, 59), and covalent labeling (Chapter 2). Native holo-Mb (hMb) at pH 7 binds a heme group and adopts a globular structure that comprises 8  $\alpha$ -helices A-H (60). Extraction of heme from the protein yields apo-myoglobin (aMb). The structures of hMb and aMb at pH 7 are similar, with exception of helix F which is partially unfolded in aMb (56, 58). At pH 4 aMb adopts a highly dynamic structure, with a partially intact A[B]GH core (56, 59)(Chapter 2). Acid-denatured aMb at pH 2 is extensively unfolded, although regions A and H retain limited transient helicity (57). Radii of gyration ( $R_g$ ) of the four Mb conformers have been reported as 18, 19, 23, and 30 Å (56). Mb has previously been examined by IM-MS (41, 54, 61, 62), but it remains unexplored how CCSs change along the hMb<sub>pH 7</sub>  $\rightarrow$  aMb<sub>pH 7</sub>  $\rightarrow$  aMb<sub>pH 4</sub>  $\rightarrow$  aMb<sub>pH 2</sub> unfolding pathway.

Our results support the view that native hMb retains a native-like structure after transfer into the gas phase. In contrast, the partially disordered species examined here (aMb at pH 7 and pH 4) generate gaseous ions with CCSs quite similar to those of the unfolded state. Some ESI charge states exhibit solution phase-dependent CCSs, but overall there is no clear relationship between the degree of unfolding in solution and in the gas phase. Our data instead reveal that partially disordered conformers undergo various Coulombically driven structural changes during and/or after the ESI process. We discuss how the observed phenomena can be linked to the ESI mechanisms that are thought to be operative for native and non-native proteins.

## 4.2 Experimental Section

### 4.2.1 Materials and Sample Preparation

Equine skeletal muscle ferri-hMb, horse heart cytochrome *c* (cyt *c*), formic acid, sodium iodide, and ammonium acetate were purchased from Sigma (St. Louis, MO). HPLC-grade glacial acetic acid and LC-MS grade water were obtained from Fisher Scientific (Nepean, Ontario, Canada). aMb was generated by 2-butanone extraction (63). Protein samples were treated as described (Chapter 2), except that ammonium acetate was used instead of phosphate buffer. All other chemicals were used as received. Optical absorption coefficients of  $\epsilon_{409} = 188,000 \text{ M}^{-1} \text{ cm}^{-1}$  for hMb and  $\epsilon_{280} = 13,980 \text{ M}^{-1} \text{ cm}^{-1}$  for aMb were used for concentration measurements (64). pH values were adjusted with formic acid, and measured using a Fisher AB-15 pH meter. Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter (Easton, MD), using solutions identical to those employed for MS. Experimental CD data were converted to mean residue ellipticity  $[\theta]$  (65).

### 4.2.2 Ion Mobility-Mass Spectrometry

All experiments were conducted on a Synapt HDMS instrument (Waters, Milford, MA) equipped with a LockSpray dual ESI source. Protein solutions were infused at a flow rate of  $5 \mu\text{L min}^{-1}$  using a syringe pump. The protein concentration was  $10 \mu\text{M}$ . All samples contained  $10 \text{ mM}$  ammonium acetate adjusted to pH 7.0, 4.0 or 2.0. The capillary voltage was maintained at  $+2.8 \text{ kV}$ , generating multiply protonated intact protein ions. IMS and transfer TWIG wave velocities were set to  $300 \text{ m s}^{-1}$  and  $247 \text{ m s}^{-1}$ , respectively. The IMS

wave height (5 V) was set to maximize the number of bins used, while the transfer wave height was held at 4 V. The source and the trap wave heights were controlled automatically. The backing pressure was adjusted to 6.5 mbar by throttling a SpeediValve on the scroll pump. The trap and transfer TWIGs ( $2.9 \times 10^{-2}$  mbar) were purged with Ar. The IMS cell was filled with 0.3 mbar N<sub>2</sub>.

Ion activation during and after ESI promotes the loss of solution structure (62, 66). The experiments were therefore conducted by employing the most gentle conditions possible, although this greatly compromised the intensity of the measured signals. Heating of the source assembly was switched off (25 °C), and the desolvation gas temperature was kept at its minimum value of 40 °C. DC voltages that control the extent of collisional activation include those on the sampling and extraction cones, the trap collision energy, and the trap bias (61). Tuning of these and all other instrument parameters was performed using *cyt c* at pH 3 as benchmark analyte (67). All voltages were lowered to ensure that the ion signals shifted to the shortest possible drift times, corresponding to the most compact gas phase structures. The optimized IMS parameters were as follows: sampling cone 10 V, extraction cone 1 V, trap collision energy 5 V, transfer collision energy 4 V, trap entrance 0 V, trap bias 11 V, IMS entrance 10 V, IMS exit 0 V, transfer entrance 1 V, transfer exit 1 V. The effects of key DC voltages on the IMS behavior of *cyt c* 10+ ions are illustrated in Figure 4.1. The conversion of experimental IMS drift time distributions to CCSs was performed using the procedure outlined in ref (61), employing a calibrant mixture of Mb, *cyt c* and ubiquitin in 49% water, 49% methanol and 2% acetic acid. Acquisition of calibration data was performed using the same gentle IMS parameters as for the actual experiments (68). The combined

calibration plot obtained for the three reference proteins showed excellent linearity ( $r^2 = 0.9985$ ). All measurements, including the acquisition of calibration data sets, were carried out in triplicate. Error bars represent standard deviations.

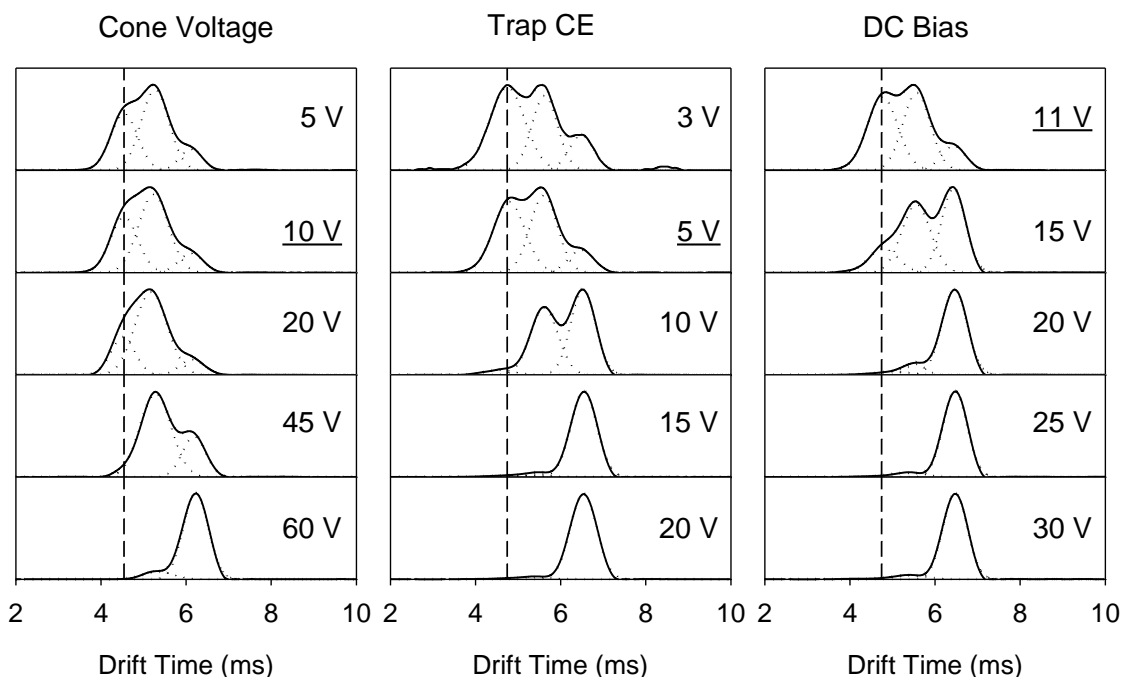


Figure 4.1 IMS drift time data acquired for of the 10+ charge state of *cyt c* at pH 3. This data set illustrates the optimization of cone voltage (left panel), trap collision energy (middle panel), and trap DC bias (right panel). While altering these parameters all other settings were left at the values summarized in the Experimental section of the main text. The optimized cone voltage, trap collision energy, and trap DC bias (underlined in each panel) used for subsequent experiments were chosen based on the value that provides the shortest possible drift time while maintaining acceptable signal intensity. The data were fitted globally using three Gaussians in Microsoft Excel (dotted lines). Dashed lines in each panel indicate the drift time of the most compact conformer.

### 4.2.3 Modeling

CCSs of theoretical model structures were calculated using the exact hard sphere scattering model implemented in MOBCAL (69, 70). PDB file 1WLA (71) was used to represent the native structure of hMb. Hydrogens were added (9) to this crystal structure using PyMol (Schrödinger, New York, NY), and water molecules were removed. As in previous work (41), we considered a linear “string” conformation with all backbone  $\Phi$  and  $\Psi$  angles set to  $180^\circ$  to represent a fully extended protein conformation, with Pro residues substituted for Ala. Random coil calculations were based on 1000 protein conformations generated by picking each of the  $\Phi$  and  $\Psi$  angles at random, subject to an algorithm that suppresses atom-atom collisions. These random conformers were produced using an in-house program. Visualization of protein structures was performed using PyMol.

## 4.3 Results and Discussion

### 4.3.1 Mb Unfolding in Solution Monitored by Optical Spectroscopy

CD spectra confirm the breakdown of protein structure along the  $\text{hMb}_{\text{pH } 7} \rightarrow \text{aMb}_{\text{pH } 7} \rightarrow \text{aMb}_{\text{pH } 4} \rightarrow \text{aMb}_{\text{pH } 2}$  unfolding pathway in solution (Figure 4.2). hMb at pH 7 shows a global minimum at 222 nm, reflecting the highly helical secondary structure of the native state (72). Heme removal at pH 7 only results in relatively minor spectral changes. A significant loss of aMb secondary structure is observed when lowering the pH to 4. At pH 2 the aMb spectrum exhibits a minimum around 200 nm, as expected for a random coil-like conformation (72). The stepwise Mb unfolding process in solution seen under the

conditions of this work is consistent with the results of earlier investigations (56-59) (Chapter 2).

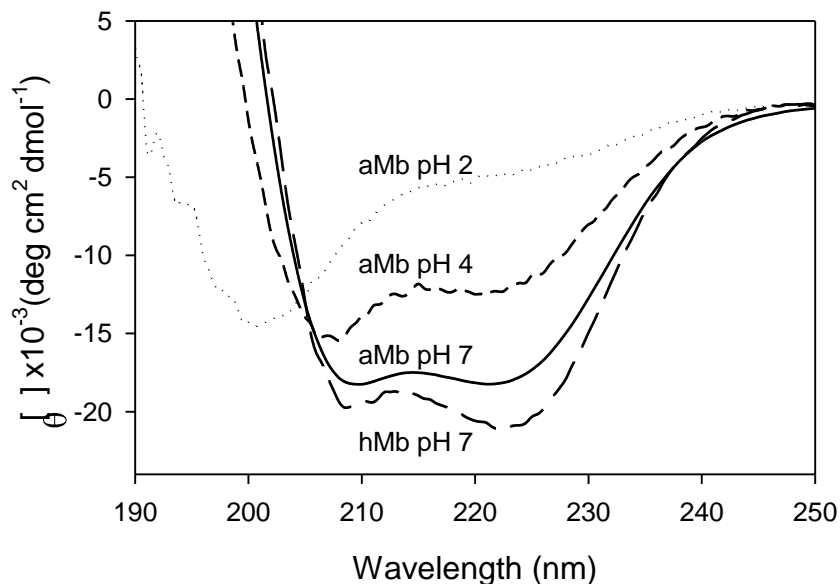


Figure 4.2 Far-UV CD spectra, reflecting the solution phase secondary structural changes along the  $\text{hMb}_{\text{pH } 7} \rightarrow \text{aMb}_{\text{pH } 7} \rightarrow \text{aMb}_{\text{pH } 4} \rightarrow \text{aMb}_{\text{pH } 2}$  unfolding pathway. Spectra were acquired as indicated in the main text.

#### 4.3.2 Mb Unfolding Probed Via ESI Charge State Distributions

ESI-MS has been used previously for monitoring the acid-induced unfolding of Mb (73, 74), but it is helpful to show spectra acquired under the conditions used here to facilitate the discussion of subsequent IM-MS data. Native hMb shows a narrow series of ions that is dominated by a 9+ peak, along with low intensity 8+ and 10+ signals (Figure 4.3A). All of the ions observed in this spectrum retain their heme group, reflecting the fact that collisional activation is minimal for the conditions used here (54). The spectrum of aMb acquired at pH 7 remains dominated by the 9+ charge state, but it exhibits a bimodal

distribution with a group of less intense ions centered at 16+ that indicate the onset of unfolding (Figure 4.3B). The absence of heme is evident from peak shifts relative to Figure 4.3A. For the bimodal distribution observed at pH 4 both groups of signals have roughly the same intensity, and the maximum of the low charge states has shifted to 11+ (Figure 4.3C). At pH 2 aMb exhibits a broad distribution of highly charged ions, centered around 19+ (Figure 4.3D). Overall, the data of Figure 4.3 illustrate the well-known effect that unfolding of a protein in solution dramatically enhances the extent of protonation during ESI, along with a significant broadening of the charge state distribution (52, 53).



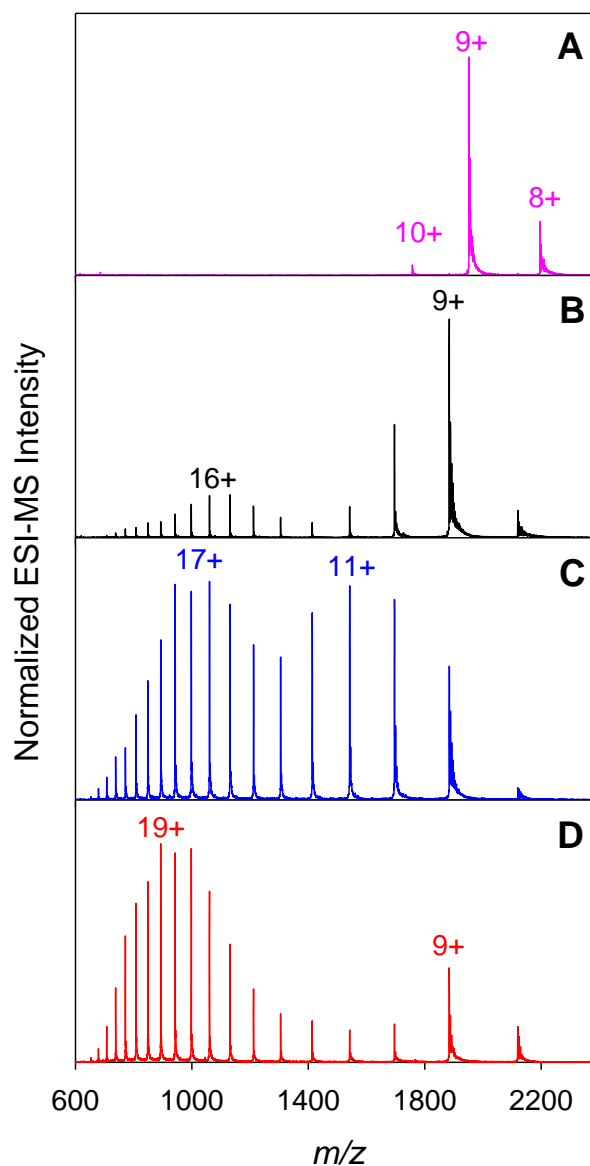


Figure 4.3 ESI mass spectra acquired for (A) hMb at pH 7, pink; (B) aMb at pH 7, black; (C) aMb at pH 4, blue; and (D) aMb at pH 2, red. The same color scheme is used to indicate solution conditions in the subsequent figures. Charge states of selected ions are indicated.

### 4.3.3 Ion Mobility Results

An overview of the CCS distributions measured for hMb and aMb under the different solution conditions is provided in Figure 4.4. The CCSs increase with increasing ESI protonation state, reaching from  $\sim 1900 \text{ \AA}^2$  for 8+ ions all the way to more than  $4000 \text{ \AA}^2$

for  $\text{aMb}^{24+}$ . This range is consistent with earlier observations (41, 61). Charge states 8+ and 9+ show CCS distributions that are unimodal, whereas those in the 10+ to 16+ range are multimodal. The latter can be modeled using two or three Gaussian bands (dotted lines in Figure 4.4). Charge states beyond 16+ have unimodal CCS distributions.

The multimodal nature seen for some of the charge states in Figure 4.5 indicates the presence of co-existing gas phase conformers (7-9). Some of the subsequent considerations can be simplified by using a single parameter to describe the overall IMS behavior of each ionic species. For this purpose average CCSs were calculated according to

$$\langle \Omega \rangle = \frac{\int I(\Omega) \Omega d\Omega}{\int I(\Omega) d\Omega} \quad 4.1$$

where  $I(\Omega)$  is the intensity profile (Figure 4.4) of each given CCS distribution. From Figure 4.5A it is seen that  $\langle \Omega \rangle$  increases monotonically with  $z$  in an almost linear fashion. Differences between the four data sets are therefore seen more clearly after dividing each  $\langle \Omega \rangle$  value by its corresponding charge state  $z$  (Figure 4.5B).

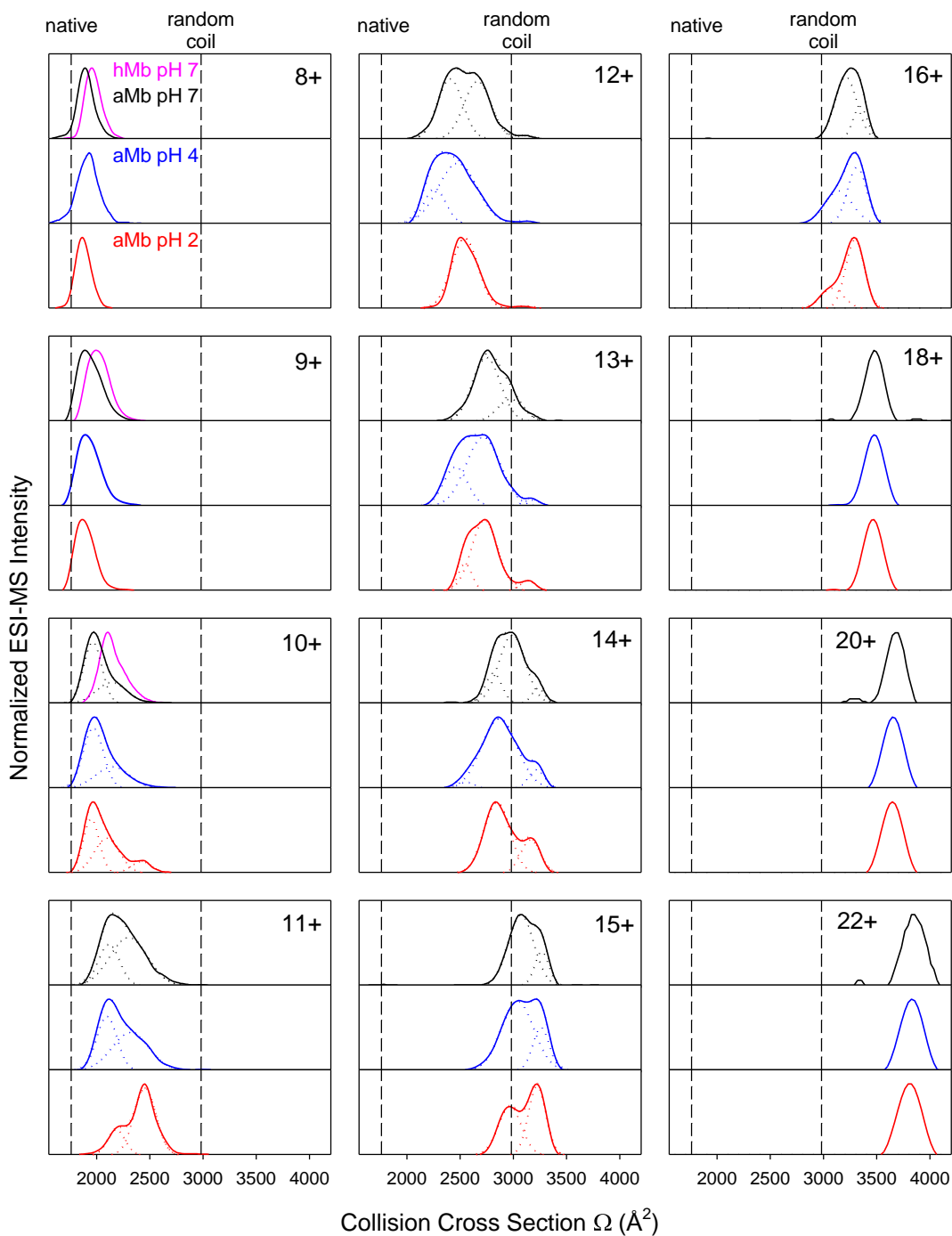


Figure 4.4 CCS distributions measured for ions generated by electrospraying hMb at pH 7 (pink), aMb at pH 7 (black), aMb at pH 4, and aMb at pH 2 (red). The different panels refer to different charge states, as indicated in the Figure. Dotted lines are the result of Gaussian decomposition. Vertical dashed lines indicate the average CCSs calculated for native hMb and for a random coil ensemble as described for Figure 4.6.

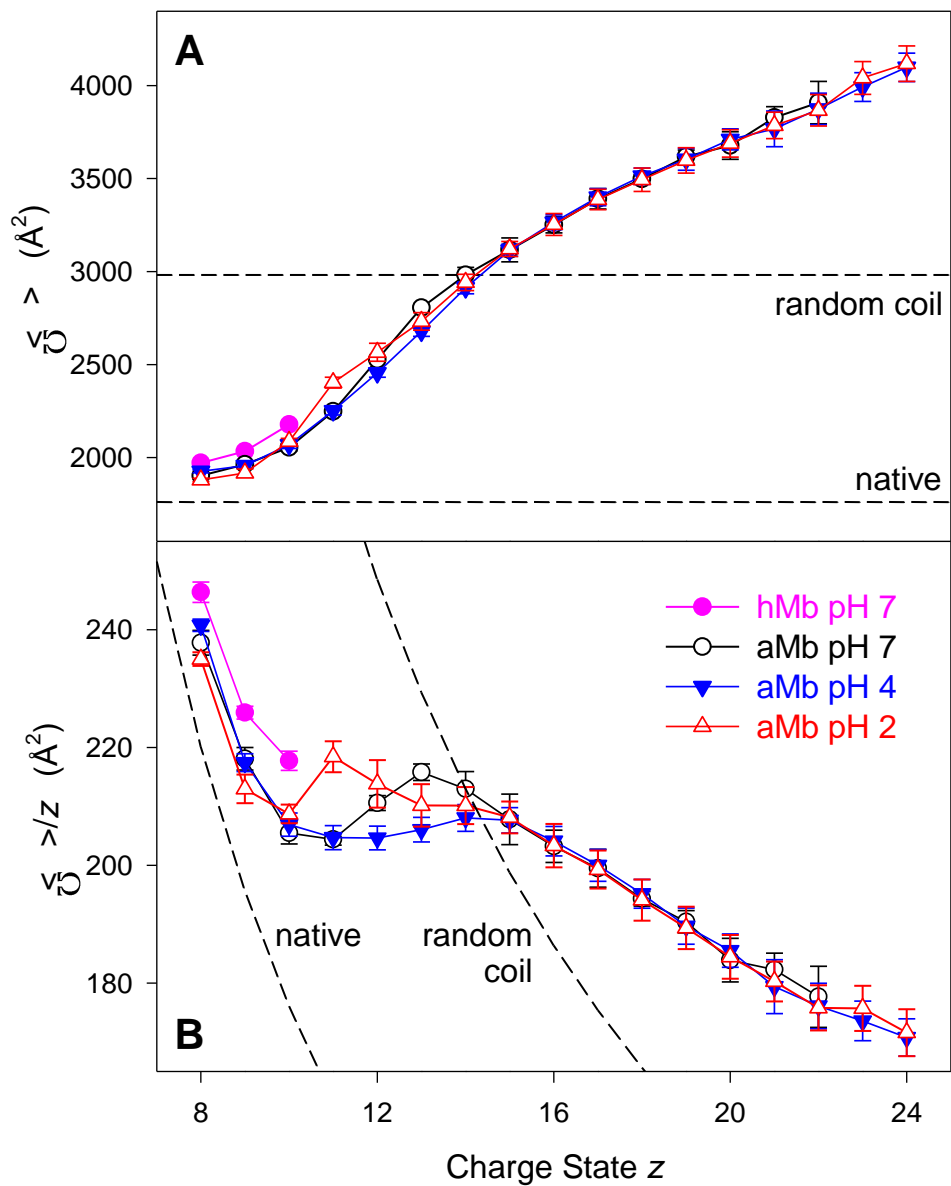


Figure 4.5 (A) Average CCS  $\langle \Omega \rangle$  of ions generated by electrospraying hMb, and aMb at different pH. These data were generated by applying Equation 4.1 to the CCS data of Figure 4.4. (B) The same data set as in panel A, but after division of  $\langle \Omega \rangle$  values by the corresponding charge state  $z$ . Lines labeled “native” and “random coil” correspond to those of Figure 4.4 and Figure 4.6.

#### 4.3.4 Collision Cross Sections of Model Structures

Before discussing the experimental IM-MS data it is instructive to consider MOBCAL-generated (69, 70) CCSs of a few model structures (Figure 4.6). The native state of hMb has a calculated CCS of  $1761 \text{ \AA}^2$ . A maximally extended “string” conformation is characterized by  $\Omega = 5081 \text{ \AA}^2$ , consistent with previously published data (41). Unfortunately, there is no X-ray structure for aMb. When analyzing the native Mb conformation without heme (and without any protein structural changes) a CCS of  $1760 \text{ \AA}^2$  is found, i.e., virtually the same as for intact hMb. As noted earlier, the actual aMb solution conformation at pH 7 is somewhat more expanded than native hMb (56). For approximating the denatured solution structure of aMb at pH 2 we use a random coil model (Figure 4.6) (57). Sampling of 1000 random coil conformers results in a bell-shaped CCS distribution ranging from  $2350 \text{ \AA}^2$  to  $3650 \text{ \AA}^2$ , with  $\langle \Omega \rangle = 2980 \text{ \AA}^2$ .

It is emphasized that the structures displayed in Figure 4.6 are only meant to serve as reference points; no claims are made that they correspond to the experimentally observed gas phase conformers. In addition, recent work related to drag enhancement and other factors implies considerable uncertainties (up to 40%) for existing  $t_d - \Omega$  conversion methods (75-77). At the current stage of development, therefore, it seems prudent to conduct comparisons between measured and calculated CCSs only in a semi-quantitative fashion.

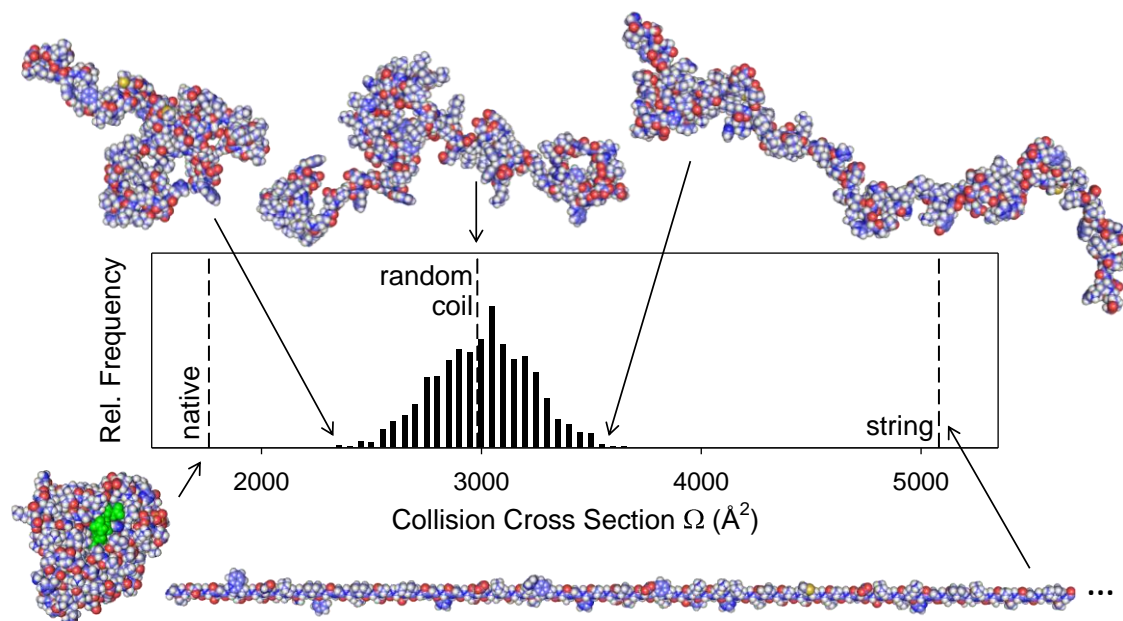


Figure 4.6 CCS values of hypothetical model structures adopted by the Mb polypeptide chain. Protein conformers are shown in spacefill representation; for visualization purposes each of them was rotated to face the observer with the largest possible projection area. Depicted along the top are three conformers that represent different regions of the random coil CCS distribution (shown as histogram). Also included are data for native hMb, with the heme group highlighted in green. In addition, the figure shows the CCS of a fully extended “string” conformation. Due to space limitations only the N-terminal half of the string is depicted (while CCS calculations were performed on the full length sequence).

#### 4.3.5 *Effects of Solution Phase Structure on Gas Phase Conformation*

The key question explored here is to what extent IM-MS analyses of partially disordered proteins can provide insights into solution phase structural properties. The pH 7 and pH 4 states of aMb represent partially disordered solution species, whereas native hMb (pH 7) and unfolded aMb (pH 2) serve as benchmarks (51, 56, 57, 59) (Chapter 2). We will discuss this question on the basis of Figure 4.4 and Figure 4.5, making reference to the MOBCAL results of Figure 4.6.

#### *4.3.5.1 Low Charge States (8+ to 10+)*

hMb ions formed at pH 7 exhibit CCSs that are slightly (~13%) larger than expected for the native conformation, consistent with the behavior seen for some other proteins (24, 78). These relatively compact CCSs, and the fact that the heme-protein complex remains intact (Figure 4.3), suggest that gaseous hMb retains a structure close to the native solution conformation (34, 35). ESI of these tightly folded conformers is believed to proceed via the charged residue mechanism (CRM), where the protein remains surrounded by a water droplet until it is liberated by solvent evaporation to dryness. Evaporative cooling and dielectric stabilization ensure that the CRM is a gentle process that promotes the preservation of native protein structure (16, 17, 79).

The situation is less clear for aMb<sup>8+</sup> to aMb<sup>10+</sup>. The CCSs of these ions show no dependence on the solution environment, and all of them are smaller than the corresponding hMb species (Figure 4.4, Figure 4.5). For aMb at pH 7 the latter effect may arise from collapse of the empty cofactor binding pocket, as suggested for other systems with internal cavities (75, 80). However, it is quite surprising that equally compact gaseous aMb ions are seen for the pH 2 sample. We cannot exclude the possibility that the unfolded pH 2 ensemble in solution extends to very compact species (26, 45), but the random coil modeling of Figure 4.6 argues against this scenario. It is more likely that low charge aMb ions generated at pH 2 reflect a sub-population of unfolded chains that underwent gas phase collapse, similar to other disordered proteins (55). Such a collapse will be favored by the low charge of the ions considered here. While most unfolded chains become highly protonated during ESI (52, 53), a certain

fraction may originate from droplets well below the Rayleigh limit, resulting in a below-average ionic charge (17). We propose that this scenario applies to aMb<sup>8+</sup> to aMb<sup>10+</sup> ions generated at pH 2. Overall, the data discussed here suggest that highly compact gas phase conformers do not necessarily correspond to native-like structures.

#### *4.3.5.2 High Charge States (17+ and beyond)*

The CCS distributions measured for highly charged aMb ions generated from pH 7, 4, and 2 are unimodal, without any discernible dependence on the solution phase conformation. The lack of solution phase memory for these species is attributed to the strong Coulombic repulsion that favors the formation of extended structures in the gas phase (7, 26, 41, 54). The ESI process for these highly charged ions likely proceeds via the chain ejection mechanism (CEM), where an unfolded polypeptide chain gets expelled from the droplet surface (17, 81). Both the droplet and the emerging chain are highly charged. Therefore, the electrostatic “stretching force” experienced by the protein is highest during the ejection process, and it diminishes somewhat after separation from the droplet (17, 82). Clearly, this stretching is detrimental to the survival of solution phase structure, consistent with the fact that highly charged aMb ions generated at pH 7, 4, and 2 share the same CCSs for any given charge state (Figure 4.4 and Figure 4.5). Electrostatic stretching is evident from the fact that, for example, aMb<sup>24+</sup> exhibits a CCS around 4100 Å<sup>2</sup> which is within ~20% of the fully extended string conformation (Figure 4.6).



#### 4.3.5.3 Intermediate Charge States (11+ to 16+)

As discussed, the CRM generates ions in low charge states and favors the retention of solution structure. Conversely, the CEM tends to generate high charge states that disrupt residual solution structure via electrostatic stretching. Intermediate charge states represent a regime in-between these two scenarios. Intriguingly,  $\text{aMb}^{11+}$  to  $\text{aMb}^{16+}$  exhibit multimodal CCS distributions (Figure 4.4). Are the more compact conformers of these distributions remnants of the solution structure? It cannot be excluded that for some of the pH 7 and pH 4 ions this is indeed the case. Importantly, however, each of the intermediate charge states in Figure 4.4 exhibits quite similar compact conformers for *all* solution conditions, *including pH 2 where aMb is extensively unfolded in solution (56-59)* (Chapter 2). Thus, a clear link between low CCS sub-populations and solution phase structural elements cannot be established. Instead, we propose that the multimodal CCS distributions comprise conformers that experienced some structural relaxation after CEM-induced stretching. The occurrence of such relaxation processes is in line with previous studies (14). Electrostatic repulsion suppresses relaxation events for higher charge states, which is why multimodal distributions are not observed for  $\text{aMb}^{17+}$  and beyond.

The preceding considerations notwithstanding, there are some differences between the  $\text{aMb}^{11+}$  to  $\text{aMb}^{16+}$  CCS distributions generated at different pH (Figure 4.4, 4.5). Qualitatively similar effects have been reported for ubiquitin (45, 78). The behavior of  $\text{aMb}^{11+}$  in our data is particularly noteworthy, since ions generated at pH 2 are more expanded than those produced at pH 4 and pH 7. When viewing this charge state in isolation it is tempting to suggest a direct correlation between solution and gas phase

structure (45). Yet, a balanced analysis has to consider the *entire* data set. The trend seen for 11+ is not reproduced by any other charge state. For example, the  $\langle\Omega\rangle$  values for 13+ are in the order pH 7 > pH 2 > pH 4 (Figure 4.5), which is incompatible with the solution phase properties of aMb. We conclude that the outcome of gas phase relaxation processes (see previous paragraph) can be somewhat affected by the solution phase history of the analyte, thus accounting for the pH-dependent effects seen for aMb<sup>11+</sup> to aMb<sup>16+</sup> (Figure 4.5B). The  $\Omega$  values resulting from these processes, however, are difficult to predict. Only some charge states yield a CCS progression that parallels the degree of unfolding in solution.

#### 4.3.6 Intensity-Weighted CCS Distributions

One aspect that has not been considered in our discussion thus far is the fact that the various CCSs of Figure 4.4 and Figure 4.5 are associated with vastly different ion intensities. An alternative view of the ion behavior can thus be obtained by displaying each CCS distribution weighted by its corresponding ESI-MS peak intensity from Figure 4.3. The results of this procedure are displayed in Figure 4.7, along with the CCS envelopes that represent the sum of individual charge state profiles for each solution condition. Also included in Figure 4.7 are the average CCSs  $\langle\Omega\rangle_{\text{tot}}$ , obtained by applying Equation 4.1 to these intensity-weighted envelope data (the subscript “tot” is used to distinguish these values from the single-charge state data of Figure 4.5, see the caption of Figure 4.7 for details of the procedure used).

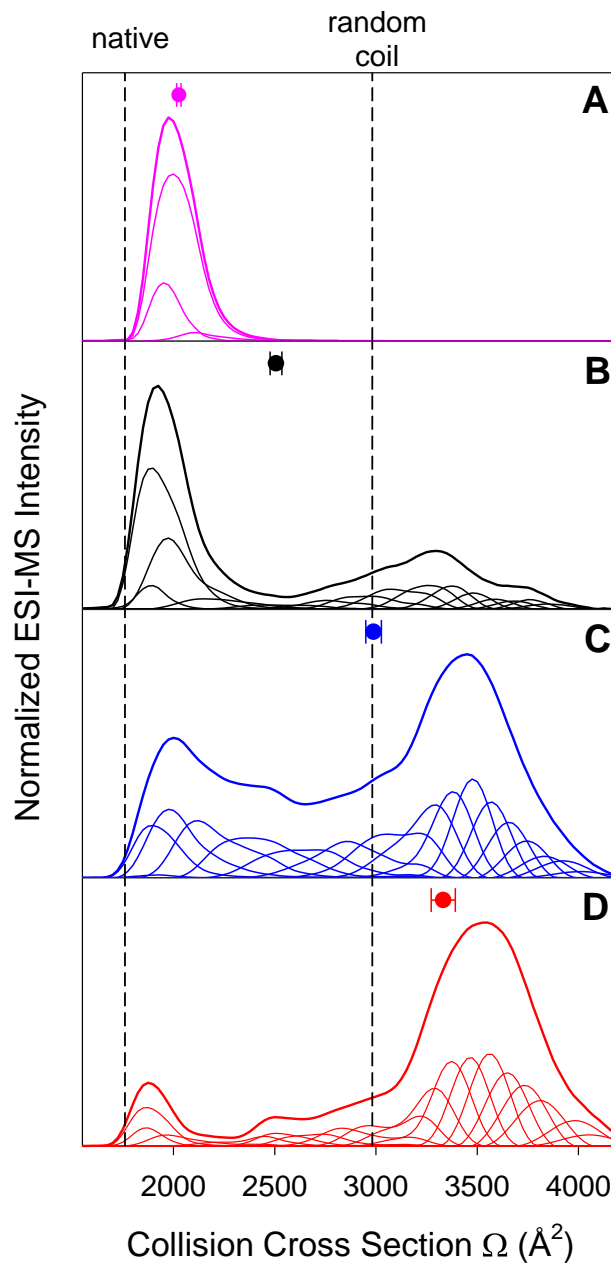


Figure 4.7 Overall CCS profiles (bold lines) of hMb at pH 7 (A), aMb at pH 7 (B), aMb at pH 4 (C), and aMb at pH 2 (D). These profiles were obtained by summation of the charge state-resolved data (thin lines). Each of the component profiles was normalized according to  $INT(z) = \int I(\Omega) d\Omega$ , where  $INT(z)$  is the intensity of the corresponding charge state in the ESI mass spectrum of Figure 4.3. The figure also displays the  $\langle \Omega \rangle_{\text{tot}}$  value for each of the four conditions, calculated as described in the text.

Displaying the IM-MS data in this intensity-weighted fashion results in dramatic differences for the four solution structures (Figure 4.7). The  $\langle\Omega\rangle_{\text{tot}}$  values shift in the expected order, from hMb at pH 7 (most compact) to aMb at pH 2 (most expanded). While interesting, this behavior is not surprising because the CCS profiles of individual charge states all exhibit comparable widths (Figure 4.4), such that the envelopes of Figure 4.7 roughly mirror the ESI mass spectrum of Figure 4.3. The progression of  $\langle\Omega\rangle_{\text{tot}}$  values is thus a manifestation of two well known facts: (1) Unfolded solution conformers generate higher charge states than compact species (52, 53); (2) the compactness of a gas phase protein ion is mainly governed by its charge state via electrostatic repulsion (7, 26, 41, 54).

#### 4.4 Conclusions

The findings of this work support the view that native proteins and biomolecular complexes can retain significant aspects of their solution structure after ESI. IM-MS is a powerful tool for examining the properties of the resulting gaseous species (14, 26, 38, 39). In contrast, our results imply that partially disordered solution conformers undergo significant structural changes, making it difficult to correlate solution and gas phase behavior. Electrosprayed semi-folded proteins in low charge state can undergo collapse (55), while highly charged species tend to undergo electrostatically driven elongation (7, 26, 41, 54). Within the CEM scenario the electrostatic stretching forces experienced by a protein are highest during ejection from the ESI droplet, due to the combined effects of droplet charge and protein charge (17). The electrostatic forces diminish after separation from the droplet, allowing ions in intermediate charge states to relax into more compact

structures that may or may not resemble the solution conformation (14). When judging the relevance of structural trends in the gas phase it is important to ensure that the corresponding phenomena are observed for more than one charge state (45).

We find a close correlation between the degree of unfolding in solution and in the gas phase only when considering CCS data in an intensity-weighted fashion (Figure 4.7). However, this correlation largely reflects the charging behavior of different protein conformers during the ESI process (52, 53), rather than a genuine memory of their solution phase conformations. The structural fate of gaseous protein ions is chiefly governed by their charge (7, 26, 40, 41, 54). Retention of solution structure is favored only for low charge states which, unfortunately, are not the norm when electrospraying semi-unfolded solution conformers. Overall, we feel that the interpretation of IM-MS data for partially disordered proteins requires a cautious approach. We are optimistic that future developments will result in the emergence of more robust strategies for the interrogation of non-native conformers by gas phase methods.

## 4.5 References

1. Fenn JB (2003) Electrospray Wings for Molecular Elephants (Nobel Lecture). *Angew. Chem. Int. Ed.* 42: 3871-3894.
2. Ben-Nissan G & Sharon M (2011) Capturing protein structural kinetics by mass spectrometry. *Chem. Soc. Rev.* 40: 3627-3637.
3. Kaltashov IA & Eyles SJ (2012) *Mass Spectrometry in Structural Biology and Biophysics: Architecture, Dynamics, and Interactions of Biomolecules* (Wiley, Hoboken, NJ) 2nd Ed.

4. Percy AJ, Rey M, Burns KM, & Schriemer DC (2012) Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry-A review. *Anal. Chim. Acta* 721: 7-21.
5. Mendoza VL & Vachet RW (2009) Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* 28: 785-815.
6. Petrotchenko E & Borchers CH (2010) Crosslinking Combined with Mass Spectrometry for Structural Proteomics. *Mass Spectrom. Rev.* 29: 862-876.
7. Bohrer BC, Merenbloom SI, Koeniger SL, Hilderbrand AE, & Clemmer DE (2008) Biomolecule Analysis by Ion Mobility Spectrometry. *Annu. Rev. Anal. Chem.* 1: 293-327.
8. Bush MF, Hall Z, Giles K, Hoyes J, Robinson CV, & Ruotolo BT (2010) Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* 82: 9667-9565.
9. Jurneczko E & Barran PE (2011) 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. *Analyst* 136: 20-28.
10. Clemmer DE & Jarrold MF (1997) Ion Mobility Measurements and their Applications to Clusters and Biomolecules. *J. Mass Spectrom.* 32: 577-592.
11. von Helden G, Gotts NG, & Bowers MT (1993) Experimental evidence for the formation of fullerenes by collisional heating of carbon rings in the gas phase. *Nature* 363(6424): 60-63.
12. Pringle SD, Giles K, Wildgoose JL, Williams JP, Slade SE, Thalassinou K, Bateman RH, Bowers MT, & Scrivens JH (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom.* 261: 1-12.
13. Giles K, Williams JP, & Campuzano I (2011) Enhancements in travelling wave ion mobility resolution. *Rapid Commun. Mass Spectrom.* 25: 1559-1566.

14. Breuker K & McLafferty FW (2008) Stepwise evolution of protein native structure with electrospray into the gas phase,  $10^{-12}$  to  $10^2$  s. *Proc. Natl. Acad. Sci. U.S.A.* 105: 18145-18152.
15. Creighton TE (1993) *Proteins* (W. H. Freeman & Co, New York).
16. Kebarle P & Verkerk UH (2009) Electrospray: From Ions in Solutions to Ions in the Gas Phase, What We Know Now. *Mass Spectrom. Rev.* 28: 898-917.
17. Konermann L, Ahadi E, Rodriguez AD, & Vahidi S (2013) Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* 85: 2-9.
18. Wolynes PG (1995) Biomolecular folding in vacuo!!!(?). *Proc. Natl. Acad. Sci. U.S.A.* 92: 2426-2427.
19. Baumketner A, Bernstein SL, Wyttenbach T, Bitan G, Teplow DB, Bowers MT, & Shea JE (2006) Amyloid  $\beta$ -protein monomer structure: a computational and experimental study. *Protein Sci.* 15: 420-428.
20. Patriksson A, Adams CM, Kjeldsen F, Zubarev RA, & van der Spoel D (2007) A direct comparison of protein structure in the gas and solution phase: The TRP-cage. *J. Phys. Chem. B* 111: 13147-13150.
21. Warnke S, von Helden G, & Pagel K (2013) Protein Structure in the Gas Phase: The Influence of Side-Chain Microsolvation. *J. Am. Chem. Soc.* 135(4): 1177-1180.
22. Ly T & Julian RR (2010) Elucidating the Tertiary Structure of Protein Ions in Vacuo with Site Specific Photoinitiated Radical Reactions. *J. Am. Chem. Soc.* 132: 8602-8609.
23. Skinner OS, McLafferty FW, & Breuker K (2012) How Ubiquitin Unfolds after Transfer into the Gas Phase. *J. Am. Soc. Mass Spectrom.* 23: 1011-1014.
24. Badman ER, Hoaglund-Hyzer CS, & Clemmer DE (2001) Monitoring Structural Changes of Protein in and Ion Trap over  $\sim 10$ -200 ms: Unfolding Transitions in Cytochrome *c* Ions. *Anal. Chem.* 73: 6000-6007.

25. Breuker K & McLafferty FW (2005) The Thermal Unfolding of Native Cytochrome c in the Transition from Solution to Gas Phase Probed by Native Electron Capture Dissociation. *Angew. Chem.* 44: 4911-4914.
26. Wyttenbach T & Bowers MT (2011) Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility–Mass Spectrometry Environment. *J. Phys. Chem. B* 115: 12266-12275.
27. Zhang X & Julian RR (2011) Investigating the gas phase structure of KIX with radical directed dissociation and molecular dynamics: Retention of the native structure. *Int. J. Mass Spectrom.* 308: 225-231.
28. Breuker K, Brüschweiler S, & Tollinger M (2011) Electrostatic Stabilization of a Native Protein Structure in the Gas Phase. *Angew. Chem. Int. Ed.* 50: 873-877.
29. Loo JA (2000) Electrospray Ionization Mass Spectrometry: a Technology for Studying Noncovalent Macromolecular Complexes. *Int. J. Mass Spectrom.* 200: 175-186.
30. Liu L, Bagal D, Kitova EN, Schnier PD, & Klassen JS (2009) Hydrophobic Protein-Ligand Interactions Preserved in the Gas Phase. *J. Am. Chem. Soc.* 131: 15980-15981.
31. Heck AJR (2008) Native mass spectrometry: a bridge between interactomics and structural biology. *Nat. Methods* 5: 927-933.
32. Benesch JLP, Ruotolo BT, Simmons DA, & Robinson CV (2007) Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chem. Rev.* 107: 3544-3567.
33. Daniel JM, Friess SD, Rajagopalan S, Wendt S, & Zenobi R (2002) Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *Int. J. Mass Spectrom.* 216: 1-27.
34. Ruotolo BT, Giles K, Campuzano I, Sandercock AM, Bateman RH, & Robinson CV (2005) Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science* 310: 1658-1661.



35. Scarff CA, Patel VJ, Thalassinos K, & Scrivens JH (2009) Probing Hemoglobin Structure by Means of Traveling-Wave Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 20: 625-631.
36. Bich C, Baer S, Jecklin MC, & Zenobi R (2010) Probing the Hydrophobic Effect of Noncovalent Complexes by Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 21: 286-289.
37. Robinson CV, Chung EW, Kragelund BB, Knudsen J, Aplin RT, Poulsen FM, & Dobson CM (1996) Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein-CoA Ligand Binding and Assembly. *J. Am. Chem. Soc.* 118: 8646-8653.
38. Ruotolo BT & Robinson CV (2006) Aspects of native proteins are retained in vacuum. *Curr. Op. Chem. Biol.* 10: 402-408.
39. Suckau D, Shi, Y., Beu SC, Senko MW, Quinn JP, Wampler FM, & McLafferty FW (1993) Coexisting stable conformations of gaseous protein ions. *Proc. Natl. Acad. Sci. U.S.A.* 90: 790-793.
40. Wood TD, Chorush RA, Wampler FM, Little DP, O'Connor PB, & McLafferty FM (1995) Gas-phase folding and unfolding of cytochrome c cations. *Proc. Natl. Acad. Sci. U.S.A.* 92: 2451-2454.
41. Shelimov KB & Jarrold MF (1997) Conformations, Unfolding, and Refolding of Apomyoglobin in Vacuum: An Activation Barrier for Gas-Phase Protein Folding. *J. Am. Chem. Soc.* 119: 2987-2994.
42. Keppel TR, Howard BA, & Weis DD (2011) Mapping Unstructured Regions and Synergistic Folding in Intrinsically Disordered Proteins with Amide H/D Exchange Mass Spectrometry. *Biochemistry* 50(40): 8722-8732.
43. Woods LA, Radford SE, & Ashcroft AE (2013) Advances in ion mobility spectrometry-mass spectrometry reveal key insights into amyloid assembly. *Biochim. Biophys. Acta* 1834(6): 1257-1268.
44. Beveridge R, Chappuis Q, MacPhee CE, & Barran P (2013) Mass spectrometry methods for intrinsically disordered proteins. *Analyst* 138(1): 32-42.

45. Shi H, Pierson NA, Valentine SJ, & Clemmer DE (2012) Conformation Types of Ubiquitin [M+8H]<sup>8+</sup> Ions from Water:Methanol Solutions: Evidence for the N and A States in Aqueous Solution. *J. Phys. Chem. B* 116: 3344-3352.
46. Hudgins RR, Woenckhaus J, & Jarrold MF (1997) High resolution ion mobility measurements for gas phase proteins: correlation between solution phase and gas phase conformations. *Int. J. Mass Spectrom. Ion Proc.* 165/166: 497-507.
47. Wright PJ, Zhang J, & Douglas DJ (2008) Conformations of Gas-Phase Ions of Ubiquitin, Cytochrome c, Apomyoglobin, and  $\beta$ -Lactoglobulin Produced from Two Different Solution Conformations. *J. Am. Soc. Mass Spectrom.* 19: 1906-1913.
48. Dill KA & MacCallum JL (2012) The Protein-Folding Problem, 50 Years On. *Science* 338: 1042-1046.
49. Hartl FU & Hayer-Hartl M (2009) Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* 16: 574-581.
50. Dyson HJ & Wright PE (2005) Intrinsically Unstructured Proteins and Their Function. *Nat. Rev. Mol. Cell Biol.* 6: 197-208.
51. Bartlett AI & Radford SE (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* 16: 582-588.
52. Chowdhury SK, Katta V, & Chait BT (1990) Probing Conformational Changes in Proteins by Mass Spectrometry. *J. Am. Chem. Soc.* 112: 9012-9013.
53. Kaltashov IA & Abzalimov RR (2008) Do Ionic Charges in ESI MS Provide Useful Information on Macromolecular Structure? *J. Am. Soc. Mass Spectrom.* 19: 1239-1246.
54. Collings BA & Douglas DJ (1996) Conformation of Gas-Phase Myoglobin Ions. *J. Am. Chem. Soc.* 118: 4488-4489.
55. Pagel K, Natan E, Hall Z, Fersht AR, & Robinson CV (2013) Intrinsically Disordered p53 and Its Complexes Populate Compact Conformations in the Gas Phase. *Angew. Chem. Int. Ed.* 52(1): 361-365.

56. Eliezer D, Yao J, Dyson HJ, & Wright PE (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* 5: 148-155.
57. Yao J, Chung J, Eliezer D, Wright PE, & Dyson HJ (2001) NMR Structural and Dynamic Characterization of the Acid-Unfolded State of Apomyoglobin Provides Insights into the Early Events in Protein Folding. *Biochemistry* 40: 3561-3571.
58. Pan J, Han J, Borchers CH, & Konermann L (2009) Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* 131(35): 12801–12808.
59. Hughson FM, Wright PE, & Baldwin RL (1990) Structural Characterisation of a Partly Folded Apomyoglobin Intermediate. *Science* 249: 1544-1548.
60. Evans SV & Brayer GD (1990) High-resolution Study of the Three-dimensional Structure of Horse Heart Metmyoglobin. *J. Mol. Biol.* 213: 885-897.
61. Ruotolo BT, Benesch JLP, Sandercock AM, Hyung S-J, & Robinson CV (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* 3: 1139-1152.
62. Hopper JTS & Oldham NJ (2009) Collision Induced Unfolding of Protein Ions in the Gas Phase Studied by Ion Mobility-Mass Spectrometry: The Effect of Ligand Binding on Conformational Stability. *J. Am. Soc. Mass Spectrom.* 20: 1851-1858.
63. Teale FWJ (1959) Cleavage of the heme-protein link by acid methylethylketone. *Biochim. Biophys. Acta* 35: 543.
64. Antonini E & Brunori M (1971) *Hemoglobin and Myoglobin in Their Reactions With Ligands* (North-Holland Publishing Company, Amsterdam, London).
65. Kelly SW, Jess TJ, & Price NC (2005) How to Study Protein by Circular Dichroism. *Biochim. Biophys. Acta* 1751: 119-139.
66. Shelimov KB & Jarrold MF (1996) "Denaturation" and Refolding of Cytochrome c in Vacuo. *J. Am. Chem. Soc.* 118: 10313-10314.

67. Smith DP, Giles K, Bateman RH, Radford SE, & Ashcroft AE (2007) Monitoring Copopulated Conformational States During Protein Folding Events Using Electrospray Ionization-Ion Mobility Spectrometry-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 18: 2180-2190.
68. Michaelievski I, Kirshenbaum N, & Sharon M (2010) T-wave Ion Mobility-mass Spectrometry: Basic Experimental Procedures for Protein Complex Analysis. *J. Vis. Exp.* 40: e1954.
69. Mesleh MF, Hunter JM, Shvartsburg AA, Schatz GC, & Jarrold MF (1996) Structural information from ion mobility measurements: Effects of the long-range potential. *J. Phys. Chem.* 100(40): 16082-16086.
70. Shvartsburg AA & Jarrold MF (1996) An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem. Phys. Lett.* 261(1-2): 86-91.
71. Maurus R, Overall CM, Bogumil R, Luo Y, Mauk AG, Smith M, & Brayer GD (1997) A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket. *Biochim. Biophys. Acta* 1341: 1-13.
72. Saxena VP & Wetlaufer DB (1971) A New Basis for Interpreting the Circular Dichroic Spectra of Proteins. *Proc. Natl. Acad. Sci. U.S.A.* 68: 969-972.
73. Katta V & Chait BT (1991) Observation of the Heme-Globin Complex in Native Myoglobin by Electrospray-Ionisation Mass Spectrometry. *J. Am. Chem. Soc.* 113: 8534-8535.
74. Dobo A & Kaltashov IA (2001) Detection of Multiple Protein Conformational Ensembles in Solution via Deconvolution of Charge-State Distributions in ESI MS. *Anal. Chem.* 73: 4763-4773.
75. Hogan CJ, Ruotolo BT, Robinson CV, & de la Mora JF (2011) Tandem Differential Mobility Analysis-Mass Spectrometry Reveals Partial Gas-Phase Collapse of the GroEL Complex. *J. Phys. Chem. B* 115(13): 3614-3621.
76. Wyttenbach T, Bleiholder C, & Bowers MT (2013) Factors contributing to the collision cross section of polyatomic ions in the kilodalton to gigadalton range: application to ion mobility measurements. *Anal. Chem.* 85(4): 2191-2199.

77. Chen Y-L, Collings BA, & Douglas DJ (1997) Collision Cross Sections of Myoglobin and Cytochrome c ions with Ne, Ar, and Kr. *J. Am. Soc. Mass Spectrom.* 8: 681-687.
78. Li J, Taraszka JA, Counterman AE, & Clemmer DE (1999) Influence of solvent composition and capillary temperature on the conformations of electrosprayed ions: unfolding of compact ubiquitin conformers from pseudonative and denatured solutions. *Int. J. Mass Spectrom. Ion Proc.* 185/186/187: 37-47.
79. de la Mora FJ (2000) Electrospray Ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta* 406: 93-104.
80. Hall Z, Politis A, Bush MF, Smith LJ, & Robinson CV (2012) Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics. *J. Am. Chem. Soc.* 134: 3429-3438.
81. Chung JK & Consta S (2012) Release Mechanisms of Poly(ethylene glycol) Macroions from Aqueous Charged Nanodroplets. *J. Phys. Chem. B* 116: 5777-5785.
82. Larriba C & Fernandez de la Mora J (2011) The Gas Phase Structure of Coulombically Stretched Polyethylene Glycol Ions. *J. Phys. Chem. B* 116(1): 593-598.

# Chapter 5: H/D Exchange Mass Spectrometry Reveals Friction-Mediated Torsional Stress During F<sub>0</sub>F<sub>1</sub> ATP Synthase Operation

## 5.1 Introduction

HDX/MS is a highly sensitive approach for examining how proteins respond to external stimuli (1, 2). This technique monitors backbone deuteration kinetics in the presence of D<sub>2</sub>O. Disordered segments exchange more rapidly than those in tightly folded regions. HDX under native conditions is mediated by protein conformational dynamics, i.e., rare excursions to locally unfolded conformers (3).

Bacteria, mitochondria, and chloroplasts share a similar ATP synthase architecture, comprising a membrane-embedded F<sub>0</sub> part and a peripheral F<sub>1</sub> portion (Figure 5.1) (4-7). The *Escherichia coli* (*E. coli*) F<sub>0</sub>F<sub>1</sub> system considered here (8) has the rotor composition  $\gamma\epsilon c_{10}$ . The stator comprises the  $\alpha_3\beta_3$  catalytic head, the  $\delta b_2$  peripheral stalk, as well as an *a* subunit.  $\gamma$  forms the central part of the shaft connecting  $\alpha_3\beta_3$  with the *a*<sub>c10</sub> proton translocator (7, 9). F<sub>0</sub>F<sub>1</sub> investigations are typically conducted under hydrolysis conditions (4, 10-12), where  $\gamma\epsilon c_{10}$  rotation is driven by the  $\beta$  subunits that cycle through their  $\beta_{TP}$ ,  $\beta_{DP}$ , and  $\beta_{EMPTY}$  states (13). ATP hydrolysis triggers power strokes that involve consecutive interactions of the  $\beta$  “DELSEED” levers with the asymmetric foot of  $\gamma$  in a stepper motor-like fashion (7, 10, 14).

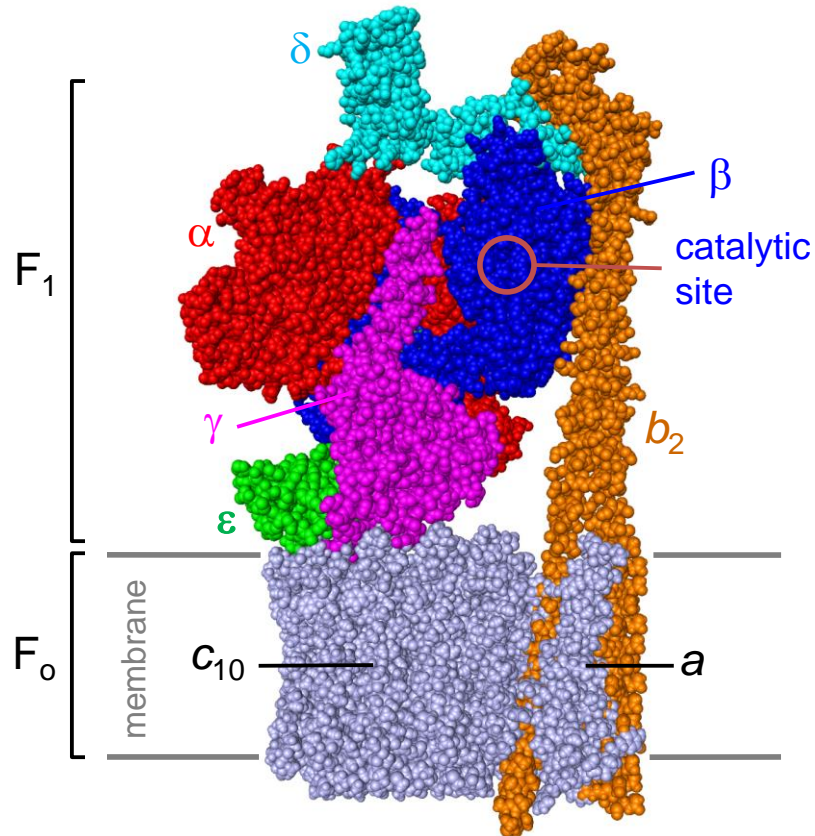


Figure 5.1 Subunit architecture of  $F_0F_1$  ATP synthase from *E. coli*. Composite model assembled from sub-complex X-ray structures 3OAA, 3J0J, 1C17, and 2WSS. One pair of  $\alpha\beta$  subunits facing the observer, as well as ten C-terminal  $\epsilon$  residues, were removed to illustrate how  $\gamma$  extends into the  $\alpha_3\beta_3$  head.

It is difficult to characterize  $F_0F_1$  function in a natural membrane environment. Crystallography has been pivotal for deciphering the structures of isolated sub-complexes (9, 13, 15). X-ray methods and cryo-electron microscopy (16) only provide static snapshots, although computer simulations can help interpret such data in a dynamic context (6, 17). With few exceptions (4, 18), past biophysical attempts to interrogate the rotary mechanism have largely focused on isolated  $F_1$ . Hence, it remains to be established how the  $F_0F_1$  components work together under conditions where ATP/ADP conversion is

coupled with transmembrane proton transport. General relationships between protein function and dynamics have been discussed in depth for regular enzymes (19-21), but not so much for molecular motors that are fuelled by an external energy source. Here we use HDX/MS to interrogate the conformational dynamics of active  $F_0F_1$  in *E. coli* membrane vesicles.

## 5.2 Experimental Section

### 5.2.1 Materials

N-(2-hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine (tricine), adenosine 5'-diphosphate (ADP) sodium salt, adenosine 5'-triphosphate disodium trihydrate salt (ATP), adenosine 5'-( $\beta,\gamma$ -imido)triphosphate lithium salt hydrate (AMP-PNP), phospho(enol)pyruvic acid cyclohexylammonium salt (PEP), (ethylenedinitrilo)tetraacetic acid (EDTA), sodium azide ( $\text{NaN}_3$ ), sodium iodide (NaI), magnesium chloride ( $\text{MgCl}_2$ ), potassium chloride (KCl), carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone (FCCP), valinomycin, rabbit muscle pyruvate kinase, lyophilized porcine pepsin, deuterium oxide ( $\text{D}_2\text{O}$ ) and LeuEnk (YGGFL) were obtained from Sigma (St. Louis, MO, USA). Formic acid (FA), acetonitrile (ACN), and water were of LC-MS grade and came from Fisher Scientific (Nepean, Ontario, Canada).

### 5.2.2 Preparation of inside-out membrane vesicles

*E. coli* strain AN1460, which carries plasmid pAN45 (22) directing moderate constitutive overproduction of wild-type ATP synthase, was grown to stationary phase at 37 °C with vigorous aeration in a salts medium (23) containing 1% glycerol, 0.05% tryptone, and



0.025% yeast extract. After harvesting and washing cells, a 50% suspension was frozen in liquid nitrogen and stored at -80 °C until use. Inside-out membrane vesicles containing ATP synthase were carried using a protocol related to that of Hertzberg and Hinkle (24). All steps carried out on ice or at 4 °C. Thawed cells were stirred for 20 min in 50 mM Tris-HCl, pH 8.0, containing 10 mM MgCl<sub>2</sub> and 1 mM EGTA. After harvesting by centrifugation, they were resuspended in cold MMMKD buffer (10 mM MOPS-KOH, 5 mM MgCl<sub>2</sub>, 10% methanol, 100 mM KCl, 1 mM DTT) supplemented with 0.2 mM CaCl<sub>2</sub>, 100 μM puromycin, and 0.05 mg/ml ribonuclease A. Cells were broken by one passage through a French pressure cell at 8,000 psi. Lysozyme was added to 0.25 mg/ml and cell debris was removed by centrifugation for 10 min at 27,000×g. The supernatant solution was layered onto a steep sucrose gradient over a pad of 2.5 M sucrose, all in the MMMKD buffer containing 100 μM puromycin, and centrifuged for 2.5 h at 100,000×g in a Ti45 rotor. The turbid areas of the sucrose gradient regions were collected and diluted with MMMKD buffer containing 100 μM puromycin, then layered over sucrose gradients in MMMKD buffer containing steps of 1.2 M, 1.6 M, and 1.8 M sucrose. After centrifugation for 18 h in an SW41 Ti rotor, vesicles in regions from 1.2/1.6 M sucrose boundary to the 1.6/1.8 M sucrose boundary were collected, pooled, diluted with MMMK buffer, and sedimented for 1.5 h at 100,000×g. Pellets were resuspended in a small volume of 10 mM MOPS-KOH, 5 mM MgCl<sub>2</sub>, 1 mM DTT, 10 % methanol, 250 mM sucrose, pH 7.5, frozen in liquid nitrogen and stored at -80 °C until use.

### 5.2.3 Functional Enzyme Assays

Membrane-bound ATP synthase activity at 37 °C was determined as described by Cipriano et al. (24). Membrane bound ATPase activity at 22 °C was determined under conditions identical to those of HDX studies, *i.e.* buffer containing 50 mM Tricine-NaOH, pH 8.0, 50 mM KCl, 4 mM MgCl<sub>2</sub>, 0.1 mg/ml pyruvate kinase, 0.5 mM ATP and 60 mM PEP. Membrane protein concentrations in the assay were in the range of 2.6-4.1 mg mL<sup>-1</sup>. Released phosphate was determined using molybdate reagent (25). The measured ATP turnover numbers were  $k_{cat} = (400 \pm 40) \text{ s}^{-1}$  for both the  $W$  and  $W_{FCCP}$  samples.

### 5.2.4 Hydrogen/Deuterium Exchange

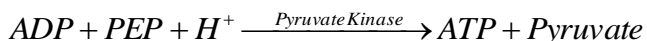
F<sub>0</sub>F<sub>1</sub> ATP synthase was studied under five biochemical conditions. Two conditions under which the enzyme is undergoing catalytic turnover and proton-pumping (i) in the presence of proton motive force (PMF) ( $W$ ); and (ii) in the absence of PMF ( $W_{FCCP}$ ). Two conditions under which the enzyme is either inhibited by (iii) ADP·Mg<sup>2+</sup> and azide ( $I_{ADP}$ ) or (iv) AMP-PNP·Mg<sup>2+</sup> and N<sub>3</sub><sup>-</sup> ( $I_{AMP-PNP}$ ). (v) One condition under which the Mg<sup>2+</sup> content of the sample is depleted  $I_{Mg^{2+} \text{ depleted}}$ . The influence of each of these conditions on structure and dynamics was probed using HDX-MS.

Each experiment was comprised of an equilibration step within which the condition of interest (e.g. inhibition) was imposed on F<sub>0</sub>F<sub>1</sub> ATP synthase via a two-fold dilution into a H<sub>2</sub>O-based buffer. This was followed by a labeling step comprised of a ten-fold dilution into a D<sub>2</sub>O-based solution with identical composition as the equilibrated mixture. The labeling step lasted between 10 seconds to 135 minutes in 90% D<sub>2</sub>O at

pH<sub>corr</sub> 8.0. HDX was quenched by acidifying to pH<sub>read</sub> 2.4 using 10% (v/v) FA. This was immediately followed by addition of 1 nmol of pepsin. Digestion was allowed to proceed for 1 minute on ice. To remove the aggregates resulting from acidification, the samples were centrifuged for 1 minute at 16100×g at 0 °C in a temperature-controlled Eppendorf 5415-R unit (Hamburg, Germany). LC-MS analysis was performed on the supernatant.

#### 5.2.4.1 The Working Conditions

Experiments on  $W$  and  $W_{FCCP}$  were conducted in the presence of an ATP-regenerating system to maintain a high ATP:ADP ratio throughout the experimental time window. This system comprised of pyruvate kinase and PEP and ensures catalytic turnover for at least 45 mins by replenishing the hydrolyzed ATP according to:



Addition of a large amount of ATP is not a viable option because F<sub>0</sub>F<sub>1</sub> ATP synthase is inhibited by ADP·Mg<sup>2+</sup> (26). The ATP-regeneration system avoids this issue by removing ADP from the solution. In addition, this system offers the added benefit of balancing the produced and consumed protons in the ATP hydrolysis and replenishment reactions, thereby avoiding pH fluctuations (see *Enzyme functional assays* for more detail). The solution composition of both working conditions was 50 mM tricine, 50 mM KCl, 60 mM PEP, 1 mg mL<sup>-1</sup> pyruvate kinase 0.5 mM ATP, and 4 mM MgCl<sub>2</sub>. In the  $W_{FCCP}$  condition PMF was abolished by equilibrating and labeling in the presence of 5 μM FCCP and 10 μM valinomycin. The length of the equilibration step was 1 minute for the  $W$  and  $W_{FCCP}$  conditions, thereby ensuring that the enzyme is undergoing catalytic turnover at the onset of labeling.

#### 5.2.4.2 *The Inhibited Conditions*

The solution composition for the  $I_{ADP}$  and  $I_{AMP-PNP}$  conditions was adjusted to mimic that of previous work by Walker and coworkers (26). Both inhibited conditions contained 0.1 mM ADP, 4 mM  $MgCl_2$ , and 3 mM  $NaN_3$ . The  $I_{ADP}$  condition contained 0.5 mM ATP whereas the  $I_{AMP-PNP}$  conditions contained 0.5 mM AMP-PNP. The length of the equilibration step was 60 minutes for the  $I_{ADP}$ ,  $I_{AMP-PNP}$ , and  $I_{Mg^{2+} \text{ depleted}}$  conditions.

#### 5.2.4.3 *The $Mg^{2+}$ -Depleted Condition*

Because  $F_0F_1$  ATP synthase requires  $Mg^{2+}$  for nucleotide binding, this condition represents the case where the  $\beta$  catalytic sites are least occupied relative to other conditions. The  $I_{Mg^{2+} \text{ depleted}}$  sample contained 50 mM tricine, 50 mM KCl, 60 mM PEP, and 1 mM EDTA.

#### 5.2.5 *Liquid Chromatography-Mass Spectrometry*

The peptic peptides in the supernatant were separated using reverse-phase liquid chromatography on a Waters nanoAcquity UPLC system with HDX technology from Waters (Milford, MA, USA) operated at 0 °C. To further minimize back-exchange the system was modified such that the digestion chamber of the instrument is entirely bypassed. The peptides loaded onto a 50  $\mu$ L loop were trapped on a BEH130 C18 (2.1 mm  $\times$  5 mm) VanGuard column at a flow rate of 100  $\mu$ L  $min^{-1}$  for 1 min. Chromatographic separation was achieved over 15 minutes using a BEH130 C8 (2.1 mm  $\times$  50 mm) at 0 °C at a flow rate of 100  $\mu$ L  $min^{-1}$  using a ACN:H<sub>2</sub>O gradient acidified

using 0.1% FA. The average column pressure and chromatographic peak resolution (FWHM) were 9500 psi and ~8.2 sec, respectively. Extensive cleaning and blank injections were performed in-between injections to eliminate carryover.

All mass spectrometry-based experiments were performed on a quadrupole time-of-flight (Q-TOF) second generation Synapt instrument operated in “resolution” ( $R \approx 24000$ ) and ion mobility modes. The instrument was fitted with a standard electrospray (ESI) Z-spray ion source that was operated at a capillary voltage of +3 V. The source block and desolvation gas temperatures were held at 80 and 150 °C, respectively. The cone and desolvation gas flows were 50 and 1000 L hr<sup>-1</sup>, respectively. The sampling and extraction cone voltages were 20 and 4 V, respectively, verified to cause minimal in-source fragmentation. Four traveling wave ion guides (TWIGs) are employed for ion transfer. All TWIGs, except for the IMS TWIG, were controlled automatically. The source TWIG was operated at  $1.0 \times 10^{-3}$  mbar, wave velocity 300 m s<sup>-1</sup>, wave height 0.2 V; the trap TWIG at  $2.4 \times 10^{-2}$  mbar, 508 m s<sup>-1</sup>, 8 V; and the transport TWIG at 508 m s<sup>-1</sup>, 0.1 V at  $2.4 \times 10^{-2}$  mbar. The IMS TWIG was operated at 3.0 mbar, with the wave height being varied between 1000-600 m s<sup>-1</sup> for 100% of the duty cycle at a constant wave height of 40 V. The trap TWIG was purged with 0.4 mL min<sup>-1</sup> of argon, the IMS TWIG with 90 mL min<sup>-1</sup> of N<sub>2</sub> and 180 mL min<sup>-1</sup> of Ar. The trap DC bias was set to 40 V. The quadrupole was set to dwell at m/z 300. The TOF analyzer employed an ADC detector operated at a pressure of  $1.06 \times 10^{-6}$  mbar. The mass spectrometer was calibrated externally using NaI. The lock mass sprayer was sampled every 1 minute and contained 2 μM LeuEnk (1+, m/z 556.2771) in 50% ACN and 0.1% FA and was delivered at 5 μL min<sup>-1</sup> using the instrument’s programmable fluidics compartment.

### 5.2.6 Peptide Mapping

Time-aligned MS<sup>E</sup> methodology was employed for peptide mapping. Over standard MS<sup>E</sup> methods it offers (i) increased peak capacity and reduced spectral complexity (ii) an additional matching parameter and hence improved confidence in assigning product ions to precursors; (iii) the ability to improve peptide fragmentation efficiency by assigning specific collision energies (CE) to each point in the mobility cycle during the high energy scan. Similar to previous work by Distler and coworkers (27), and via analysis of undeuterated test samples and systematically changing the transfer CE, the following transfer CE profile was found to provide the best results. (i) Ion-mobility bins 0–35: transfer CE of 20-25 eV, (ii) ion-mobility bins 36–100: transfer CE ramp from 25 eV to 40 eV, (iii) ion-mobility bins 101–200: transfer CE ramp from 40 eV to 60 eV. Data was acquired over the 50-2000 m/z range and with a scan time of 0.4 sec (for MS<sup>E</sup>) and 1 sec (for non MS<sup>E</sup>) experiments.

Five replicates of MS<sup>E</sup> were recorded for peptide identifications purposes and the data were analyzed using PLGS 3.5.3 from Waters. Raw data were processed with low and high intensity thresholds 125 and 50 counts, respectively. The bin intensity threshold value was set to 750 counts. The data were searched against the entire *E. coli* proteome with a nonspecific protease.

### 5.2.7 Sequence coverage and proteomic profile of membranes

The use of natural *E. coli* membranes in our experiments implies that the peptic digests contained a large number of peptides originating from proteins other than F<sub>0</sub>F<sub>1</sub>. 850 kDa

worth of unique sequence of contaminant proteins associated with bacterial membranes was also identified in all trials. These included proteins comprising the 30S particle of the ribosome, NADH-quinone oxidoreductase, succinate dehydrogenase, aerobic glycerol-3-phosphate dehydrogenase, phosphoenolpyruvate-dependent sugar phosphotransferase system, etc. For proteins corresponding to F<sub>0</sub>F<sub>1</sub> ATP synthase, only those that were identified in all five replicates and with at least 3 consecutive product ions were selected. Under these conditions peptic digestion of F<sub>0</sub>F<sub>1</sub> consistently yielded a total of 203 peptides with S/N ratios that were adequate for providing highly reproducible HDX/MS measurements. Sequence coverage for the extra-membrane subunits of F<sub>0</sub>F<sub>1</sub> was high ( $\alpha$  - 83%,  $\beta$  - 81%,  $\gamma$  - 74%,  $\delta$  - 77%,  $\epsilon$  - 48%, b - 58%). Only a few peptides were detected for membrane-embedded subunits, such that a meaningful characterization of *a* and *c* was not possible. Low digestion yields for transmembrane segments are a common occurrence in HDX/MS. The following considerations will therefore focus on the  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ , and b subunits of F<sub>0</sub>F<sub>1</sub>.

### 5.2.8 HDX Data Analysis

The level of deuterium uptake at time *t* is presented as %D values according to:

$$\%D(t) = \frac{m_t - m_0}{m_{100} - m_0}$$

where  $m_t$  is the absolute deuterium uptake at time *t*,  $m_0$  and  $m_{100}$  correspond to minimally and fully deuterated controls, respectively. That addition of these controls ensures that spurious in-exchange and back-exchange are accounted for. The  $m_{100}$  sample was prepared by following the digestion protocol described above in H<sub>2</sub>O based solution. The

generated peptides were then lyophilized and resuspended in 90% D<sub>2</sub>O at pH<sub>corr</sub> 8 for 1 hour prior to quenching, centrifugation and LC-MS analysis. All %D values are averages of three independent biological and analytical replicates. All HDX data analysis was performed using DynamX 3.0 (Waters).

### 5.3 Results and Discussion

The deuteration behavior of F<sub>0</sub>F<sub>1</sub> was studied under various settings, with particular attention to three conditions (Figure 5.2). (i) Enhanced ADP concentrations produce the inactive state  $I_{ADP}$  where ADP·Mg (without Pi) remains permanently bound in at least one catalytic site (28). (ii)  $W_{FCCP}$  represents working F<sub>0</sub>F<sub>1</sub>, where ATP hydrolysis drives proton transport into the vesicle. PMF buildup is prevented by the uncoupler FCCP (29). (iii)  $W$  represents working F<sub>0</sub>F<sub>1</sub> without uncoupler, such that the enzyme must overcome a PMF-mediated counter-torque.



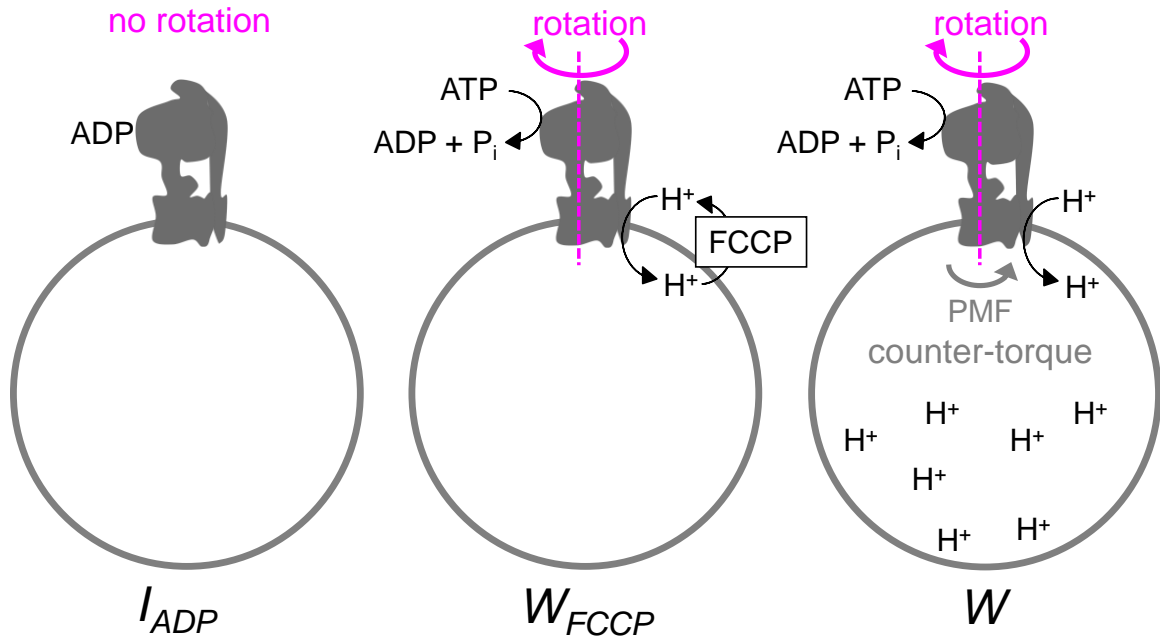


Figure 5.2 Cartoon depiction of membrane vesicle-bound  $F_0F_1$  under different experimental conditions.  $I_{ADP}$  represents ADP-inhibited  $F_0F_1$ ;  $W_{FCCP}$  refers to “working” conditions where  $F_0F_1$  pumps protons in the presence of the uncoupler FCCP which prevents PMF buildup;  $W$  also represents “working”  $F_0F_1$ , but under conditions where protons are pumped against a counter-torque generated by PMF.

An ATP regeneration system was used to ensure that both  $W$  and  $W_{FCCP}$  underwent ATP hydrolysis with  $k_{cat} = (400 \pm 40) \text{ s}^{-1}$  for 45 min. The sequence coverage for peripheral  $F_0F_1$  subunits was  $\sim 80\%$ , while transmembrane coverage was much lower. Thus, we will only consider data for the  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ , and  $b$  subunits. Control experiments confirmed that HDX/MS can pinpoint changes in  $F_0F_1$  structure and dynamics with high fidelity (Figure 5.3).

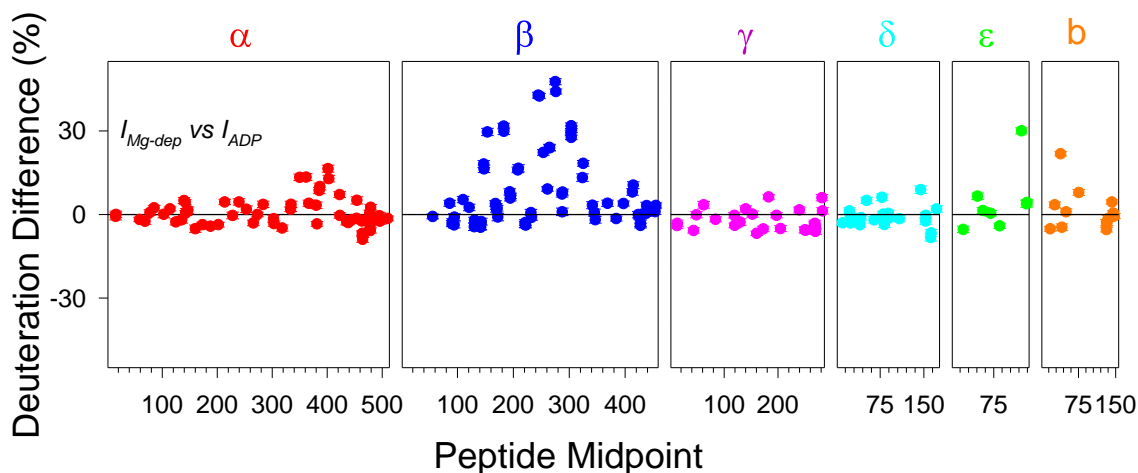


Figure 5.3 HDX behavior of  $F_0F_1$  in an  $Mg^{2+}$ -depleted state  $I_{Mg\_dep}$  relative to the reference state  $I_{ADP}$ . The differences in deuteration percentage are displayed for peptide segments of individual subunits (noted along the top), for a labeling time of  $t = 45$  minutes. Data points represent the average of three independent measurements; error bars represent standard deviations. The  $I_{Mg\_dep}$  exhibits catalytic site nucleotide binding affinities that are orders of magnitude lower than for  $I_{ADP}$  (30). As a result,  $I_{Mg\_dep}$  exhibits greatly enhanced HDX, in the  $\beta$  subunit where the catalytic sites are located (positive deuteration differences) (31).

We will initially discuss some unprocessed mass spectra. The catalytic site P-loop ( $\beta^{148}GGAGVVGKTV^{159}$ ) (13) displays highly asymmetric HDX distributions under  $I_{ADP}$  conditions (Figure 5.4A). This pattern is consistent with three distinct conformations, attributed to the  $\beta_{TP}$ ,  $\beta_{DP}$ , and  $\beta_{EMPTY}$  states (13). Rotational averaging causes coalescence of these distributions into a unimodal envelope ( $W_{FCCP}$  and  $W$ , Figure 5.2). HDX changes similar to those discussed here were observed for two additional peptides close to the catalytic sites,  $\beta^{177-189}$  and  $\alpha^{360-373}$ . Regions that are insensitive to change in conditions include the  $\beta$  lever (Figure 5.4B). In contrast, the  $\gamma$  C-terminal helix shows greatly enhanced HDX under  $W$  conditions (Figure 5.4C).

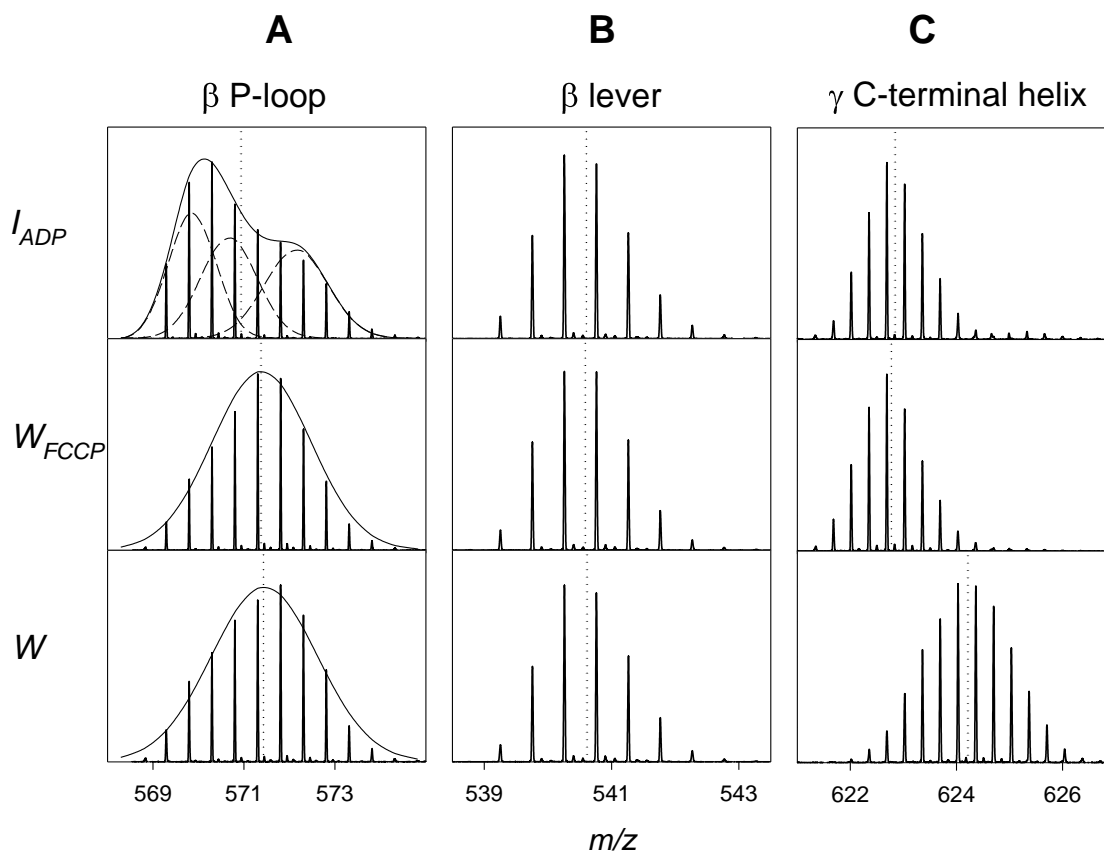


Figure 5.4 HDX behavior of  $W_{FCCP}$  and  $W$  compared to that of  $I_{ADP}$ . (A) Mass spectra of a peptide comprising the active site P-loop at  $t = 1.5$  min ( $\beta^{148}\text{FGGAGVVGKTVNM}^{159}$ ). Gaussian deconvolution of the  $I_{ADP}$  spectra reveals the presence of three distinct conformers, assigned to  $\beta_{TP}$ ,  $\beta_{DP}$ , and  $\beta_{EMPTY}$ . (B)  $\beta$  lever region ( $\beta^{380}\text{DELSEEDKL}^{388}$ ) at  $t = 45$  min. (C)  $\gamma$  C-terminal region ( $\gamma^{260}\text{LQLVYNKARQASITQE}^{275}$ ) at  $t = 45$  min. Vertical dotted lines represent centroid  $m/z$  values.

Differences in HDX patterns (Figure 5.5) can be visualized by adopting  $I_{ADP}$  as reference state, yielding the color maps of Figure 5.6. The overall differences between  $W_{FCCP}$ ,  $W$  and  $I_{ADP}$ , are small throughout the observable sequence range (Figure 5.5 and Figure 5.6A). The one prominent exception is the  $\gamma$  C-terminal helix, which is ~50% more deuterated under  $W$  conditions than in the  $W_{FCCP}$  and  $I_{ADP}$  states (Figure 5.6B). This

dramatic effect implies that under  $W$  conditions the  $\gamma$  C-terminal helix becomes destabilized and spends a considerable fraction of time in locally unfolded conformations.

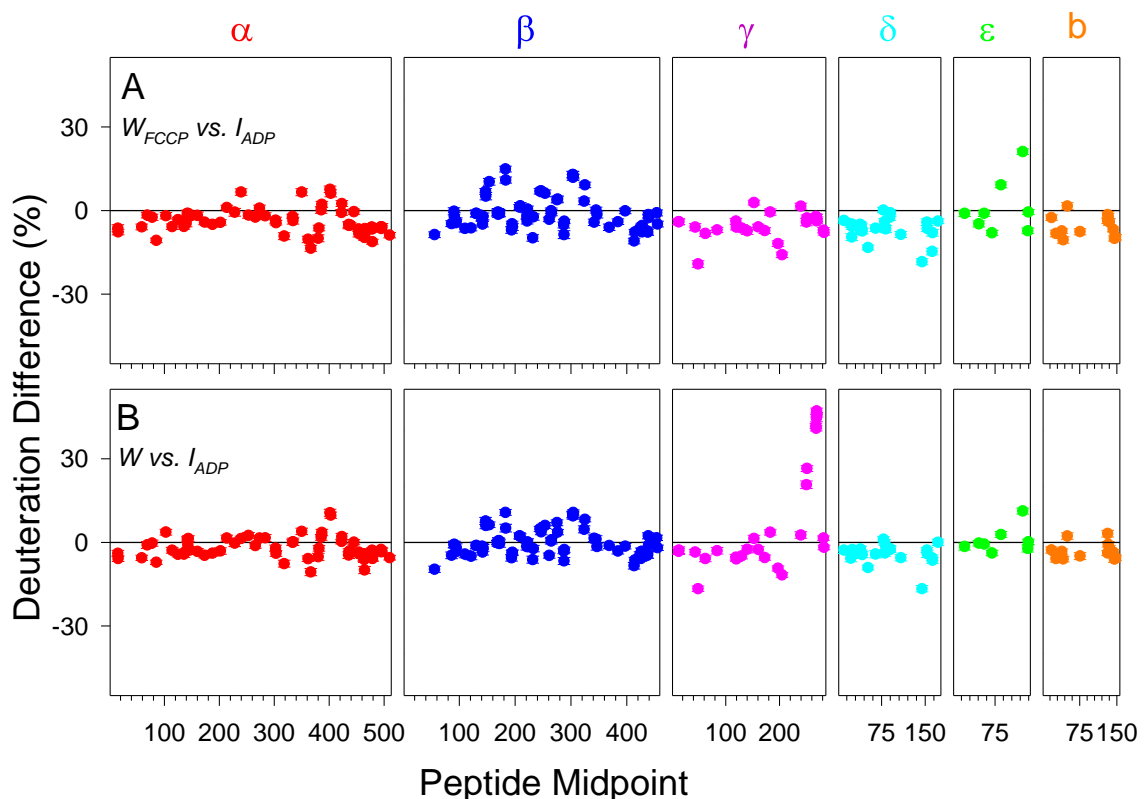


Figure 5.5 HDX behavior of  $F_0F_1$  in under  $W$  and  $W_{FCCP}$  conditions relative to the reference state  $I_{ADP}$ . The differences in deuteration percentage are displayed for peptide segments of individual subunits (noted along the top), for a labeling time of  $t = 45$  minutes. Data points represent the average of three independent measurements; error bars represent standard deviations. (A) Only relatively subtle differences are observed for  $W_{FCCP}$  (working without PMF) relative to  $I_{ADP}$ . (B)  $W$  (proton pumping against PMF) shows dramatically enhanced deuteration in the  $\gamma$  C-terminal helix relative to  $I_{ADP}$ .

$W$  conditions impose enhanced mechanical stress on the  $F_0F_1$  power transmission components, because rotation is opposed by a  $c_{10}$  counter-torque arising from proton

accumulation inside the vesicle (7) (Figure 5.2). This situation is analogous to an automotive powertrain operated under load, e.g., when a combustion engine propels a car uphill. Under  $W_{FCCP}$  conditions the  $c_{10}$  counter-torque is eliminated (Figure 5.2), equivalent to an idling motor that is known to experience much lower mechanical stress (32). This comparison, along with mechanical unfolding studies on other proteins (33), suggests that  $F_0F_1$  elements involved in power transmission should exhibit different HDX characteristics when PMF is present or not. Surprisingly, the main  $F_0F_1$  powertrain elements do not show this effect ( $\beta$  levers and  $\gamma$  foot, Figure 5.4, Figure 5.6). We conclude that large parts of the  $F_0F_1$  H-bonding network are insensitive to intramolecular forces encountered during rotational catalysis.

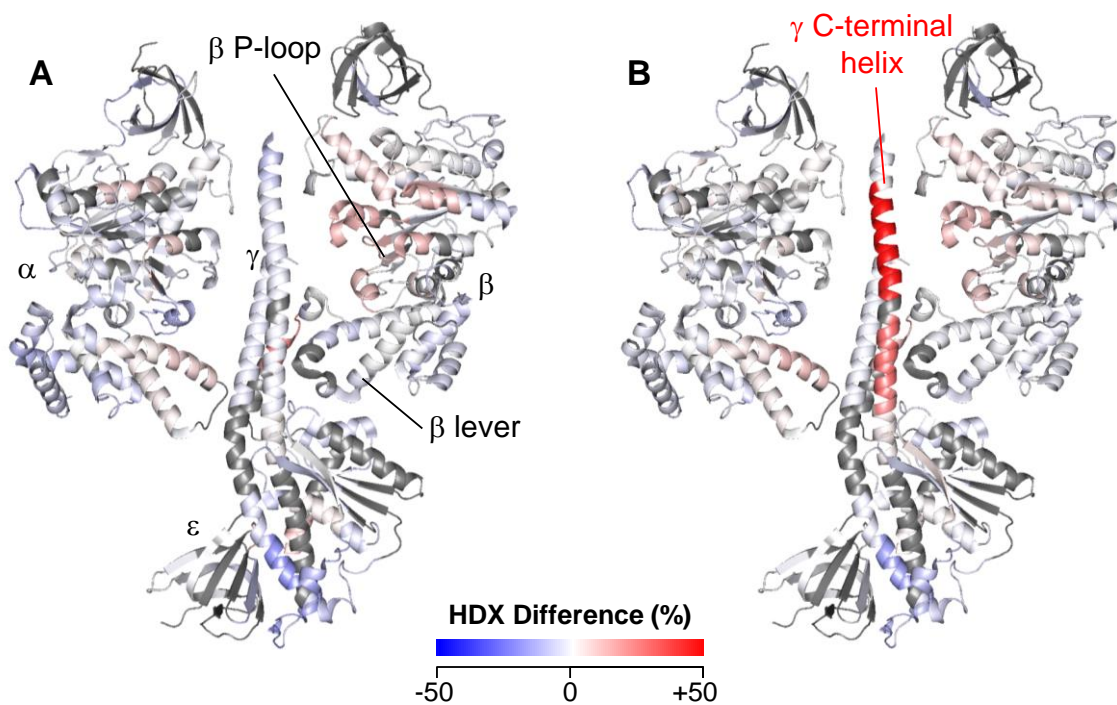


Figure 5.6 (A) Deuterium difference map of  $W_{FCCP}$  vs.  $I_{ADP}$ . (B) Deuterium difference map of  $W$  vs.  $I_{ADP}$ . Red coloring of the  $\gamma$  C-terminal helix highlights dramatically enhanced deuteration under  $W$  conditions compared to  $I_{ADP}$ .

Truncation experiments demonstrated that the  $\gamma$  C-terminal segment is *not* directly involved in power transmission (7). Instead, we attribute the destabilization of this region (Figure 5.4C and Figure 5.6B) to friction effects. An eccentric force acting on a shaft will only result in a stable rotation axis if the shaft is supported by suitable bearings (32). This general principle applies to piston/crankshaft interactions in combustion engines, and to  $\beta/\gamma$  contacts in  $F_0F_1$ . Combustion engine measurements revealed that bearing forces are much greater under load than under idling conditions (32). Bearing forces encountered for high loads cause increased kinetic friction during crankshaft rotation. This can lead to bearing failure, when surface protrusions on either side of the interface scrape past one

another under high pressure such that friction-induced degradation ensues (34). We propose that related friction effects are encountered by the  $\gamma$  rotor, as illustrated in the following by a simple model.

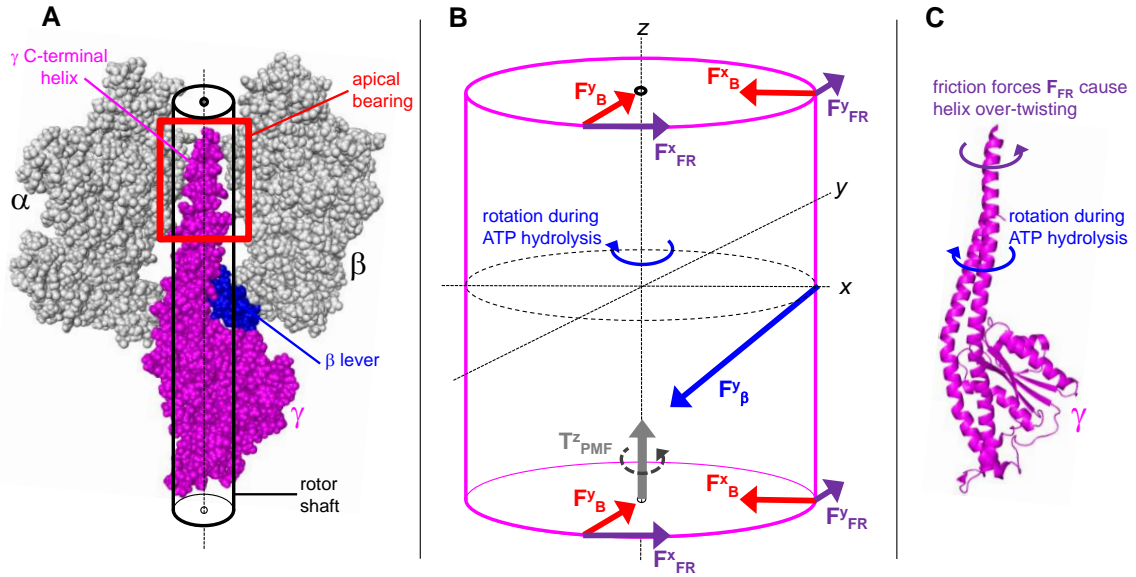


Figure 5.7 Mechanical model, illustrating why PMF destabilizes the  $\gamma$  C-terminus during rotation. (A) Close-up view of  $\gamma/\alpha_3\beta_3$  contacts (pdb file 3OAA after hydrogen addition; only one  $\alpha\beta$  pair is shown). Surface roughness in the apical bearing causes kinetic friction during rotation. (B) Forces ( $F$ ) and torques ( $T$ ) acting on a simplified “ $\gamma$  shaft” under ATP-hydrolysis conditions. All vectors point along principal axes, as indicated by  $xyz$  subscripts. Mechanical equilibrium requires that  $\Sigma F^{xyz} = 0$  and  $\Sigma T^{xyz} = 0$ . PMF generates the torque  $T_{PMF}^c$  (via  $c_{10}$  which is not shown). Shaft rotation against this torque is driven by the  $\beta$  lever force  $F_{\beta}^y$  that is associated with bearing forces  $F_B^y$ . The latter generate friction forces  $F_{FR}^x = k F_B^y$ , ( $k$  = kinetic friction coefficient).  $F_{FR}^x$  has to be compensated by  $F_B^x$ , generating the friction force  $F_{FR}^y = k F_B^x$ . Equations for all forces are derived in the SI. (C) Rotation is opposed by friction forces ( $F_{FR}^x, F_{FR}^y$ ) within the apical bearing, thereby causing over-twisting of the  $\gamma$  C-terminal helix. Destabilization of the  $\gamma$  C-terminal helix takes place only in the presence of PMF because bearing forces and friction forces are proportional to  $T_{PMF}^c$ .

Figure 5.7A highlights the  $\alpha_3\beta_3$  region that serves as apical bearing of the  $\gamma$  shaft. This bearing comprises the sleeve surrounding the very end of  $\gamma$  (13), as well as adjacent contact elements such as the  $\beta$  catch loop (35). It is seen that the bearing interface is not atomically flat. Instead, numerous side chains protrude into the annular gap from either side. Hence, rotation must be accompanied by steric clashes as side chains of  $\gamma$  are forced past those of  $\alpha_3\beta_3$ . Kinetic friction arises whenever moving surfaces come in contact with each other (36). Friction is exacerbated by the fact that, unlike for macroscopic engines (32, 34, 36), the dimensions of the  $\gamma$  bearing leave no space for a film of lubricant molecules (13). Figure 5.7B outlines a cylindrical shaft (representing  $\gamma$ ) that experiences an eccentric force  $F^\gamma_\beta$  (mimicking  $\beta$  lever interactions). The cylinder edges are supported by contact points (bearings) that exert reaction forces which are associated with kinetic friction. These friction forces are proportional to the PMF counter-torque  $T^c_{PMF}$  produced by  $c_{10}$ . This proportionality reflects the fact that the bearings have to press against the  $\gamma$  shaft much more strongly when  $F_oF_1$  operates under load, i.e., in the presence of PMF. The friction forces oppose rotation, thereby over-twisting the  $\gamma$  C-terminal helix (Figure 5.7C). We propose that this over-twisting destabilizes backbone H-bonds, thereby causing enhanced HDX rates as seen in Figure 5.4C and Figure 5.6B. The dependence of friction forces on  $T^c_{PMF}$  implies that destabilization of  $\gamma$  will not occur in the absence of PMF. This model prediction explains why the  $\gamma$  C-terminus undergoes enhanced HDX only in the  $W$  state, but not under  $W_{FCCP}$  conditions.

One of the simplifying assumptions in our model is that the bearings supporting the rotor shaft on either end are identical. In reality the membrane-embedded  $c_{10}$  terminus is very different from the apical region. It would be interesting to examine if the  $c_{10}$



bearing experiences HDX effects similar to those seen for the  $\gamma$  C-terminus. Unfortunately, this aspect remains inaccessible due to the limited transmembrane sequence coverage.

Living cells always have PMF. Thus, our results suggest that kinetic friction is intrinsic to  $F_0F_1$  operation *in vivo*. This contrasts earlier proposals that envision smooth  $\gamma$  rotation facilitated by nonpolar side chains lining the apical sleeve (7, 13). The presence of a “squeaky bearing” does not necessarily imply functional disadvantages. Rotational catalysis in isolated  $F_1$  proceeds unimpeded even after cross-linking the  $\gamma$  C-terminus with  $\alpha$  (37). In these cross-linked constructs the torque provided by  $\alpha_3\beta_3$  unfolds the  $\gamma$  C-terminus, causing  $\gamma$  to undergo swivel motions instead of rotating as a whole (37). We do not propose that swiveling takes place in (non-crosslinked)  $F_0F_1$  under normal operating conditions. Nonetheless, our data show that rotation of  $\gamma$  in the apical bearing is not as smooth as envisioned previously (13). This basic conclusion should hold for both directions, although our experiments only focused on ATP hydrolysis-driven rotation. The apical bearing, as defined in Figure 5.7, comprises the  $\beta$  catch loop which may help coordinate rotation with ATP/ADP binding and release (35). Catch loop- $\gamma$  interactions undoubtedly contribute to kinetic friction, such that some of the  $\gamma$  destabilization seen here may reflect an essential component of  $F_0F_1$  function. In addition, twisting of  $\gamma$  C-terminus may contribute to elastic energy transmission, as previously suggested for other regions of the  $\gamma\epsilon c_{10}$  rotor (4, 7).

## 5.4 References

1. Pirrone GF, Iacob RE, & Engen JR (2015) Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Anal. Chem.* 87(1): 99-118.
2. Konermann L, Pan J, & Liu Y (2011) Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem. Soc. Rev.* 40: 1224-1234.
3. Englander SW, Mayne L, & Krishna MMG (2007) Protein folding and misfolding: mechanism and principles. *Quart. Rev. Biophys.* 40: 287-326.
4. Sielaff H & Borsch M (2013) Twisting and subunit rotation in single FOF1-ATP synthase. *Philos. Trans. R. Soc. B-Biol. Sci.* 368(1611).
5. Wachter A, Bi YM, Dunn SD, Cain BD, Sielaff H, Wintermann F, Engelbrecht S, & Junge W (2011) Two rotary motors in F-ATP synthase are elastically coupled by a flexible rotor and a stiff stator stalk. *Proc. Natl. Acad. Sci. U. S. A.* 108(10): 3924-3929.
6. Czub J & Grubmuller H (2011) Torsional elasticity and energetics of F-1-ATPase. *Proc. Natl. Acad. Sci. U. S. A.* 108(18): 7408-7413.
7. Junge W, Sielaff H, & Engelbrecht S (2009) Torque generation and elastic power transmission in the rotary FOF1-ATPase. *Nature* 459(7245): 364-370.
8. Ballhausen B, Altendorf K, & Deckers-Hebestreit G (2009) Constant c(10) Ring Stoichiometry in the Escherichia coli ATP Synthase Analyzed by Cross-Linking. *J. Bacteriol.* 191(7): 2400-2404.
9. Symersky J, Pagadala V, Osowski D, Krah A, Meier T, Faraldo-Gomez JD, & Mueller DM (2012) Structure of the c(10) ring of the yeast mitochondrial ATP synthase in the open conformation. *Nat. Struct. Mol. Biol.* 19(5): 485-U440.
10. Martin JL, Ishmukhametov R, Hornung T, Ahmad Z, & Frasch WD (2014) Anatomy of F1-ATPase powered rotation. *Proc. Natl. Acad. Sci. U.S.A.* 111(10): 3715-3720.

11. Uchihashi T, Iino R, Ando T, & Noji H (2011) High-Speed Atomic Force Microscopy Reveals Rotary Catalysis of Rotorless F<sub>1</sub>-ATPase. *Science* 333(6043): 755-758.
12. Noji H, Yasuda R, Yoshida M, & Kinosita K (1997) Direct observation of the rotation of F<sub>1</sub>-ATPase. *Nature* 386(6622): 299-302.
13. Abrahams JP, Leslie AGW, Lutter R, & Walker JE (1994) Structure at 2.8 Å resolution of F<sub>1</sub>-ATPase from bovine heart mitochondria. *Nature* 370: 621-628.
14. Mnatsakanyan N, Krishnakumar AM, Suzuki T, & Weber J (2009) The Role of the beta DELSEED-loop of ATP Synthase. *J. Biol. Chem.* 284(17): 11336-11345.
15. Cingolani G & Duncan TM (2011) Structure of the ATP synthase catalytic complex (F<sub>1</sub>) from *Escherichia coli* in an autoinhibited conformation. *Nat. Struct. Mol. Biol.* 18: 701-708.
16. Baker LA, Watt IN, Runswick MJ, Walker JE, & Rubinstein JL (2012) Arrangement of subunits in intact mammalian mitochondrial ATP synthase determined by cryo-EM. *Proc. Natl. Acad. Sci. U. S. A.* 109(29): 11675-11680.
17. Pu J & Karplus M (2008) How subunit coupling produces the  $\gamma$ -subunit rotary motion in F<sub>1</sub>-ATPase. *Proc. Natl. Acad. Sci. U.S.A.* 105: 1192-1197.
18. Sambongi Y, Iko Y, Tanabe M, Omote H, Iwamoto-Kihara A, Ueda I, Yanagida T, Wada Y, & Futai M (1999) Mechanical rotation of the c subunit oligomer in ATP synthase (FOF<sub>1</sub>): Direct observation. *Science* 286(5445): 1722-1724.
19. Eisenmesser EZ, Miller O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Sosco DA, Skalicky JJ, Kay LE, & Kern D (2005) Intrinsic dynamics of an enzyme underlies catalysis. *Nature* 438: 117-121.
20. Liuni P, Jeganathan A, & Wilson DJ (2012) Conformer Selection and Intensified Dynamics During Catalytic Turnover in Chymotrypsin. *Angew. Chem. Int. Ed.* 51: 9666 –9669.
21. Bhabha G, Lee J, Ekiert DC, Gam J, Wilson IA, Dyson HJ, Benkovic SJ, & Wright PE (2011) A Dynamic Knockout Reveals That Conformational

- Fluctuations Influence the Chemical Step of Enzyme Catalysis. *Science* 332: 234-238.
22. Downie JA, Langman L, Cox GB, Yanofsky C, & Gibson F (1980) Subunits of the adenosine-triphosphatase complex translated in vitro from the Escherichia-Coli unc operon. *J. Bacteriol.* 143(1): 8-17.
  23. Tanaka S, Lerner SA, & Lin ECC (1967) Replacement of a phosphoenolpyruvate-dependent phosphotransferase by a nicotinamide adenine dinucleotide-linked dehydrogenase for utilization of mannitol. *J. Bacteriol.* 93(2): 642-648.
  24. Israelachvili JN (1974) Van der Waals forces in biological systems. *Q. Rev. Biophys.* 6: 341-387.
  25. Taussky HH & Shorr E (1953) A microcolorimetric method for the determination of inorganic phosphorus. *J. Biol. Chem.* 202(2): 675-685.
  26. Bowler MW, Montgomery MG, Leslie AGW, & Walker JE (2006) How azide inhibits ATP hydrolysis by the F-ATPases. *Proc. Natl. Acad. Sci. U. S. A.* 103(23): 8646-8649.
  27. Distler U, Kuharev J, Navarro P, Levin Y, Schild H, & Tenzer S (2014) Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* 11(2): 167-170.
  28. Saita E, Iino R, Suzuki T, Feniouk BA, Kinoshita K, & Yoshida M (2010) Activation and Stiffness of the Inhibited States of F-1-ATPase Probed by Single-molecule Manipulation. *J. Biol. Chem.* 285(15): 11411-11417.
  29. Cipriano DJ & Dunn SD (2006) The Role of the  $\epsilon$  Subunit in the *Escherichia coli* ATP Synthase: The C-terminal Domain is Required for Efficient Energy Coupling. *J. Biol. Chem.* 281: 501-507.
  30. Weber J, Wilkemounts S, & Senior AE (1994) Cooperativity and stoichiometry of substrate-binding to the catalytic sites of escherichia-coli F1-ATPase: Effects of magnesium, inhibitors, and mutation. *J. Biol. Chem.* 269(32): 20462-20467.
  31. Powell KD, Ghaemmaghani S, Wang MZ, Ma L, Oas TG, & Fitzgerald MC (2002) A General Mass Spectrometry-Based Assay for the Quantitation of

Protein-Ligand Binding Interactions in Solution. *J. Am. Chem. Soc.* 124: 10256-10257.

32. Kapulainen M, Ronkainen H, Hokkanen A, Stuns I, Varjus S, Nyysönen S, Turunen R, & Halme J (2014) Fibre Optic Sensors for Long-Term Monitoring of Oil Film Pressure in Diesel Engine Main Bearing. *Tribol. Lett.* 56(1): 47-54.
33. Cecconi C, Shank EA, Bustamante C, & Marqusee S (2005) Direct Observation of the Three-State Folding of a Single Protein Molecule. *Science* 309: 2057-2060.
34. Wang QA (1997) Seizure failure of journal-bearing conformal contacts. *Wear* 210(1-2): 8-16.
35. Greene MD & Frasch WD (2003) Interactions among gamma R268, gamma Q269, and the beta subunit catch loop of Escherichia coli F-1-ATPase are important for catalytic activity. *J. Biol. Chem.* 278(51): 51594-51598.
36. Mosey NJ, Muser MH, & Woo TK (2005) Molecular mechanisms for the functionality of lubricant additives. *Science* 307(5715): 1612-1615.
37. Hilbers F, Junge W, & Sielaff H (2013) The Torque of Rotary F-ATPase Can Unfold Subunit Gamma If Rotor and Stator Are Cross-Linked. *PLoS One* 8(1).

## Chapter 6: Conclusions

### 6.1 Summary

Understanding biological processes requires characterization of the molecular players that participate in them. Proteins often take up a central role in these. Structural biologists have managed to amass a vast repertoire of tools to characterize protein structure, folding and dynamics (Section 1.2). Despite these incredible advances, some aspects remain intractable. Structural MS methods represent an increasingly popular set of tools that hold great promise for filling in the gaps where other methods fall short (1). This dissertation aims to add to this rapidly growing toolbox by developing complementary MS-based methods and applying them to solve problems of fundamental and biological interest.

The aim of Chapter 1 was to assess the applicability of fast photochemical oxidation of proteins (FPOP) to characterization of partially disordered conformers. This is of particular interest because these species are not amenable to classical high-resolution techniques discussed in Sections 1.2.1 and 1.2.3. Such equilibrium intermediates can often be populated by exposure to mildly acidic pH. In FPOP,  $\cdot\text{OH}$  generated by laser photolysis of  $\text{H}_2\text{O}_2$  introduces oxidative modifications at solvent accessible side chains (2). By contrast, buried sites are protected from radical attack. The resulting labeling pattern can be analyzed by MS. In  $\cdot\text{OH}$  labeling, it can be challenging to separate structurally induced labeling changes from pH-mediated "secondary" effects (3). Chapter 2 demonstrated that with a proper choice of conditions (e.g., in the absence of pH-dependent  $\cdot\text{OH}$  scavengers) such undesired phenomena can be almost completely eliminated. Using apomyoglobin (aMb) as a model system, the structure of an

intermediate that is formed at pH 4 was mapped. This species retains a highly protected helix G that is surrounded by partially protected helices A, B, and H. Chapter 2 demonstrated the utility of FPOP for the structural characterization of equilibrium intermediates. The near-absence of an intrinsic pH dependence represents an advantage compared to hydrogen/deuterium exchange (HDX) MS (4).

Chapter 3 developed a method to enable studying early events in protein folding. A number of ultrarapid triggering techniques have been available for some time (5, 6), but coupling of these techniques with detection methods that are capable of providing detailed structural information has proven to be difficult. Chapter 3 addresses this issue by combining submillisecond mixing with FPOP. aMb served as a model system for these measurements. The submillisecond mixer used improved the time resolution by a factor of 50 compared to earlier ·OH labeling experiments (7). Spatially-resolved changes in solvent accessibility follow the folding process. Data obtained in this way indicate that early aMb folding events are driven by both local and sequence-remote docking of hydrophobic side chains. Assembly of a partially formed A(E)G(H) scaffold after 0.2 ms is followed by stepwise consolidation that ultimately yields the native state. It was shown that by employing submillisecond mixing in conjunction with slower mixing techniques; it is possible to observe complete folding pathways, from fractions of a millisecond all the way to minutes.

Chapter 4 explored the structural relationship between solution and gas phase protein conformers. Collision cross sections (CCSs) measured by ion mobility mass spectrometry (IM-MS) provide a measure of analyte size (8). It is well established that for native proteins and their complexes many structural features can be preserved in the

gas phase (9). It is unclear in how far IM-MS is suitable for exploring structural properties of semi-folded species (10). Chapter 4 addresses this question, using myoglobin as model system. This protein follows a sequential unfolding pathway that comprises two partially disordered states, i.e. aMb at pH 7 and pH 4 (11). IM-MS data acquired for these two conformers were compared to those of native hMb at pH 7, and extensively unfolded aMb at pH 2. When examining individual aMb charge states, it was found that the degree of gas phase unfolding is not strongly correlated with the corresponding solution behavior. It was found that that the high ESI charge states acquired by non-native conformers generate result in conformational transitions driven by electrostatics. Overall, Chapter 4 revealed that gas phase unfolding as well as collapse events can lead to disparities between gaseous and solution structures for partially unfolded proteins. IM-MS data on non-native conformers should therefore be interpreted with caution.

Chapter 5 examined the role of conformational dynamics for the function of multi-protein molecular machines such as  $F_0F_1$  ATP synthase using HDX-MS. HDX-MS is a highly sensitive approach for monitoring the structural dynamics of proteins (12, 13). HDX-MS monitors backbone deuteration kinetics in the presence of  $D_2O$ . Disordered segments exchange more rapidly than those in tightly folded regions. Measurements of spatially-resolved deuterium are performed using LC-MS.  $F_0F_1$  is a membrane-bound multi-protein complex that uses proton-motive force (PMF) to drive the synthesis of ATP from ADP and inorganic phosphate (14). Reverse operation of the enzyme generates PMF via ATP hydrolysis. Operation in either direction involves rotation of the  $\gamma$  subunit which is part of a central shaft connecting the  $\alpha_3\beta_3$  head and the membrane-anchored  $c_n$



ring. X-ray crystallography (15, 16) and other techniques (17-19) have provided detailed insights into the structure and function of  $F_0F_1$  sub-complexes. However, interrogating the overall conformational dynamics of intact membrane-bound  $F_0F_1$  during rotational catalysis has proven to be difficult. Chapter 5 shows that HDX-MS (12, 13) provides a unique window into the inner workings of  $F_0F_1$  in its natural membrane-bound state. The backbone H-bonding network of key power transmission elements is insensitive to PMF-induced mechanical stress. Unexpectedly, HDX-MS reveals a pronounced destabilization of the  $\gamma$  C-terminus during rotational catalysis. This partial unfolding of  $\gamma$  occurs only when  $F_0F_1$  operates against a PMF-induced torque; the effect disappears when PMF is eliminated by an uncoupler. The behavior of  $\gamma$  is attributed to kinetic friction within the apical rotor bearing. Contrary to earlier proposals, the data presented in this Chapter demonstrate that friction-mediated torsional stress is an intrinsic factor associated with  $F_0F_1$  operation under physiologically relevant conditions. PMF effects on  $\gamma$  are analogous to those encountered in combustion engines, where the crankshaft bearings exert much greater forces when operated under load than when the engine is idling (20).

## 6.2 Future Directions

### 6.2.1 Characterization of the Labeling Pulse in FPOP

One of the most important figures of merit of covalent labeling methods is the length of the labeling step  $\tau$ . Using methods with short  $\tau$  is desirable because labeling-induced unfolding of proteins can be circumvented. In FPOP,  $\cdot\text{OH}$  are generated from photolysis of  $\text{H}_2\text{O}_2$  by a pulsed excimer laser (2).  $\tau$  is equivalent to the lifetime of the generated  $\cdot\text{OH}$  and the multistep chemistry that ensues (21). Hambly and Gross employed kinetic

analysis to estimate  $\tau$  because direct experimental measurements are extremely challenging. Calculations using published rate constants showed that significant [ $\cdot\text{OH}$ ] persist in solution after hundreds of  $\mu\text{s}$  from the initial laser pulse. Using the same analysis method, it was shown that the lifetime of these radicals can be limited to *ca.* 1  $\mu\text{s}$  by the addition of 15 mM Gln as a radical scavenger (2). FPOP owes much of its popularity to this short labeling time. This is shorter than most protein conformational changes, such that a “snapshot” of the protein's state is taken before it can structurally respond to the labeling (22). Also, this feature paves the way for monitoring fast kinetics and characterizing fleeting events as demonstrated in Chapter 3.

In reality, the multi-step nature of the radical reactions dictates that oxidative modifications can be caused by any radical and not just  $\cdot\text{OH}$  (21). The implicit assumption behind the FPOP methodology and Hambly's estimation of  $\tau$  is that the offspring radicals generated by the reaction of  $\cdot\text{OH}$  with Gln are stable and nonreactive towards protein side chains. The only way to remove radicals from the environment is via recombination with other radicals. Radical traps such as 2,2,6,6-Tetramethylpiperidin-1-yl)oxyl (TEMPO) have been developed to slow or stop radical reactions by the formation of persistent and stable radicals stabilized by a combination of electronic, steric factors (23). To the best of my knowledge, no such properties have been found for Gln radicals. The lower than expected amounts of unmodified proteins Chapter 2, Chapter 3, and references (22, 24) further expose our lack of understanding of FPOP. Thus, it would be of immense interest to directly measure  $\tau$  using an experimental approach.

A possible strategy is a pump/probe approach. A UV-transparent quartz cuvette can be used as a reaction chamber that contains the typical constituents of an FPOP

experiment in addition to a dye present at  $\mu\text{M}$  quantities. The firing of the excimer (pump) laser acts as the trigger and produces  $\cdot\text{OH}$ . Upon modification by any radical species in solution, the extensive pi bonding network of the dye is disrupted. This would result in a change in its spectroscopic properties which the probe laser monitors. For instance, the absorbance of the dye monitored by the probe laser operated at a suitable wavelength is taken as a reporter of  $\tau$ . Detection of probe laser signal can be performed using a photodiode and a high frequency oscilloscope. Similar experiments can be designed that exploit fluorescence. It seems reasonable that the strategy described above advances our understanding of FPOP and reveals how established radical traps such as TEMPO (25) fair in comparison to Gln.

### *6.2.2 Pushing the HDX-MS Envelope*

Successful completion of biological processes in a cellular environment often entails several players acting as part of a team. This calls for methods capable of uncovering the interaction of proteins of interest in a large background of other biomolecules. The data presented in Chapter 5 showed that HDX-MS fits this description, where the dynamics of the 550 kDa  $\text{F}_0\text{F}_1$  molecular machine was interrogated in the presence of *ca.* 800 kDa worth of unique sequence of background proteins. Impressive as it is, this does not represent a limit. Improved LC and IM separation can provide the necessary enhancements in resolving power (peak capacity) to cross this barrier and study more complex systems. The bottleneck often is the limited separation power of LC at 0 °C. Modern UPLCs capable of providing increasingly higher backpressures have alleviated the situation by enabling higher flow rates in columns with smaller stationary phase (26).

This provides higher peak capacity without increasing the separation time and deuterium label loss. Utilizing microchip capillary electrophoresis (27) in tandem with IMS promises to be a powerful combination for rapid sample analysis. Also desirable is improving the sequence coverage for regions that reside in the membranes. HDX-MS of membrane proteins continues to lag far behind that of soluble systems (1). Poor digestion under typical HDX quench conditions is the main culprit. Thus, studies of membrane proteins are often performed on detergent-solubilize systems (28) or in nanodiscs (29). The application of HDX-MS to proteins in their natural lipid environment is rare. It would be interesting to see if the use of pepsin-immobilized columns operated at high pressure is a more effective strategy (30).

The improvements outlined above should enable HDX-MS of ever larger and complex systems. An interesting opportunity as a follow-up to what was reported in Chapter 5 is studying cellular respiration using the membrane vesicles. For instance, membranes energized with excess protons can be prepared by oxidizing NADH or succinate by complexes I and II of the respiratory chain, respectively. The generated PMF is then employed to drive  $F_0F_1$  in the synthesis direction to produce ATP. Continued development of HDX-MS promise to enable complete characterization of macromolecular complexes such as the ribosome, the nucleosome, and photosynthetic complexes.

### 6.3 References

1. Konermann L, Vahidi S, & Sowole MA (2014) Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. *Anal. Chem.* 86: 213-232.

2. Hambly DM & Gross ML (2005) Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* 16: 2057-2063.
3. Tong X, Wren JC, & Konermann L (2008)  $\gamma$ -Ray-Mediated Oxidative Labeling for Detecting Protein Conformational Changes by Electrospray Mass Spectrometry. *Anal. Chem.* 80: 2222-2231.
4. Goswami D, Devarakonda S, Chalmers MJ, Pascal BD, Spiegelman BM, & Griffin PR (2013) Time Window Expansion for HDX Analysis of an Intrinsically Disordered Protein. *J. Am. Soc. Mass Spectrom.* 24: 1584-1592.
5. Yang WY & Gruebele M (2003) Folding at the speed limit. *Nature* 423: 193-197.
6. Xu M, Beresneva O, Rosario R, & Roder H (2012) Microsecond Folding Dynamics of Apomyoglobin at Acidic pH. *J. Phys. Chem. B* 116: 7014-7025.
7. Stocks BB & Konermann L (2009) Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* 81: 20-27.
8. Wytttenbach T, Pierson NA, Clemmer DE, & Bowers MT (2014) Ion Mobility Analysis of Molecular Dynamics. *Annu. Rev. Phys. Chem.* 65: 175-196.
9. Ruotolo BT & Robinson CV (2006) Aspects of native proteins are retained in vacuum. *Curr. Op. Chem. Biol.* 10: 402-408.
10. Wytttenbach T & Bowers MT (2011) Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility-Mass Spectrometry Environment. *J. Phys. Chem. B* 115: 12266-12275.
11. Eliezer D, Yao J, Dyson HJ, & Wright PE (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* 5: 148-155.
12. Pirrone GF, Iacob RE, & Engen JR (2015) Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Anal. Chem.* 87(1): 99-118.

13. Konermann L, Pan J, & Liu Y (2011) Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem. Soc. Rev.* 40: 1224-1234.
14. Junge W, Sielaff H, & Engelbrecht S (2009) Torque generation and elastic power transmission in the rotary F<sub>0</sub>F<sub>1</sub>-ATPase. *Nature* 459(7245): 364-370.
15. Abrahams JP, Leslie AGW, Lutter R, & Walker JE (1994) Structure at 2.8 Å resolution of F<sub>1</sub>-ATPase from bovine heart mitochondria. *Nature* 370: 621-628.
16. Cingolani G & Duncan TM (2011) Structure of the ATP synthase catalytic complex (F<sub>1</sub>) from *Escherichia coli* in an autoinhibited conformation. *Nat. Struct. Mol. Biol.* 18: 701-708.
17. Sielaff H & Borsch M (2013) Twisting and subunit rotation in single F<sub>0</sub>F<sub>1</sub>-ATP synthase. *Philos. Trans. R. Soc. B-Biol. Sci.* 368(1611).
18. Martin JL, Ishmukhametov R, Hornung T, Ahmad Z, & Frasch WD (2014) Anatomy of F<sub>1</sub>-ATPase powered rotation. *Proc. Natl. Acad. Sci. U.S.A.* 111(10): 3715-3720.
19. Uchihashi T, Iino R, Ando T, & Noji H (2011) High-Speed Atomic Force Microscopy Reveals Rotary Catalysis of Rotorless F<sub>1</sub>-ATPase. *Science* 333(6043): 755-758.
20. Kapulainen M, Ronkainen H, Hokkanen A, Stuns I, Varjus S, Nyysönen S, Turunen R, & Halme J (2014) Fibre Optic Sensors for Long-Term Monitoring of Oil Film Pressure in Diesel Engine Main Bearing. *Tribol. Lett.* 56(1): 47-54.
21. Xu G & Chance MR (2007) Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* 107: 3514-3543.
22. Gau BC, Sharp JS, Rempel DL, & Gross ML (2009) Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* 81: 6563-6571.
23. Matyjaszewski K (1998) Radical Nature of Cu-Catalyzed Controlled Radical Polymerizations (Atom Transfer Radical Polymerization). *Macromolecules* 31(15): 4710-4717.

24. Chen J, Rempel DL, & Gross ML (2010) Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* 132: 15502-15504.
25. Pattison DI, Lam M, Shinde SS, Anderson RF, & Davies MJ (2012) The nitroxide TEMPO is an efficient scavenger of protein radicals: Cellular and kinetic studies. *Free Radical Biology and Medicine* 53(9): 1664-1674.
26. Wales TE, Fadgen KE, Gerhardt GC, & Engen JR (2008) High-Speed and High-Resolution UPLC Separation at Zero Degrees Celsius. *Anal. Chem.* 80(17): 6815-6820.
27. Black WA, Stocks BB, Mellors JS, Engen JR, & Ramsey JM (2015) Utilizing Microchip Capillary Electrophoresis Electrospray Ionization for Hydrogen Exchange Mass Spectrometry. *Anal. Chem.* 87: 6280-6287.
28. Pan Y, Piyadasa H, O'Neil JD, & Konermann L (2012) Conformational Dynamics of a Membrane Transport Protein Probed by H/D Exchange and Covalent Labeling: The Glycerol Facilitator. *J. Mol. Biol.* 416: 400-413.
29. Parker CH, Morgan CR, Rand KD, Engen JR, Jorgenson JW, & Stafford DW (2014) A Conformational Investigation of Propeptide Binding to the Integral Membrane Protein  $\gamma$ -Glutamyl Carboxylase Using Nanodisc Hydrogen Exchange Mass Spectrometry. *Biochemistry* 53(9): 1511-1520.
30. Ahn J, Jung MC, Wyndham K, Yu YQ, & Engen JR (2012) Pepsin Immobilized on High-Strength Hybrid Particles for Continuous Flow Online Digestion at 10 000 psi. *Anal. Chem.* 84(16): 7256-7262.

# Appendix I-Permissions



RightsLink®

Home

Create Account

Help



**Title:** Unraveling the Mechanism of Electrospray Ionization

**Author:** Lars Konermann, Elias Ahadi, Antony D. Rodriguez, et al

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Jan 1, 2013

Copyright © 2013, American Chemical Society

LOGIN  
If you're a **copyright.com user**, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink user** or want to [learn more?](#)

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK

CLOSE WINDOW

Copyright © 2015 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)





RightsLink®

Home

Create Account

Help



ACS Publications  
Most Trusted. Most Cited. Most Read.

**Title:** Mapping pH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry

**Author:** Siavash Vahidi, Bradley B. Stocks, Yalda Liaghati-Mobarhan, et al

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Nov 1, 2012

Copyright © 2012, American Chemical Society

LOGIN

If you're a **copyright.com user**, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink user** or want to [learn more?](#)

#### PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK

CLOSE WINDOW

Copyright © 2015 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at [customer@copyright.com](mailto:customer@copyright.com)



**Title:** Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling

**Author:** Siavash Vahidi, Bradley B. Stocks, Yalda Liaghati-Mobarhan, et al

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Sep 1, 2013

Copyright © 2013, American Chemical Society

[LOGIN](#)  
If you're a [copyright.com user](#), you can login to RightsLink using your copyright.com credentials. Already a [RightsLink user](#) or want to [learn more?](#)

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)



**Title:** Partially Disordered Proteins Studied by Ion Mobility-Mass Spectrometry: Implications for the Preservation of Solution Phase Structure in the Gas Phase

**Author:** Siavash Vahidi, Bradley B. Stocks, Lars Konermann

**Publication:** Analytical Chemistry

**Publisher:** American Chemical Society

**Date:** Nov 1, 2013

Copyright © 2013, American Chemical Society

[LOGIN](#)

If you're a [copyright.com user](#), you can login to RightsLink using your copyright.com credentials. Already a [RightsLink user](#) or want to [learn more?](#)

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)

[CLOSE WINDOW](#)

# Curriculum Vitae

Siavash Vahidi

Ph.D. Candidate

Department of Chemistry

*The University of Western Ontario (UWO)*

Research Group Homepage: <http://publish.uwo.ca/~konerman>

---

## Education

<u>Degree</u>	<u>Institution</u>	<u>Department</u>	<u>Year</u>
Ph.D.	University of Western Ontario London, ON, Canada (with Lars Konermann)	Chemistry	2010-present
B.Sc. <i>with Honours</i>	Shahid Beheshti University (National University of Iran)- Tehran, Iran	Chemistry	2006-2010

## Employment History

<u>Date</u>	<u>Rank and Position</u>	<u>Department</u>	<u>Institution</u>
2010-Present	Research Assistant (Ph.D.)	Chemistry	University of Western Ontario
2010-Present	Teaching Assistant	Chemistry	University of Western Ontario

## Honours, Awards and Scholarships

<u>Date</u>	<u>Award Name</u>	<u>Type</u>	<u>Location of Tenure</u>	<u>Amount (\$CAD)</u>
2014-May to April-2015	Ontario Graduate Scholarship (OGS)/ Queen Elizabeth II Graduate Scholarship in Science and Technology	Provincial (Research)	University of Western Ontario	\$15,000
2013-Sept	Canadian Society for Mass Spectrometry (CSMS) Award	National (Research)	University of Western Ontario	\$2000
2013-July	Various Travel Awards	Institutional (Research)	University of Western Ontario	\$1160

### Selected Articles in Peer-Reviewed Journals

1. "Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules" L. Konermann, **S. Vahidi**, and M. A. Sowole, *Anal. Chem.* 86, 213-232 (2014).
2. "Partially Disordered Proteins Studied by Ion Mobility-Mass Spectrometry: Implications for the Preservation of Solution Phase Structure in the Gas Phase" **S. Vahidi**, B. B. Stocks, and L. Konermann, *Anal. Chem.* 85, 10471-10478 (2013).
3. "Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling" **S. Vahidi**, B. B. Stocks, Y. Liaghati-Mobarhan, and L. Konermann, *Anal. Chem.* 85, 8618-8625 (2013). **Cover Feature.**
4. "Effects of Ammonium Bicarbonate on the Electrospray Mass Spectra of Proteins: Evidence for Bubble-Induced Unfolding" J. B. Hedges, **S. Vahidi**, X. Yue, and L. Konermann, *Anal. Chem.* 85, 6469-6476 (2013).
5. "Unraveling The Mechanism of Electrospray Ionization" L. Konermann, E. Ahadi, A. D. Rodriguez, and **S. Vahidi**, *Anal. Chem.* 85, 2-9 (2013). **Cover Feature.**
6. "Mapping pH-Induced Protein Structural Changes Under Equilibrium Conditions by Pulsed Oxidative Labeling and Mass Spectrometry" **S. Vahidi**, B. B. Stocks, Y. Liaghati-Mobarhan, and L. Konermann, *Anal. Chem.* 84, 9124-9130 (2012).

### Selected Conference Presentations: (Presenting author in bold)

1. Oral: "Investigating the Importance of Protein Conformational Dynamics During Catalysis by HDX-MS: Focus on the F<sub>0</sub>F<sub>1</sub>-ATP Synthase Molecular Machine" **S. Vahidi**, Y. Bi, S. D. Dunn, and L. Konermann, *Proceedings of the 63rd ASMS Conference on Mass Spectrometry and Allied Topics*; St. Louis, MO, (2015).
2. Oral: "Strategies for Minimizing Spurious In-Source CID for Peptides during ESI-MS" **S. Vahidi**, and L. Konermann, *Proceedings of the 62nd ASMS Conference on Mass Spectrometry and Allied Topics*; Baltimore, MD, (2014).
3. Oral: "Do Electrosprayed Protein Ions Retain Memory of Their Solution Phase Structure? Insights from Ion Mobility Mass Spectrometry" **S. Vahidi**, B. B. Stocks, and L. Konermann, *Proceedings of the 61st ASMS Conference on Mass Spectrometry and Allied Topics*; Minneapolis, MN, (2013).
4. Poster: "Structural Interrogation of Partially Folded Proteins by Pulsed Oxidative Labeling and ESI-MS" **S. Vahidi**, B. B. Stocks, and L. Konermann, *Proceedings of the 60th ASMS Conference on Mass Spectrometry and Allied Topics*; Vancouver, BC, (2012).
5. Poster: "Conformational Transitions of Apo-Myoglobin Probed by Hydroxyl Radical Labeling and Mass Spectrometry" **S. Vahidi**, Y. Liaghati-Mobarhan, and L. Konermann, *Proceedings of the 59th ASMS Conference on Mass Spectrometry and Allied Topics*; Denver, CO, (2011).