



The University of
Nottingham

UNITED KINGDOM • CHINA • MALAYSIA

The Use of Protein Modification and Ion Mobility-Mass Spectrometry
to Probe Protein Structure

By

Asia Khalaf HN Al-jabiry

Thesis submitted to the University of Nottingham for the degree
of Doctor of Philosophy
March 2021

Abstract

Mass spectrometry (MS) is considered to be indispensable technology for the use in modern pharmaceutical drug discovery and development processes. However, MS is rarely used as a screening technology for protein structure. In this project, ion mobility-mass spectrometry (IM-MS) methods are developed to investigate protein structure with the use of chemical modification and genetic modification.

Collision induced unfolding (CIU) method was optimised for measuring the mobility of ubiquitin (Ub) drift traces and the collision cross section (CCS) was calculated. The mobility was measured in the trap by acquiring several voltages and monitoring the drift trace of the lower state ion ($[M+6H]^{6+}$ and $[M+5H]^{5+}$). By combining the CIU method and chemical modification of proteins we can enhance the understanding of protein structure in the gas phase.

Acetylation was carried out first on ubiquitin, the results showed a difference in the drift trace for ubiquitin after acetylation. This led to inspection of the MS/MS spectrum of intact Ub. The b-ion, corresponding to fragmentation at lysine residue K6, showed this residue to have importance for the structural integrity of the protein. Therefore, different K6 mutant were obtained and their CIU were acquired. The results confirmed that the K6 residue is indeed crucial in the ubiquitin unfolding pathway. Acetylation of this residue, or its replacement with alanine (K6A Ub) produced a conformationally destabilised form of the protein, which unfolded at lower collision energies. Wild type Ub and its mutant K6O mutant shared the fact the K6 is present, and the result showed they have the same CIU unfolding profile, In contrast the NoK and K6R mutant where the K6 has been modified to R, resulted in a more stable compact structure as evidenced by the CIU profile.

Diethylpyrocarbonate modification of the single histidine residue in Ub, which was postulated to interact with K6 in the gas phase structure resulted in modest destabilisation of compact Ub, while succinylation of the N-terminus had no clear effect on stability of the protein structure. Studies of molecular dynamics and charge distribution support the experimental data by rationalising the importance of protonated K6 and H68 interaction in the gas-phase stabilization of the native folding of Ub.

Finally, the ubiquitin associating domain UBA2 was destabilised by adding an acetyl group to the N-terminus of the protein. The observation was interpreted by the breaking of a key intramolecular interaction between the N-terminus and the glutamate residue E22. Moreover, the behaviour of different in charge states showed the importance of addition of charge on the structure of proteins.

Publications

Combined Chemical Modification and Collision Induced Unfolding Using Native Ion Mobility-Mass Spectrometry Provides Insights into Protein Gas-Phase Structure. Al-

jabiry, A., Palmer, M., Langridge, J., Bellamy-Carter, J., Robinson, D., & Oldham, N. J.

(2021).. *Chemistry – A European Journal*, 27(55), 13783-13792.

Acknowledgment

First and foremost, I would like to express my great appreciation to Professor Neil Oldham for his excellent supervision, his patience and guidance through all these years and through all the hardship, he was always there supporting and was exceptionally great mentor.

I want to thank my fellow student Jedd Bellamy-Carter, for his constant assistance and help through this project so far, and for sharing his thoughts in countless valuable discussions. I would like to thank all the members of the Oldham group in particular: Dr Vanderson Barbosa Bernado, Dr Mariana Breda, Bruna da Silva Granja and the many project students to have come through our lab for providing an inspiring and fun environment in which to learn.

I'm grateful to Dr. Robert Layfield and Dr. Dan Scott for their collaboration through their help in the expression and purification of the UBD proteins, and for providing the protein mutants and purification experiments.

I am particularly thankful for Prof. David Robinson, for his work and time spent on the Molecular Dynamics simulation and for the exciting collaboration.

I wouldn't be able to continue this journey without my sisters and my brother they were my friends during the hardest parts. Finally, to my Father and Mother who always encouraged me and fought for me to pursue my dreams.

Contents

Abstract	3
Publications.....	4
Acknowledgment	5
1 Introduction.....	2
1.1 Protein structure and interactions	2
1.1.1 Levels of protein structure	3
1.1.2 Protein folding and unfolding	10
1.1.3 Ubiquitin	12
1.1.4 Ubiquitin-binding domains (UBD).....	14
1.2 Mass spectrometry.....	16
1.2.1 Methods of ionization	17
1.2.2 Mass analysers	20
1.2.3 Tandem mass spectrometry.....	22
1.3 Ion mobility mass spectrometry	26
1.3.1 Ion mobility.....	26
1.3.2 Traveling wave ion mobility spectrometry	28
1.3.3 Collision cross section calculation.....	29
1.3.4 Collision Induced Unfolding.....	33
1.4 Protein structure studies by ion mobility mass spectrometry.....	36
1.5 Chemical modification of proteins	38
1.5.1 Acetylation.....	38
1.5.2 Carboethoxylation of Histidine.....	40
1.5.3 Succinylation.....	41
1.6 Molecular dynamics simulation	43
1.7 Aims and objectives	45
2 Materials and methods	47
2.1 Ubiquitin.....	47
2.1.1 Ubiquitin mutants.....	47
2.2 Protein expression and purification.....	49
2.2.1 Ubiquitin binding domains	49
2.2.2 Preparation of ultra-competent cells	49
2.2.3 Preparation of solutions for protein overexpression and purification.....	51
2.2.4 Protein expression and sample preparation.....	52

2.3	Expression and purification UBA2s.....	53
2.4	Site-directed mutagenesis.....	54
2.5	Acetylation reaction	55
2.6	Diethylpyrocarbonate reaction	55
2.7	Succinylation reaction	56
2.8	Fourier Transform Ion Cyclotron Resonance (FT-ICR)	56
2.9	Ion mobility-mass spectrometry	57
2.9.1	Mobility calibration standard.....	57
2.9.2	Calibration and CCS calculation.....	57
2.9.3	Software and application.....	59
2.10	Molecular dynamics simulation.....	59
3	Chemical modification.....	62
3.1	Introduction	62
3.1.1	The optimisation of acetylation on wild type ubiquitin (WT-Ub).....	63
3.1.2	Tandem mass spectrometry.....	65
3.2	Diethylpyrocarbonate modification.....	69
3.2.1	Modification of histidine and lysine residues in proteins by diethylpyrocarbonate.....	69
3.2.2	Tandem mass spectrometry ECD fragmentation for DEPC-based modification	71
3.3	Succinylation modification	74
3.3.1	The use of succinic anhydride to modify lysine residues in proteins	74
3.3.2	Tandem mass spectrometry ECD fragmentation for mapping succinylation sites	76
3.4	Conclusions	79
4	Collision-induced unfolding of chemically modified ubiquitin.....	81
4.1	Collision-induced unfolding of acetylated ubiquitin.....	81
4.1.1	CCS for WT-Ub vs Acetylated WT-Ub for $[M+5H]^{5+}$	82
4.1.2	CCS for WT-Ub vs Acetylated WT-Ub for $[M+6H]^{6+}$	87
4.2	Collision-induced unfolding for DEPC and SUCC modification	89
4.2.1	$[M+5H]^{5+}$ ion CCS comparison for WT-Ub vs. DEPC-modified-Ub	89
4.2.2	$[M+5H]^{5+}$ ion CCS comparison for the WT-Ub vs. SUCC-Ub.....	91
4.3	Conclusions	93
5	Collision-induced unfolding of ubiquitin mutants.....	95

5.1	Introduction	95
5.2	[M+5H] ⁵⁺ ion CCS comparison for mutants vs. WT-Ub	96
5.2.1	Ubiquitin mutant (K6A).....	96
5.2.2	Ubiquitin mutant (K6O).....	98
5.2.3	Ubiquitin mutant (K6R).....	100
5.2.4	Ubiquitin mutant (NoK).....	103
5.3	[M+6H] ⁶⁺ ion CCS comparison for the mutants vs. WT-Ub	104
5.3.1	Ubiquitin mutant (K6A).....	105
5.3.2	Ubiquitin mutant (K6O).....	106
5.3.3	Ubiquitin mutant (K6R).....	107
5.3.4	Ubiquitin mutant (NoK).....	108
5.4	Conclusions	109
6	Molecular dynamics and charge distribution.....	113
6.1	Introduction	113
6.2	Experimental <i>CCS</i> equated to theoretical <i>CCS</i>	115
6.3	K6A and lysine histidine interaction	118
6.4	Conclusions	123
7	Ubiquitin-binding domains UBA2.....	126
7.1	Introduction	126
7.2	Acetylation of UBA2	126
7.2.1	CCS of Acetylated of UBA2.....	129
7.3	Collision Induced Unfolding of UBA2 Mutant.....	131
7.3.1	CCS for UBA2 mutant K14A	131
7.3.2	CCS for UBA2 mutant K14H.....	132
7.4	Conclusions	132
8	Conclusions and Outlook.....	135
	References.....	139

List of Abbreviations

ATD Arrival time distribution

CCS Collisional Cross Section

CID Collision-induced dissociation

CIU Collision-induced unfolding

CSD Charge State Distribution

DEPC Diethylpyrocarbonate

ECD Electron Capture Dissociation

ESI Electrospray ionisation

ETD Electron Transfer Dissociation

FT-ICR Fourier Transform Ion Cyclotron Resonance

HDX Hydrogen-Deuterium Exchange

HPLC High Performance Liquid Chromatography

IMS Ion Mobility Mass Spectrometry

K_d Dissociation constant

LB Luria broth

LC Liquid chromatography

m/z Mass-to-charge ratio

MALDI Matrix Assisted Laser Desorption Ionisation

MD Molecular dynamics

MS Mass Spectrometry

MS/MS Tandem mass spectrometry

nESI Nano-electrospray ionisation

NMR Nuclear Magnetic Resonance

PCR Polymerase chain reaction

PDB Protein databank

PTM Post-translational modification

Q-ToF Quadropole–Time-of-flight hybrid

SDS-PAGE Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

td Drift time

ToF Time of flight

TWIG Travelling Wave Ion Guide

TWIMS Travelling-wave ion mobility spectrometry

Ub Ubiquitin

UBD Ubiquitin Binding Domain

Chapter 1

Introduction

1 Introduction

1.1 Protein structure and interactions

Generally, a protein molecule consists of a long chain of amino acids linked by covalent peptide bonds. Depending on organism between 20 and 22 of the 300 amino acids present in nature are known as proteinogenic amino acids, and they are found in the structure of proteins seen in various forms of life. Each protein has a unique sequence of amino acids that define its structure and function. As seen in Figure 1-1, the general formula of an amino acid is $\text{H}_2\text{N}-\text{CH}(\text{R})-\text{COOH}$. Herein, R is used to describe the amino acid side chain, which is composed of one of 20 amino side chain options. For example, one of the simplest amino acids, glycine, has just one hydrogen atom in the position of the R-group (Butterworth, 2005).

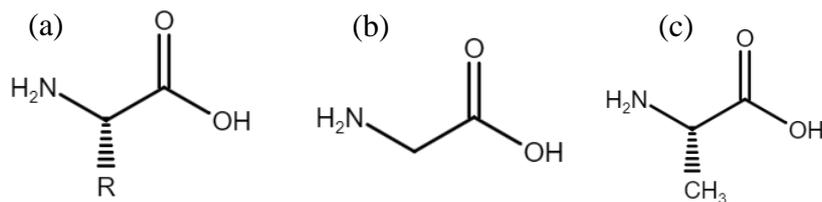
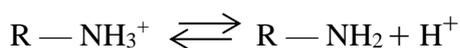
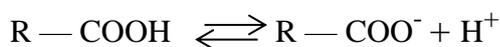


Figure 1-1. (a) Generic amino acid structure, (b) glycine, and (c) alanine.

The chemical structure of amino acids consists of charged and uncharged forms of the ionizable $-\text{COOH}$ and $-\text{NH}_3^+$ weak acid groups that are in equilibrium in a solution:



As shown, both $\text{R}-\text{COOH}$ and $\text{R}-\text{NH}_3^+$ are weak acids, while $\text{R}-\text{COOH}$ is a far stronger acid than $\text{R}-\text{NH}_3^+$. At physiological pH (pH 7.4), carboxyl groups exist almost entirely as $\text{R}-\text{COO}^-$ and amino groups predominantly as $\text{R}-\text{NH}_3^+$. Molecules that contain an equal number of ionizable groups of opposite charge – and hence carry no net charge – are termed

‘zwitterions.’ In the context of peptides and proteins, where the amino and carboxylate groups are linked to form a neutral peptide bond, only the free N-terminal amino group and C-terminal carboxylate group carry a charge. The majority of charges in the molecule arise from protonation or deprotonation of ionisable sidechains, classification of amino acids based on their structure and chemical properties, Based on their polarity, amino acids are normally divided into four categories. The polarity of amino acids in protein structure reflects their functional role (Butterworth, 2005).

The amino acid sequence of a protein is taken from the translation of mRNA, which in turn is derived from the transcription of a DNA sequence. The process is then governed by DNA, the cell's genetic material. Some of the incorporated amino acids undergo modification to create their derivatives during protein synthesis. When gene mutations occur, the DNA sequence is altered and can cause a change in the amino acid sequence of a protein. Changing a single amino acid in a protein's sequence can have a dramatic effect on the protein's structure and function (Butterworth, 2005).

1.1.1 Levels of protein structure

As indicated before, the shape of a protein is crucially important to its function. There are four levels of protein structure: primary, secondary, tertiary, and quaternary. The primary structure determines the other three levels of structure. A protein's primary structure consists of variable combinations of the 20 different amino acids to form polypeptide chains. These are translated from mRNA by the ribosome. Thus, there is a direct link between gene sequence and protein structure.

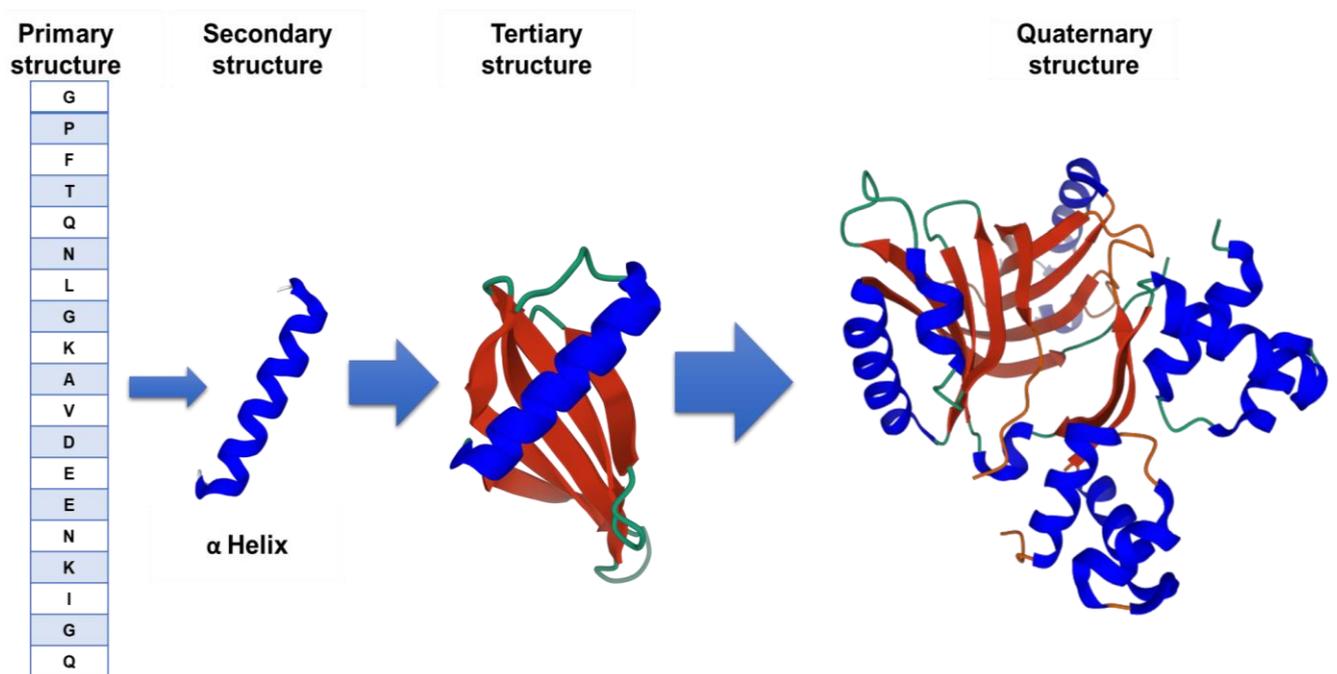


Figure 1-2. the four level of protein structure. first the primary structure represents the amino acid sequence, the secondary structure shows the of crystal structure of α helix Monellin, third the tertiary structure shows the α helix and the β pleated sheet of single chain Monellin, finally the quaternary structure shows all the bonded sheet to make a protein.

Along the protein sequence, it is possible for some amino acid to be different in some regions without impacting the overall shape the biological function, while other regions of the protein contain crucial sites that are essential for structure/function and the amino acid sequence must be preserved during transcription. In cases where mutation results in the change of one amino acid, serious genetic diseases can occur (Alberts & Press, 2017).

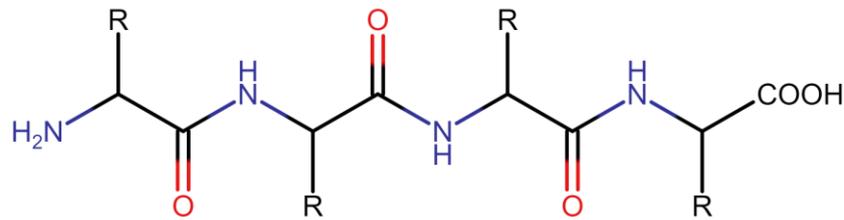


Figure 1-3. Protein primary structure showing backbone and covalent bonding.

The next level of protein structure is the secondary structure. Although each protein has a unique conformation; two regular folding patterns are commonly found within proteins.

These are the α helix and the β pleated sheet. Many different sets of weak noncovalent bonds that develop between one segment of the chain and another determine how a protein chain folds. Atoms in the polypeptide backbone and atoms in the amino acid side chains are involved. The strength of a large number of such noncovalent connections determines the stability of each folded shape in this way (Alberts & Press, 2017).

The propensity is calculated as the ratio of the frequency with which an amino acid occurs in a specific fold of a secondary structure to the frequency with which that same amino acid occurs in all proteins. The interchangeability of amino acids is achieved by maintaining their secondary structural propensity, and the mutability of amino acids is directly proportional to the amount of energy required for their biosynthesis while being inversely proportional to the frequency of their most probable conformation. (Bohórquez, 2017).

The global three-dimensional structure of a polypeptide constitutes the tertiary structure. Generally, the tertiary structure forms due to interactions between the R groups of the amino acids in the primary structure. The structure of protein is mainly formed by the non-covalent bonds affecting the binding between the amino acids in the protein molecule to form the more complex conformation, which gives the protein its function. Forces such as ionic interactions,

hydrogen bonds, Van der Waals forces and the hydrophobic effect all contribute to the tertiary structure. Noncovalent bonds are 30–300 times weaker than the covalent connections that hold biological molecules together. However, sections of a polypeptide chain can be held together strongly by a number of weak bonds operating in combination. The strength of a large number of such noncovalent connections determines the stability of each folded shape in this way. These relatively weak bonds also enable fast, reversible binding between proteins and between proteins with small molecules. The sequence of amino acids and how they interact to form the conformation with the lowest energy are the principal factors that give a protein its global tertiary structure (Alberts & Press, 2017).

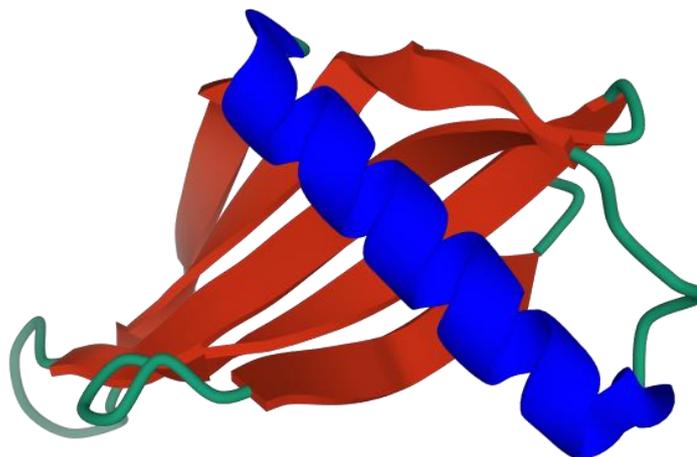


Figure 1-4. The tertiary structure structure of the polypeptide backbone forming the α helix (blue) and the β sheet (red).

A fourth weak force, known as the hydrophobic clustering force, which is really driven by entropic considerations, is also important in defining how a protein is shaped. In an aqueous environment, hydrophobic molecules, such as the nonpolar side chains of specific amino acids, are pulled together to limit their disruption of the hydrogen-bonded network of water molecules. As a result, the distribution of polar and nonpolar amino acids in any protein plays an important role in its folding.

Nonpolar (hydrophobic) side chains of proteins, such as those found in phenylalanine, leucine, valine, and tryptophan, tend to congregate in the molecule's interior (just as hydrophobic oil droplets coalesce in water to form one large droplet). They are able to avoid contact with the water that surrounds them inside a cell as a result of this (Andersen, 2001).

Polar groups, on the other hand, such as those found in arginine, glutamine, and histidine, tend to cluster at the molecule's surface, where they can form hydrogen bonds with water and other polar molecules. Polar amino acids that are buried deep inside the protein are frequently hydrogen-bonded to other polar amino acids or the polypeptide backbone (Alberts & Press, 2017).

In the example of ionic interactions, the ions in aqueous solution are surrounded by water molecules, and these water molecules shield the ionic charges from each other, resulting in the ionic interactions on the surface of a protein being weaker than those inside, where there is usually little to no water. In hydrogen bonds the interaction is formed between a hydrogen atom and an electronegative atom. Van der Waals forces are considered one of the weakest forces in nature, as they are formed by the attraction of induced dipole moments in non-polar molecules (Andersen, 2001). The attraction between the adjacent atoms generates electron clouds of opposite polarity that favours charge distributions in the next atom, which then forms a weak bond that is referred to as “dispersion” forces (also known as “London” forces), which are weaker than H-bonds (Andersen, 2001). In proteins they play an important role due to their weak nature and allow for flexible movement in large macromolecules. Most proteins contain hydrophobic amino acid side chains, which do not favour the aqueous environment since they cannot form hydrogen or ionic bonds with water. Polar molecules interact strongly with each other, which results in the hydrophobic residues being forced into the core of the protein. A combination of enthalpic interactions and entropic effects (principally the hydrophobic effect)

leads to hydrophobic residue being buried inside the protein. Figure 1-5 shows the bonds that hold the tertiary structure together (Andersen, 2001).

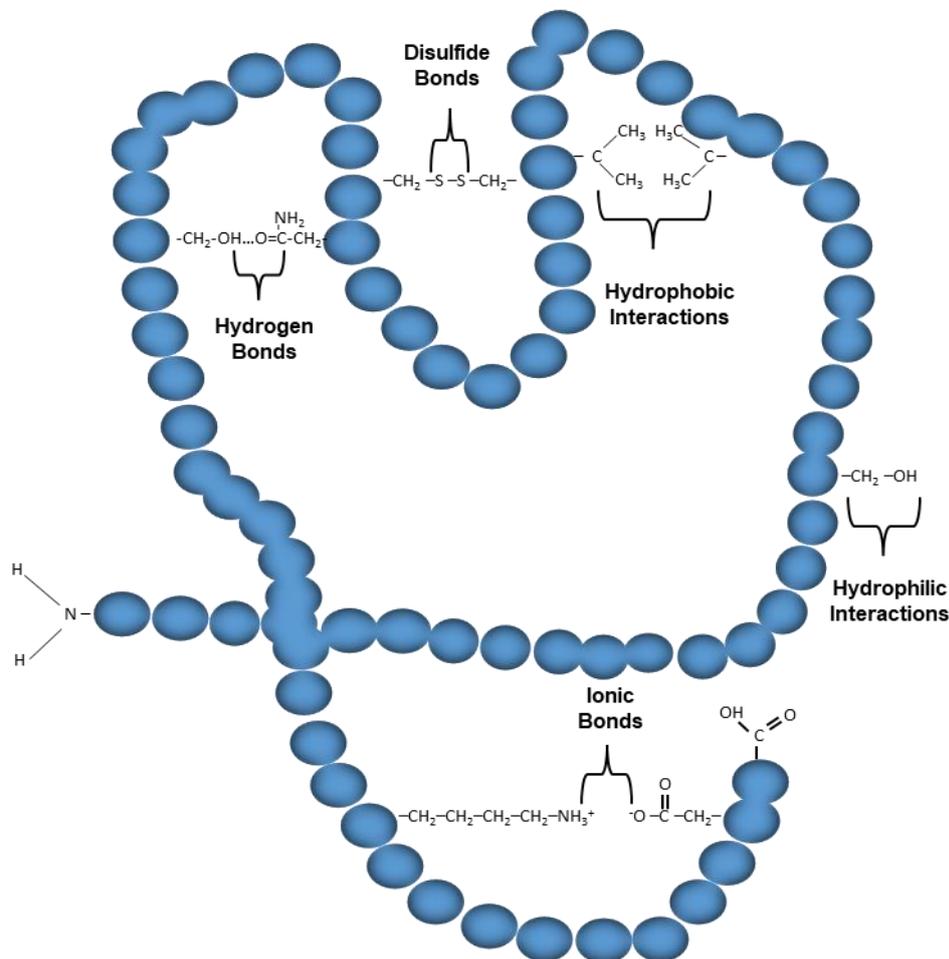


Figure 1-5. Types of side chain interactions.

These interactions can also result in quaternary structure, which is a protein molecule forming a complex of more than one polypeptide chain to form a multi subunit protein or oligomer structure. Oligomeric proteins can have many biological activities.

Examples of proteins with quaternary structure include enzymes, cell membrane proteins, and contractile proteins such as actin and myosin. Haemoglobin is an example of quaternary protein, and consists of a total of four subunits (two alpha subunits and two beta subunits). These units are held together by hydrophobic interactions, hydrogen bonding, and salt bridges,

which gives it confirmation, thereby allowing each subunit that contains a central haeme group to bind one oxygen molecule.

Finally, an intrinsically disordered protein (IDP) is a protein that lacks a fixed or ordered three-dimensional structure, due to the absence of macromolecular interaction partners like other proteins or RNA. Random coil, molten globule-like aggregates, or flexible linkers in massive multi-domain proteins are examples of unstructured to partially organized IDPs. Along with globular, fibrous, and membrane proteins, they are sometimes regarded a different class of proteins (Dunker et al., 2001).

The ordered state, the molten globule, and the random coil are three different forms that proteins can take. The five examples below show that natural protein structure can correspond to any of the three states (not just the ordered state), and that protein function can derive from any of the three states and transitions. protein that are in a process that likely mimics infection, Nucleosome hyperacetylation is crucial to DNA replication and transcription, Clustering contains an ordered domain and also a native molten globular region and in a critical signalling event, a helix in calcineurin becomes bound and surrounding (Dunker et al., 2001).

1.1.2 Protein folding and unfolding

The study of protein folding, structure, as well as thermodynamic and kinetic properties is important due to the close relationship between genetic diseases commonly found in humans, such as misfolded proteins that form extracellular or intracellular aggregates to initiate profound cellular dysfunction. Because of advances in experimental and theoretical methodologies, as well as their integration, scientific understanding of protein folding has increased in the last decade. Understanding how a protein folds to its active conformation is only possible if the pathway between the folded and unfolded states is known; the intermediate states and transition states that occur between the beginning and final states define a pathway.

Protein folding pathways have the same potential for complexity as the variety of three-dimensional conformations that can be achieved. Within any practical population of unfolded protein molecules, each molecule is likely to have a conformation distinct from that of every other molecule in the population at any one time; a corresponding diversity of routes must exist, at least in the early stages of folding.

Studying protein folding and unfolding mechanisms can help us understand how such diseases arise and can potentially lead to preventive cures. (Selkoe, 2003). Several experimental methods have been developed through time to measure the kinetics of protein folding and unfolding such as stopped-flow methods and continuous-flow techniques. (Crabtree & Shammas, 2018; Roder, Maki, & Cheng, 2006).

1.1.2.1 Stopped-flow methods

Stopped-flow is an experimental technique for studying chemical processes with a half time of less than 1 millisecond. The stopped-flow technique is performed by mixing solutions in a small cell. A stopped-flow, in its most basic form, combines two solutions, for example a protein and a denaturant. Small amounts of solutions are introduced into a high-efficiency mixer in a continuous and quick manner. This mixing method then initiates a fast reaction. The newly mixed solution makes its way to the observation cell, pushing the contents out (Clark, 1997).

The mixing time is typically between 1 and 10 ms, while the reaction is being monitored by real-time optical absorbance (UV through IR), fluorescence emission, or circular dichroism (CD) spectroscopy. Despite the short measurement time, the reaction is sensitive and considered a powerful kinetic method of measuring the protein unfolding mechanism (Roder et al., 2006). However, the stopped-flow technique has its disadvantages, and in some cases the timescale of the reaction can be altered by changing the apparatus, systems, or experimental conditions, e.g., temperature, pH, ionic strength. (Crabtree & Shamma, 2018)

1.1.2.2 Continuous-flow techniques

The advantages of continuous-flow processes include safer handling and metering of hazardous substances, as well as better throughput and batch reproducibility. The utilization of multistep or telescoped flow speeds up the synthesis process and eliminates the need for interim storage. In a continuous-flow experiment, the reaction is performed under steady-state flow conditions as a function of the distance down-stream from the mixer, which can be beneficial for insensitive detection methods. Therefore, the continuous-flow technique can achieve faster end results than stopped-flow, yet it would lead to the use of more sample of protein (Roder et al., 2006).

1.1.3 Ubiquitin

Ubiquitin is a small regulatory protein that is widely found in almost all cells of eukaryotic organisms, Aaron Ciechanover, Avram Hershko, and Irwin Rose discovered the basic roles of ubiquitin and the components of the ubiquitylation pathway in the early 1980s, Aaron Ciechanover, they were awarded the Nobel Prize in Chemistry in 2004 for their discovery of ubiquitin-mediated proteolysis. The ubiquitin pathway has been proven to have a critical function in cellular metabolism and regulation during many years of research. (Hershko & Ciechanover, 1998).

The ubiquitin protein consists of 76 amino acids and has a molecular mass of 8.5 kDa. Ub contain seven lysines (K6, K11, K27, K29, K33, K48, and K63) that, function as attachment sites for polyubiquitin chain assembly, together with the N-terminus. These lysines are positioned in different directions on the surface of Ub. As a result, Ub chains with various lysine connections have distinct orientations.

The ubiquitin fold is distinguished by a complicated folding consisting of a five-stranded β sheet, an α helix, and a short helix. Ubiquitin is extremely soluble and thermostable ($T_M > 360$ K), with no complex compounds or disulfide linkages (Vijay-Kumar, Bugg, & Cook, 1987).

Ubiquitin is highly conserved among eukaryotic species: Human and yeast ubiquitin share 96% sequence identity; Ubiquitin is encoded by four distinct genes in mammals. The UBA52 and RPS27A genes, respectively, code for a single copy of ubiquitin fused to the ribosomal proteins L40 and S27a. Polyubiquitin precursor proteins are encoded by the UBB and UBC genes. (Goldstein et al., 1975) (Kimura & Tanaka, 2010).

(a)

```

sp|P62975|UBIQ_RABIT      MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG
sp|P68197|UBIQ_CERCA     MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG
sp|Q865C5|UBIQ_CAMDR     MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG
tr|A5JUZ1|A5JUZ1_MOUSE  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG
tr|A0A2K6NCN6|A0A2K6NCN6_RHIRO MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG
tr|Q71MT9|Q71MT9_SPOLT  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG

```

(b)

```

Ubiquitin sequence:      6      11      27 29 33
MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQ
LEDGRTLSDYNIQKESTLHLVLRGG
                        63

```

Figure 1-6. (a) Sequences aligned of ubiquitin of different species using *ClustalOmega*, (b) lysine residue are highlighted and numbered in red.

Ubiquitin acts as a versatile cellular signal that controls a wide range of biological processes, including protein degradation, DNA repair, endocytosis, autophagy, transcription, immunity, and inflammation (Hershko & Ciechanover, 1998). Much of this function is through the generation of polyUb chains, where links between Ub's C-terminus and one of the seven Lys residues on another Ub monomer are produced. Different linkages lead to different topology and function.

Ub is a very important protein and a structurally preserved protein that regulates most of the eukaryotic cell processes. It carries out its many tasks by conjugating to a wide range of target proteins. A wide range of alterations can occur. Ub and other members in the family class of ubiquitin (e.g., SUMO) (Pickart & Eddins, 2004) function by attachment of Ub's C-terminus to Lys residues on the target protein through an isopeptide bond. (Hicke, 1997). This can change the stability or localization of the target protein, or at least inhibit its activity. Because Ub contains seven lysines, it can link to other Ubs several times, allowing polyubiquitin chains to develop (Pickart & Eddins, 2004).

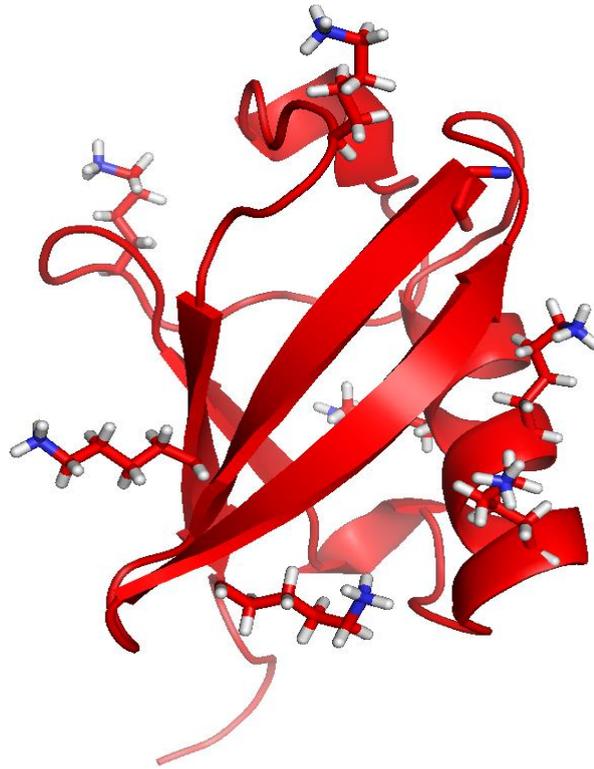


Figure 1-7. Ubiquitin pymol structure (1UBQ) showing the lysine residues that are able to form isopeptide bonds to other Ub monomers and generate polyUb chains.

1.1.4 Ubiquitin-binding domains (UBDs)

Ubiquitin-binding domains (UBDs) are proteins that interact with monoubiquitin and ubiquitin chains to achieve signalling in order to interpret ubiquitinated target signals into biochemical cascades in the cell.

The *Saccharomyces cerevisiae* Rad23 (UBDs) protein is involved in the regulation of nucleotide excision repair, proteolysis, and cell cycle progression in the proteasome system. Rad23 was originally identified as the primary degradation machinery in the cytosol and nucleus of eukaryotic cells. There are some similarities between the structure of Rad23 with the ubiquitin proteasome system domains (Chen, Shinde, Ortolan, & Madura, 2001). The structure of Rad23 consists of a N-terminal ubiquitin-like (UbL) domain and its connected to two ubiquitin-associated (UBA) domains: an internal UBA1 domain and a C-terminal UBA2 domain (Heessen, Masucci, & Dantuma, 2005).

The UBA domain was originally identified as a sequence motif present in proteins linked to the ubiquitination system. Chen et al. (2001) reported that, regardless of their low sequence homology, all UBA domains share the ability to bind ubiquitin. Based on its ability to simultaneously bind ubiquitinated proteins and the proteasome, it has been proposed that Rad23 may function as a platform that facilitates interactions between substrates and the proteasome. Recently, direct biochemical evidence has been provided for a role of Rad23 in targeting ubiquitinated substrates to the proteasome. (Chen et al., 2001)

1.2 Mass spectrometry

Mass spectrometry (MS) is an essential analytical tool in experimental chemistry and biochemistry. The applications of mass spectrometry cover a wide range from routine mass measurement to obtaining information on the structure of biological molecules and large-scale studies of entire proteomes. The field of biochemistry is undoubtedly one of the main beneficiaries of innovative MS (Kanu, Dwivedi, Tam, Matz, & Hill, 2008).

MS is based on the phenomenon of the motion of charged particles (ions) under the influence of magnetic or electric fields. This motion depends on the mass to charge ratio (m/z) of the ion.

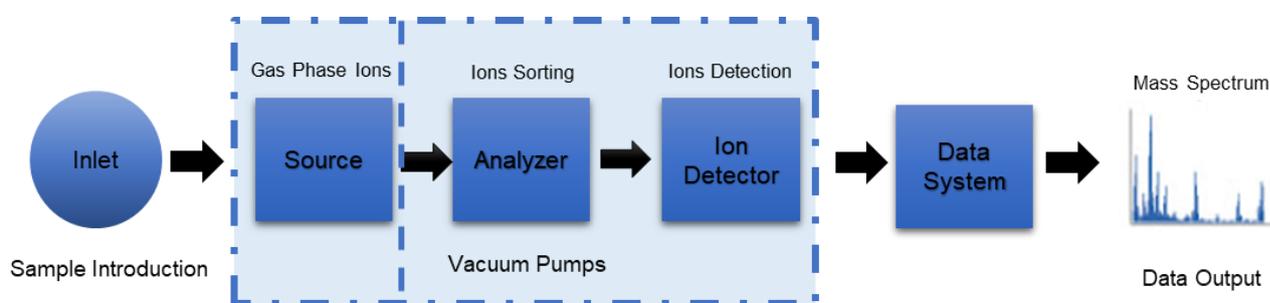


Figure 1-8. Block diagram that shows the basic parts of a mass spectrometer.

Figure 1-8 illustrates the elementary components of MS. First, the inlet is used to introduce samples into the vacuum region, if a vacuum ionization source is used (Banerjee & Mazumdar, 2012). Then, the sample molecules are ionized and accelerated to the mass analyser, which is located after the source region. The mass analyser separates ions according to their m/z ratio. Ions are then detected, and the signal is transferred to a data system. Mass spectrometers have a vacuum system to preserve the low pressure (10^{-3} to 10^{-10} mbar) to minimize ion-molecule interactions and maximize ion transmission. (Gross, 2006)

Recent advancements in MS technology can also be used to characterize secondary, tertiary, and quaternary structures of proteins, thus making it a fast and sensitive complementary method to nuclear magnetic resonance (NMR) and X-ray crystallography (Gross, 2006).

1.2.1 Methods of ionization

1.2.1.1 Electrospray ionization

Electrospray ionization (ESI) is a mass spectrometry technique for producing ions using an electrospray, which involves applying a high voltage to a liquid to create an aerosol. It is particularly useful for making ions from macromolecules since it avoids the molecules' tendency to fragment when ionized (Dole et al., 1968).

ESI starts with an analyte solution, which is injected (often by a mechanical syringe pump) through a capillary at a flow rate 5~1000 $\mu\text{l}/\text{min}$. High voltage (2-5 kV) is applied to the tip of the capillary relative to the MS inlet. This strong electric field causes the dispersion of the sample solution into an aerosol of highly charged electrospray droplets. (Gross, 2006)

Desolvated and partially desolvated ions pass through a sampling cone or the entrance of a heated capillary toward the analyser of the mass spectrometer (Figure 1-9). The heated sampling cone of the capillary causes the complete desolvation of the ions passing through it. Desolvation may be assisted by gas flow (N_2) or heat (50-300°C). The charged droplets are reduced in size by solvent evaporation supported by the flow of nitrogen gas. (Banerjee & Mazumdar, 2012)

Electrospray ionization works well for both small and large polar molecules and produces mostly multiply charged ions for large peptides and proteins. However, the greatest advantage of using ESI is in protein analysis, since proteins can be ionized without denaturation, for example non-covalent, receptor covalent and receptor-ligand complexes remain intact. The main benefits of ESI are its ability to integrate into chromatographic and electrophoretic separation methods, as it can generate multiple charged analyte ions, thus enabling instruments such quadrupole and ion trap to detect high molecular mass analytes. (Ashcroft, 2003)

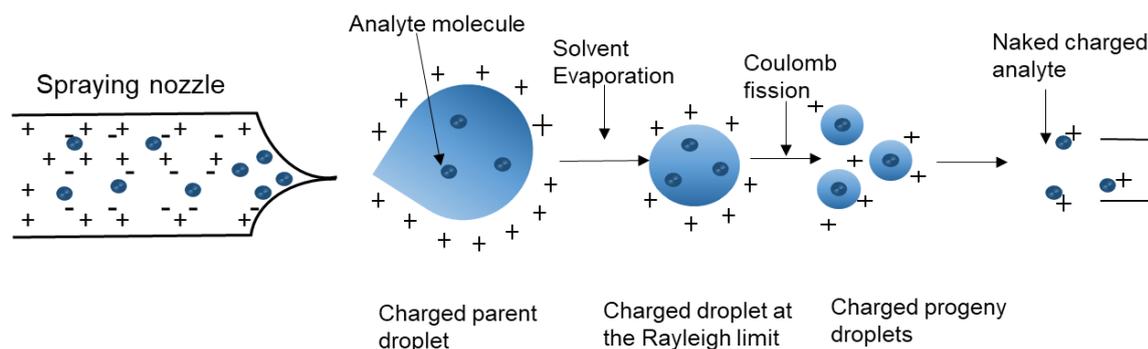


Figure 1-9. Schematic representation of the electrospray ionization process.

In ESI-MS analyses, there are several model of ionization that explain the final production of gas-phase ions: the ion evaporation model (IEM) and the charge residue model (CRM) (Kearle & Peschke, 2000).

In CRM a voltage applied to the tip of the capillary generates an aerosol of analyte-containing droplets; solvent evaporation leaves positive ions originating from the addition of one (MH^+) or more $[(M + nH)^{n+}]$ protons to the analyte, which are subsequently extracted into the high vacuum of the mass spectrometer analyser system, where they are then separated according to their m/z ratio. (Kearle & Peschke, 2000). The Charged Residue Mechanism has permitted quantitative predictions of protein charge state in the gas phase and is well supported for big proteins of varied mass. It is anticipated to be useful in the investigation of bigger supramolecular and polymeric systems as well (Santos, 2009).

The Ion Evaporation Model (IEM) is another mechanism used to explain the ESI process. Direct ion emission from the droplets is predicted to occur once the radii of the droplets have shrunk to a very tiny size (less than 10 nm) according to this model. Coulomb fission is now replaced by the ion evaporation mechanism (Santos, 2009).

In conclusion, the Ion Evaporation Model is well supported experimentally for low molecular mass ions seen in inorganic and organic chemistry. The Charged Residue Model (CRM)

becomes considerably more credible when the ions are large, such as polymers, dendrimers, or biological supramolecular complexes like proteins and enzymes (Santos, 2009).

1.2.1.2 Nano-electrospray ionization

Nano-electrospray ionization (nano-ESI) is likely to be the most efficient liquid sampling method for a direct protein analysis with MS. Nano-ESI is a type of electrospray with the same fundamental ionization process of droplet formation. (El-Faramawy, Siu, & Thomson, 2005)

One to two microliters of sample is deposited into a glass or quartz tube that has a tip diameter in the order of 1 μm , and is sprayed from the tip by applying a voltage to the solution. The actual flow rate is usually a few nL/min to a few tens of nL/min, controlled by the diameter of the tip, the voltage applied, and the backpressure which is sometimes applied to the tube content. Nano-ESI is typically used for peptide and protein analysis because of the ability to analyse a small volume of sample, and to make the sample last for many minutes so that various experiments can be performed. (El-Faramawy et al., 2005).

When nanospray was developed for electrostatic dispersion, the mass spectra of analytes exhibited substantially less dependency on background electrolytes (such as sodium salts) in experiments. Charged droplets form gas-phase ions only when the droplets are very small. This is true for both IEM and CRM. As a result, if one starts with small initial droplets, such as those formed by nanospray, substantially less solvent evaporation is necessary to reach the final droplet size required for the formation of gas-phase ions. As a result, in the presence of contaminants such as sodium salts, the salt concentration rise with nanospray will be significantly smaller. (Santos, 2009).

1.2.2 Mass analysers

The mass analyser is the most fundamental part of the mass spectrometer. In principle, the mass analyser separates ions by m/z , which are then passed to the detector. The proper selection of the mass analyser depends on the resolution, mass range, scan rate, and detection limit required for an application. (Cotter, Gardner, Ilchenko, & English, 2004)

The combination of ESI with several types of mass analysers provides a wide variety of specialized mass spectrometers. Six types of analysers are commonly used in proteomics: quadrupole (Q), ion trap (IT), time-of-flight (TOF), Fourier transform ion-cyclotron resonance FT-ICR, Orbitrap, and Magnetic sector mass analyser. Analysers are selected as a function of the analytical problems and their price. The choice of a mass spectrometer will strongly depend on the strategy preferred for protein identification and on the biological question. Once these are clearly defined, the key characteristics and performances of the instrument should be considered.

Ion trap and FT-ICR can be used in MS and MS/MS analysis, since the same analyser is used sequentially as MS1 and MS2. Q-TOF and TOF-TOF are hybrid instruments which can also perform MS/MS analysis.

1.2.2.1 Time-of-flight mass analyser

The Time-of-flight mass spectrometer (TOF MS) is one of the most common types of analysers used in MS because of its sensitivity, speed of analysis, resolving power, and high signal to-noise ratio. TOF is a mass spectrometry method that uses a time of flight measurement to calculate an ion's mass-to-charge ratio. An electric field of known strength accelerates ions (Wolff & Stephens, 1953).

TOF MS is a common method of peptide and protein characterization and can analyse either multiply charged or singly charged protein ions of high mass (i.e., greater than 200 kDa) in a

high throughput mode. In TOF, ions are accelerated by a voltage V , with the time required for the ion to cross a fixed distance in the flight tube being measured as a function of m/z . A pulse that separates packets of ions along the flight tube towards the detector is needed for the TOF to be recorded (Banerjee & Mazumdar, 2012). Frequently this is done by using an orthogonal pusher to convert a continuous ion beam into an interrupted one in the ESI-TOF spectrometer (source). TOF mass spectrometers operate on the simple principle that ions of differing m/z ratios, but equal kinetic energy, when projected into an electric field free region, will separate according to their m/z ratios (Hoaglund, Valentine, Sporleder, Reilly, & Clemmer, 1998).

1.2.2.2 Fourier transform ion cyclotron resonance (FT-ICR)

The basic principle of the FT-ICR is to measure ion cyclotronic frequency in a magnetic field, which allows ion mass to be calculated. For this, a pulsed radiofrequency signal is used to excite the ions while they are orbiting. (Marshall, Hendrickson, & Jackson, 1998)

Excited ions generate signals that are processed by a Fourier transform (FT) to obtain the component frequency of the different ions, which correspond to their m/z ratio. Because ion frequency can be measured with high accuracy, their corresponding m/z ratio is also calculated with high accuracy. (Marshall et al., 1998). One major drawback of these instruments is their high cost, which is partly due to the superconducting magnet required to trap ions and induce ion circular motion. However, FT-ICR instruments have the highest resolution capabilities. (Schaeffer-Reiss, 2008)

1.2.3 Tandem mass spectrometry

Tandem mass spectrometry, commonly known as MS/MS or MS², is a technique in instrumental analysis in which two or more mass analysers are linked together to expand their ability to examine chemical samples by adding an additional reaction step. The investigation of biomolecules such as proteins and peptides is a common use of tandem MS (de Hoffmann, 1996).

It is consequently necessary to create analytical instruments and procedures with sufficient sensitivity, selectivity, and accuracy to identify and quantify them PTMs. Several articles detailed the good results obtained by MS and tandem MS (MS/MS) investigations, demonstrating that PTMs may be assigned to specific sites in proteins even at the resolution of individual amino acids (Jensen, 2004). Most common type of MS/MS ion activation methods are collision-induced dissociation (CID), electron transfer dissociation (ETD), and electron capture dissociation (ECD) (Elviri, 2012).

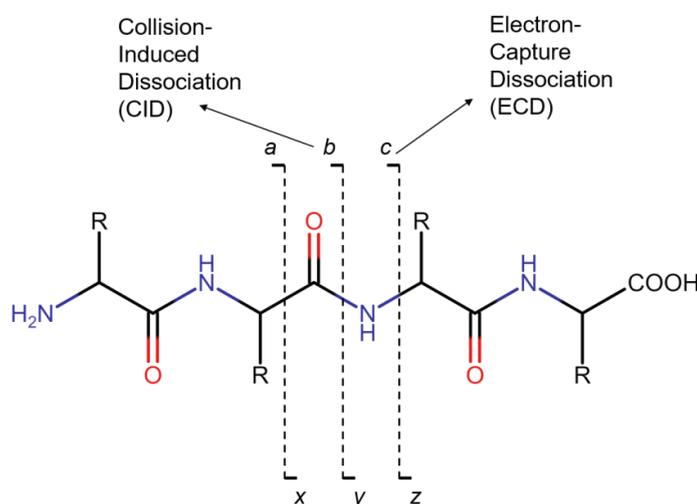


Figure 1-10 Peptide backbone fragmentation: peptide structure, selected tandem mass spectrometric (MS/MS) techniques, Fragment ions retaining the positive charge on the amino terminus are termed a-, b-, or c-type ions. Fragment ions retaining the positive charge on the carboxy terminus are termed x-, y-, or z-type ions.

1.2.3.1 Collision induced dissociation (CID)

Collision-induced dissociation (CID) is one the most common ion activation methods used in present day MS instruments. In CID, activation of the selected ions occurs by collisions with neutral gas molecules in a collision cell.

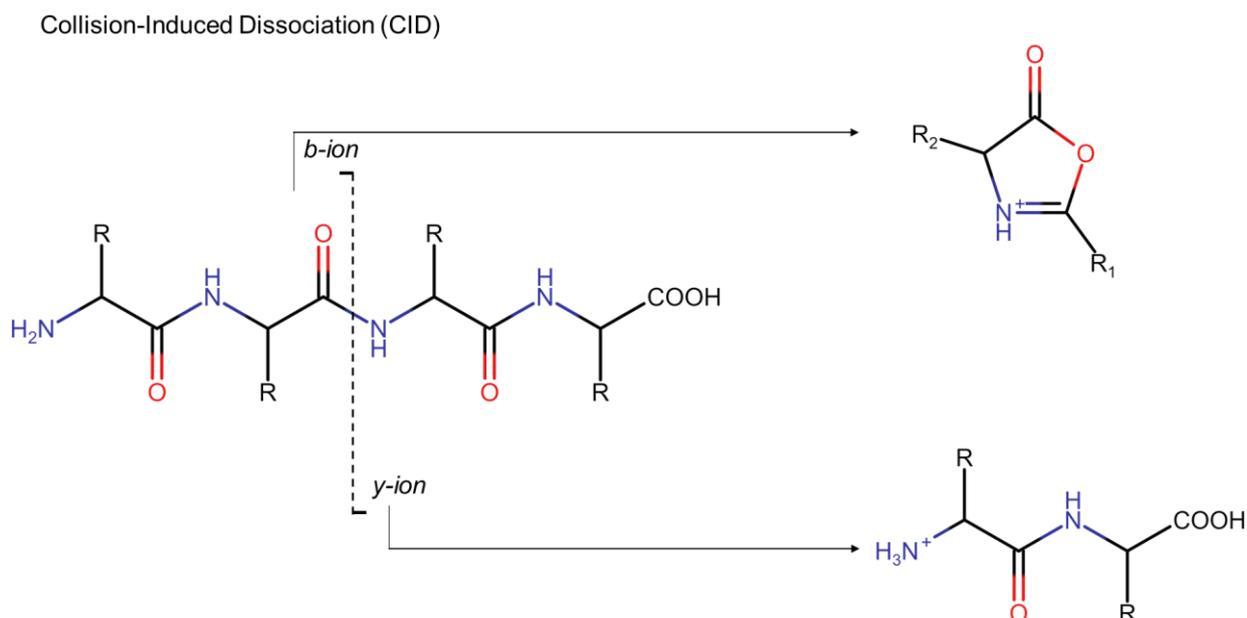


Figure 1-11 Peptide backbone fragmentation: CID Fragment ions retaining the positive charge on the amino terminus are termed *b-ion*, Fragment ions retaining the positive charge on the carboxy terminus are termed *y-ion* type ions.

Collisions between the precursor ion and a neutral target gas are accompanied by an increase in internal energy, which induces dissociation to a number of fragments. When precursor ion with a high kinetic energy undergoes collision with a neutral gas (helium, nitrogen or argon), part of the precursor ion kinetic energy is converted into internal energy, which leads to the breakdown of the bond and the fragmentation of the molecular ion into smaller fragments called product ions. These fragment ions can then be analysed by tandem mass spectrometry. Tandem mass spectrometry activates precursor ion formed in the ion source. The ion activation method defines what types of products result. (Sleno & Volmer, 2004)

The masses are then used to identify the peptides by their fingerprint. The fragmentation of the peptide is a calculated procedure controlled by the physical and chemical properties of the peptides and amino acid sequence. Peptides normally break at the peptide bond when fragmented, the resulted spectrum shows the information of the amino acids component of the peptide. (Bafna & Edwards, 2003)

Techniques such as CID can be used to fragment molecular ions and proteins to produce specific fragments or peptides in MS/MS measurement. Specific fragments are important for the identification of amino acid or protein structures. This will help in the interpretation of peptide to identify the protein sequence. (de Hoffmann, 1996)

1.2.3.2 Electron-capture dissociation (ECD)

Classical peptide sequencing by “tandem” mass spectrometry mainly uses automated instruments including Q-TOF, IT and OT, TOF-TOF, and seldom FT-ICR. MS/MS instruments offer additional possibilities and give access to sophisticated experiments for the characterization of peptide families (phosphopeptides, peptide glycosylation, etc.). To improve peptide sequencing, fragmentation techniques alternative to classical CID have been developed, for example electron capture dissociation (ECD). In MS/MS, ECD is one of the most extensively used procedures for mass chosen precursor ion activation and dissociation of multiply-charged ions such as those produced by peptides or intact proteins. It entails the application of low-energy electrons directly to trapped gas-phase ions (Elviri, 2012).

The advantage of ECD is to generate fragments that are evenly distributed along the peptide backbone. In contrast, CID-induced fragments are usually restricted to a more limited number of cleavage points in the peptide and, therefore, yield less sequence information. This is a major advantage for the study of PTMs. Indeed, the combination of CID and ECD fragmentation

methods can be used, for example, to localize PTM on the peptide backbone. (Schaeffer-Reiss, 2008)

ECD fragments precursor ions by adding free electron to multiply protonated molecule M^{n+} to form an odd-electron ion. Releasing of the electric potential energy results in fragmentation of the product ion. (Mirgorodskaya, Roepstorff, & Zubarev, 1999). The odd-electron species experience reconfiguration, followed by N-C α backbone cleavage $[M + nH]^{n+} + e^- \rightarrow [M + nH]^{(n-1)+} \rightarrow \text{fragment}$ (Biemann, 1990).

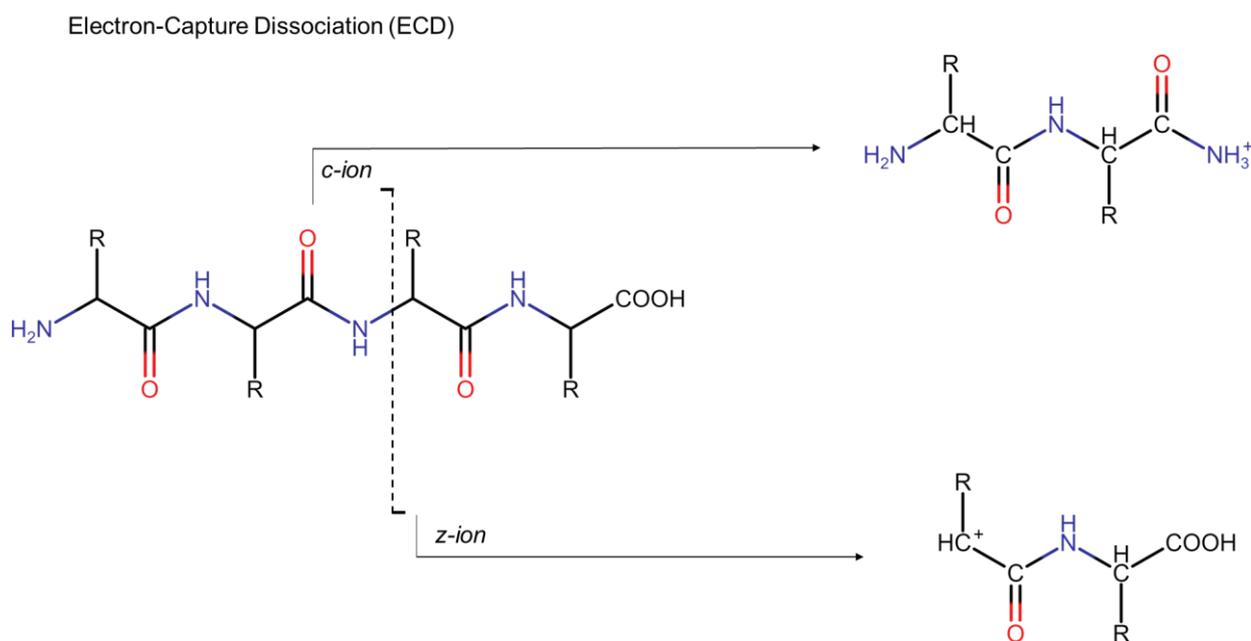


Figure 1-12. Peptide backbone fragmentation: ECD Fragment ions retaining the positive charge on the amino terminus are termed c-ion, Fragment ions retaining the positive charge on the carboxy terminus are termed z-ion type ions.

1.3 Ion mobility mass spectrometry

Many protein MS measurements are relatively straightforward experiments aimed at determining protein mass in the hypothesis that a protein is distinct by its covalent structure (amino acid sequence and post-translational modifications, PTMs). Nevertheless, covalent structure alone does not provide a comprehensive characterization of the protein molecules since the biological activity of proteins is most frequently determined by their higher structure.

MS plays a significant role in the characterization and identification of protein sequence and structure, and its effect continues to grow. Ion mobility (IM) spectrometry coupled with mass spectrometry (MS) provides a method for detailed investigation for the structural and conformational properties of bio-molecules in the gas phase, particularly intact large biomolecules such as proteins. (Lanucara, Holman, Gray, & Eyers, 2014)

1.3.1 Ion mobility

Ion mobility spectrometry separation depends on the mobility differences of a packet ions as they drift through an inert gas under the influence of a weak electric field. The first ion mobility devices or drift tubes that have been used employed a uniform, static, electric field to drive ions through the background gas. The physical principles behind drift tube devices are well researched. The mobility values, or drift times, obtained can be used to derive gas-phase collision cross-sections (Ω) for comparison with theory (McDaniel & Mason, 1973).

Pioneering research (Konstantin B Shelimov, Clemmer, Hudgins, & Jarrold, 1997) showed that IMS can be used in probing protein structure. Its potential for interrogating the structure of proteins and multiprotein complexes has only recently been utilized. Ion mobility– mass spectrometry is often applied to the structural elucidation of protein assemblies that have failed high-throughput crystallization or NMR spectroscopy screens (Zhong, Hyung, & Ruotolo, 2012).

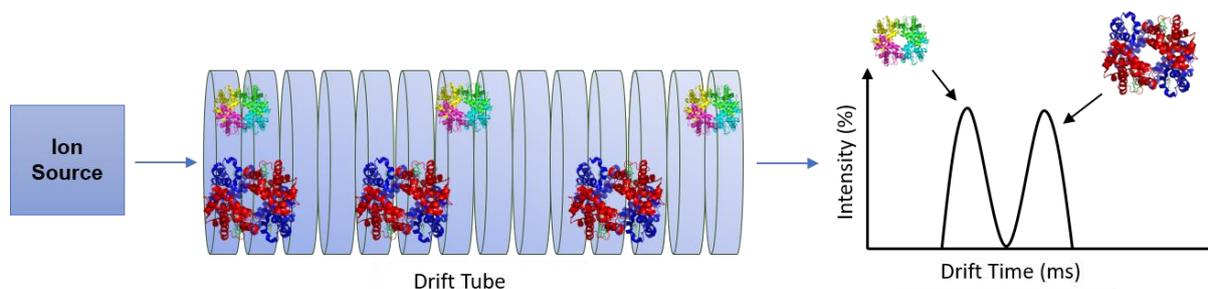


Figure 1-13. Fundamental separation principles and data acquisition schematic of ion mobility-mass spectrometry. Ions are generated at the ion source, then induced to drift in an ion guide filled with neutral gas molecules under the effect of an electric field. The ions travel through this region based on their size-to-charge ratio. Then they reach the time-of-flight (ToF) mass analyser via a vacuum to be analysed based on their mass-to-charge (m/z).

Early work on ESI-MS of proteins showed a close relationship between the protein charge state distribution in the mass spectrum and the ESI solution conditions, with denaturing conditions yielding high charge states and native conditions lower charge states (Chowdhury, Katta, & Chait, 1990; Lars Konermann & Douglas, 1998; Valentine, Anderson, Ellington, & Clemmer, 1997).

Additional extensive work by Clemmer's group on IMS-MS of several proteins, such as cytochrome *c*, lysozyme, and apomyoglobin, which discovered well-defined structures from compact to extended and some in between. (Badman, Myung, & Clemmer, 2002)

Clemmer's research has made huge advancement in the field of Ion mobility in the past two decades, especially on protein folding/unfolding in the gas phase, highly focused on understanding the nature of protein conformation in the gas phase using a combination of physical (collision cross section measurements) and chemical probes (H/D exchange and proton transfer reactions), results observed for conformational changes with charge state for cytochrome *c*, As charge state increase the coulombic repulsion energy increased as well. For high charge states, an insufficient number of intramolecular stabilizing interactions are possible to maintain a compact conformation (Valentine et al., 1997).

Collisional activation of the ions can result in the initiation of the conversion of protein from a compact conformation into a partially unfolded or elongated structure. Lower activation energy

converts a compact into a partially folded structure, while a stronger activation energy unfolds the protein to a series of well-resolved extended conformations, a result that is similar to what happens as the charge state increases (Merenbloom, Koeniger, Valentine, Plasencia, & Clemmer, 2006).

These results indicate that the unfolding of protein in IMS occurs in stages of low well-defined intermediates of protein conformations, however the final structure obtained depends on the starting structure of initial compact conformations. Whereas native solution conditions yield low charge states and denaturing conditions yield high charge states (Shi & Clemmer, 2014).

1.3.2 Traveling wave ion mobility spectrometry

Ion mobility (IM), separation is used extensively to assess the structure and identity of molecular analytes, ranging from atomic species to multiprotein complexes (B. T. Ruotolo et al., 2005). The combination of this separation technology with mass spectrometry (IM-MS) has allowed more detailed and multidimensional analysis of both complex mixtures and biomolecules.

IM-MS is capable of rapidly measuring changes in protein structure, oligomeric state, and binding stoichiometry from complex mixtures at relatively low concentrations. Since the introduction of the Synapt platform (the Synapt HDMS system-Waters Corp., Milford, USA), the majority of IM-MS datasets for multi-protein complexes have been generated on IM-MS instruments. A schematic diagram of the hybrid quadrupole/IM separator/Q-ToF instrument is shown in Figure 1-14.

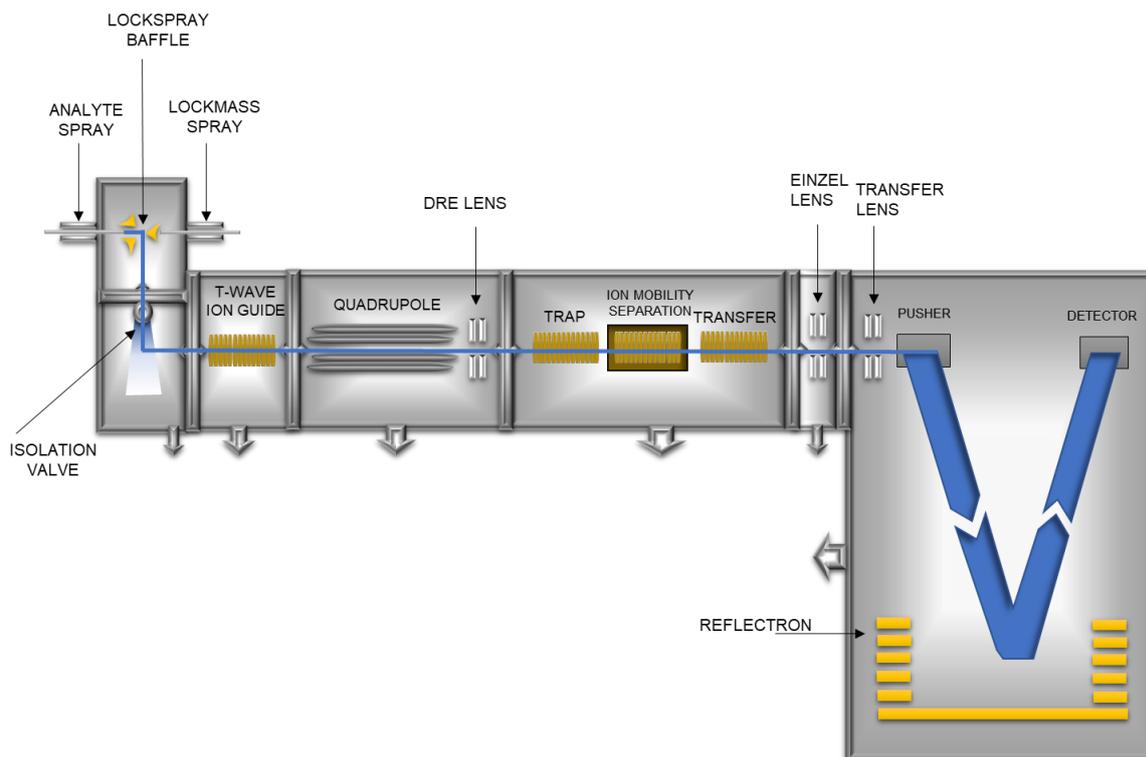


Figure 1-14. Schematic diagram of the Waters Synapt HDMS (Waters, Altrincham, UK) system.

The trap and transfer compartments operate at pressures similar to that of a standard collision cell, thus, CID fragmentation can operate in these regions. The advantage of this system is that the three operation modes can be used; i) fragmentation of mobility separated precursor ions, ii) mobility separation of product ions and iii) fragmentation of mobility separated product ions. Another advantage is the presence of a quadrupole mass analyser which is positioned before the trap region which makes the analysis perform on the entire range of species present or on ions at a pre-selected m/z of interest. This range of operation modes makes the Synapt instrument ideal for the study of conformer rearrangement following collisional activation on the millisecond timescale. (Garcia, Giles, Bateman, & Gaskell, 2008)

1.3.3 Collision cross section calculation

In IMS, ions move through a drift tube containing a buffer gas. The buffer gas delays the ion movement, while the acceleration of the ions happens in an electric field. Thus, ions travel at a terminal velocity proportional to the inverse of their collision cross section (CCS), a parameter

that describes the average of collision rate. The CCS is smaller for more compact (more spherical) molecules than for more extended ones (e.g., planar structures, extended chains, helices, etc.) and thus yields insight into the overall shape. (Ewing, Glover, & Clemmer, 2016)

IM separates protein ions based on shape, charge and mass ratio under the effect of an electric field. Larger protein ions experience a greater number of collisions with the inert gas that is in the chamber, and consequently have a larger CCS values, than more compact protein ions of identical m/z . (Zhong et al., 2012).

Furthermore, ion mobility of a gas phase ion is subject to its collision cross section and charge state. Highly charged ions will experience a greater drift force than singly charged ions, therefore having drift times. For ions of the same charge state and similar mass compact ion conformations will have shorter drift times than elongated conformations. Collision cross sections calculation can be informative in structural determinations. (T. Wytttenbach, N. A. Pierson, D. E. Clemmer, & M. T. Bowers, 2014)

To determine the experimental CCS of an ion, the mobility must be determined:

$$K = \frac{v}{E} \quad 1.1$$

where K is the proportion between the speed an ion moves, v , and the electric field to which it is subjected, E (Mason & McDaniel, 1988). Because the mobility depends on the temperature and pressure of the measurement, a reduced mobility, K_0 which has been normalized to standard temperature and pressure, is typically reported.

In low electric fields, the Nernst–Townsend–Einstein relation, relates the mobility and the diffusion coefficient (Revercomb & Mason, 1975). The diffusion coefficient or mobility can be used to derive a CCS for an ion, the net effect for a linear drift tube is Eq. (2), the Mason-Schamp equation.

$$\Omega = \frac{ze}{16} \left(\frac{18\pi}{k_b T} \right)^{1/2} \left(\frac{1}{m_1} + \frac{1}{m_B} \right)^{1/2} \frac{tE}{L} \frac{1}{N} \quad 1.2$$

where L is the length of the drift tube, t is the drift time, k_b is the Boltzmann constant, T is the temperature, m_1 is the mass of the ion, m_B is the mass of the buffer gas molecules, and N is the number density of the buffer gas at the temperature and pressure of the experiment (Revercomb & Mason, 1975). However, for TWIMS calibration against proteins of known CCS value (Scarff, Thalassinou, Hilton, & Scrivens, 2008), such as myoglobin, ubiquitin or bradykinin is needed because the electric field E in a TWIMS cell is complex in nature and cannot be directly substituted into equation 1.2.

By measuring the changes in analyte collisional cross section CCS, valuable information such as conformational dynamics (Jenner et al., 2011) folding and unfolding intermediates (Shi, Pierson, Valentine, & Clemmer, 2012) ligand-induced stability/conformation changes (Hyung, Robinson, & Ruotolo, 2009) aggregation intermediates and protein topologies can be obtained (S. Niu & Ruotolo, 2015).

Work done by the Oldham group (Hopper & Oldham, 2009) showed that collisional cross section CCS measurements were found to be comparable with those predicted for folded protein structures in the absence of activation. Partially unfolded conformations were formed under collisional activation energies that were determined. The stability of noncovalent intra- and intermolecular interactions within protein complexes is related to the degree of unfolding and dissociation generated by these specific collision energy. These findings show that ligand binding increases the conformational stability of protein ions in the gas phase.

1.3.4 Collision Induced Unfolding

Collision Induced Unfolding (CIU) can be used to investigate the stability and conformational difference of protein ions, CIU methods have been used broadly in ligand binding identification. Ions are activated in the ion trap region by collision with argon gas in Waters Synapt G1 HDMS (Waters, Altrincham, UK) and Nitrogen gas in Waters Synapt G2, the resulting ion conformations then get separated in the IMS drift cell and detected by ToF MS.

Through adjusting the trap collision voltage, the kinetic energies of protein ions entering the trap region can be controlled and regulated. Activation profiles are then collected, and the collision cross section relatively calculated for each energy.

CIU experiments were described earlier for small monomeric proteins ions like apomyoglobin (Konstantin B. Shelimov & Jarrold, 1997) CIU was performed by Bowers on polyatomic ions in the absence of a solvent such as a dinucleotide, a sodiated polyethylene glycol chain, the peptide bradykinin, the protein ubiquitin, and two types of peptide oligomers. (Thomas Wyttenbach, Nicholas A. Pierson, David E. Clemmer, & Michael T. Bowers, 2014).

Earlier research employed CIU on small proteins but this quickly expanded to include more detailed instrumentation and applications covering large multi-protein complexes like intact antibodies (Tian, Han, Buckner, & Ruotolo, 2015) and the effect of ligand binding on conformational stability (Hopper & Oldham, 2009).

By introducing traveling wave mobility separators to IMS-MS instrumentation, enhances the mobility separation without compromising MS sensitivity (Hopper & Oldham, 2009) through applying Direct Current (DC) voltages as a series of high-speed traveling waves in combination with Radio Frequency (RF) radial confinement. Utilizing such method showed that the

conformations of large multiprotein complexes can be preserved upon transition to the gas phase (Pringle et al., 2007).

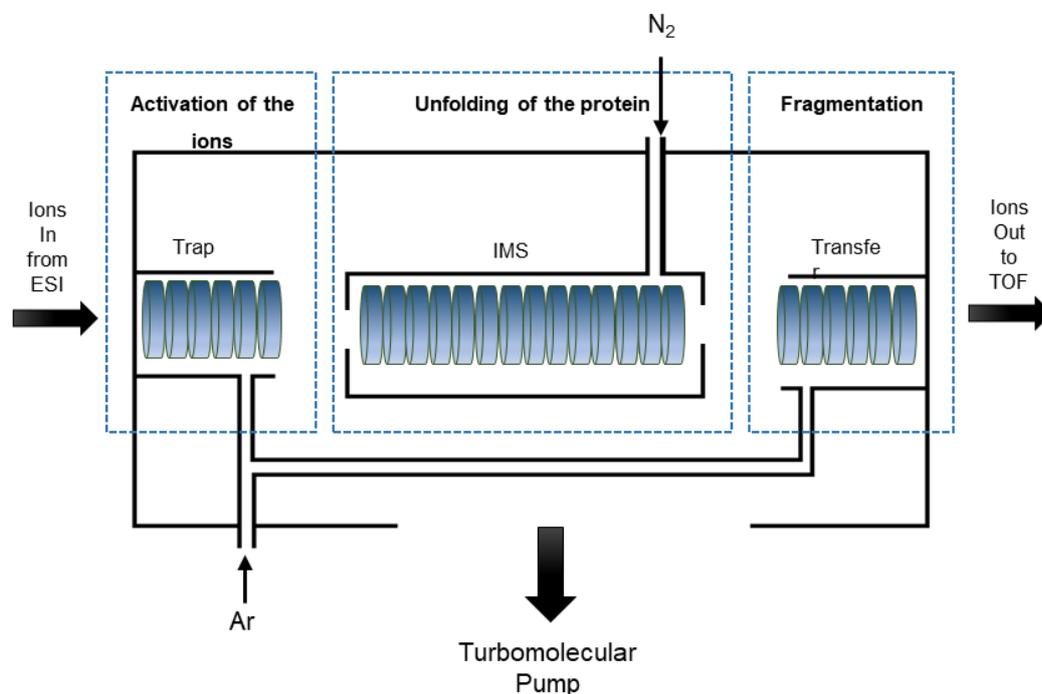


Figure 1-15 A schematic diagram of the IMS section of the Waters Synapt HDMS (Waters, Altrincham, UK), comprising three travelling wave ion guides labelled, trap, IMS and transfer.

The IM section comprises three travelling wave-enabled stacked ring ion guides as shown in more detail in Figure 1-5. The trap ion guide is used to accumulate ions during the previous mobility separation then release an ion packet into the IM ion guide for mobility separation. The transfer ion guide is used to convey the mobility separated ions to the ToF for mass analysis.

Protein structure and function are closely related, and conformational changes often carry changes in the activity of a protein. IM-MS is being used increasingly to supplement NMR and crystallography to provide a dynamic view of the structural changes caused by protein-ligand binding.

As previously mentioned, detailed CIU measurements have also been applied in the studies of protein–ligand stabilities. In a study by Oldham, ESI-IM-MS was used to measure the structural

stability of natively compact protein ions (FK-binding protein, hen egg-white lysozyme, and horse heart myoglobin) as a function of small molecule binding (Hopper & Oldham, 2009) The results show clear shifts in the CIU stabilities of ligand bound complexes relative to apo protein.

There are many variables that affect the gas-phase structure of desolvated protein ions. For example, studies by Clemmer's showed that altering the internal temperature of the ions caused dramatic influence on the size of the protein ion recorded by IM. Mainly leading to a positive correlation between protein ion CCS and their internal temperature, with protein ions of high internal temperatures adopting large, string-like conformational states (Clemmer & Jarrold, 1997).

There are many beneficial uses of CIU for differentiating isomers and biosimilars . Studies by Ruotolo's group provided a comprehensive investigation utilizing CIU fingerprinting and CID stabilities for studying protein–ligand complexes. The data indicate separate ligand stabilization mechanisms for wilde type and L55P TTR, a disease-associated mutant with enhanced amyloidogenic properties (Hyung et al., 2009).

In addition, (CIU) was used to show the unfolding pathway of the protein and followed in detail to generate additional points of comparison between either apo-states or alternate conformational families of the protein, since many possible tertiary structures project identical ion CCS values (Shuai Niu, Rabuck, & Ruotolo, 2013).

1.4 Protein structure studies by ion mobility mass spectrometry

These methods have been used in identifying the protein structure. Its potential for interrogating the structure of proteins and multiprotein complexes has only recently been utilized. (Hopper & Oldham, 2009).

Several researches have investigated collision induced unfolding where a single subunit within the assembly unfolds, inhabiting a number of intermediate structures that are relatively stable on the timescale of the IM measurement (Hyung et al., 2009; Brandon T. Ruotolo, Benesch, Sandercock, Hyung, & Robinson, 2008).

To provide a general overview, ion mobility application ranges from structural identification of a large molecule, atomic species (Bowers, Kemper, von Helden, & van Koppen, 1993) to multiprotein complexes (B. T. Ruotolo et al., 2005).

The combination of ion mobility (IM) separation with mass spectrometry is becoming powerful structural identification method, capable of measuring the structure, topology, dynamics and composition of large protein assemblies within complex mixtures (Linjie Han, Hyung, Mayers, & Ruotolo, 2011).

Recent advancements have seen huge developments in both IM-MS and the IM and MS technologies. (Merenbloom et al., 2006) (Kemper, Dupuis, & Bowers, 2009). Using ion mobility Sobott group found that the more extended states are more prone to fragment release. Products of gas-phase collapse are, however, less stabilized towards unfolding than the native conformation, indicating that the ions retain a memory of previous conformational states. (Frederik Lermyte, Łacki, Valkenborg, Gambin, & Sobott, 2017) Furthermore, this collapse of charge-reduced ions is promoted if the ions are 'preheated' using collisional activation, with

possible implications for the kinetics of gas-phase compaction. (F. Lermyte, Lacki, Valkenborg, Gambin, & Sobott, 2017)

Recently, new application of ion mobility has emerged for the analysis of the intact antibodies to distinguish between antibody isoforms involving the collision induced unfolding (CIU), where collisional heating in the gas-phase is used to generate unfolded conformers, which are then separated by IM and then analysed by MS. (S. Niu & Ruotolo, 2015; Tian et al., 2015; Tian & T. Ruotolo, 2017)

CIU has been used in the research of protein stability shifts upon anion or cation binding (Linjie Han et al., 2011) tertiary structure and differences in kinase–inhibitor complexes.(Shuai Niu et al., 2013)

In summary, CIU has many applications in structural biology, measurements to capture cooperative increases in protein stability as a function of ligand attachment in a multiprotein–ligand binding system (S. Niu & Ruotolo, 2015).(L. Han & Ruotolo, 2015; Zhong et al., 2012)

1.5 Chemical modification of proteins

Protein modification is an area of huge interest for many scientific field, especially mass spectrometry analysis. Specific chemical modification of the protein can change the protein structure and function, intracellular and extracellular function of the protein varied based on these modifications (Drahl, Cravatt, & Sorensen, 2005).

Naturally occurring modifications in the cell are commonly referred to as posttranslational modifications (PTMs), since they occur after the protein biosynthesis (the translation step). PTMs and proteins involved in mediating their incorporation into target proteins add additional layers to functional properties and diversities. In nature the covalent modification exists to modulate the proteins function. Subsequently these methods were incorporate on study of protein structure and functions (Pucheault, 2008).

Techniques that enable artificial chemical modification of proteins has been widely studied. These modifications include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation. This allowed the understanding of molecular mechanisms, structure and function of protein (Baslé, Joubert, & Pucheault, 2010).

1.5.1 Acetylation

Of the 20 ribosomal coded amino acid residues, lysine is the most frequently post-translationally modified, which has important functional and regulatory consequences (Zhang et al., 2011). Natural acetylation of proteins normally happens on the lysine residues or the N-terminus with the presence of acetyl-coenzyme A in the reaction as the acetyl group donor. For example, in histone acetylation, this protein acetylated and deacetylated on lysine residues in the N-terminal tail after the gene regulation. Naturally, acetylated and deacetylated reactions are catalysed by enzymes with histone acetyltransferase (HAT) or histone deacetylase (HDAC)

activity, HATs and HDACs can contribute in the reaction of the acetylation on non-histone proteins as well (Sadoul, Boyault, Pabion, & Khochbin, 2008).

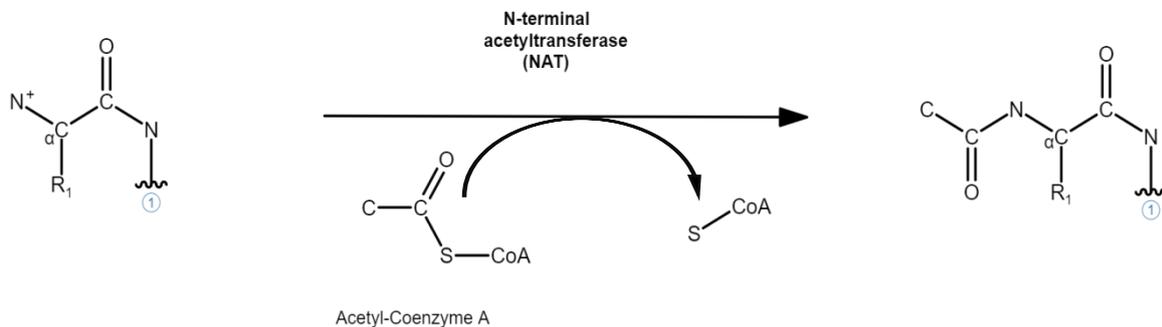


Figure 1-16. An NAT acetylates the N-terminal α -amino group of a protein or polypeptide by transferring an acetyl moiety (Ac) from Ac-CoA. The acetylation removes the positive charge of the N-terminal amino group, thus changing the chemical properties of the protein N terminus.

N-terminal or lysine acetylation modification can be easily identified on the post-translational modification sites due to the 42.01 Da mass shift, resulting from the replacement of one of the hydrogens from the amine group with $COCH_3$ on the modified amino acid site. Lysine is a widely distributed amino acid in almost all proteins in living organism and it is commonly found at the protein surface. Modification of lysine residues is one of the most common method used as covalent labelling strategy to probe protein surface structure.

Early researches done by Przybylski and co-workers showed that MS can be exclusively use to determine specific lysine modification sites, by using acetic anhydride to probe the surface topology of hen-egg white lysozyme (Suckau, Mak, & Przybylski, 1992). This research showed that the reactivity of the lysine residues correlates well with the calculated solvent accessibilities from crystal structure data, by using ^{252}Cf plasma desorption (PD) and MS peptide mapping methods.

Bioactivity assays coupled with modification results suggested that the N-terminal amino acid plays a role in the protein's function. Later researchers used anhydride acetylation/trideuteroacetylation for the acetylation and succinylation of amino groups to

characterize the surface topology of model proteins. These early results by Przybylski and co-workers were very significant in founding the building block in the method of covalent labelling when combined with MS in modelling protein structure (Suckau et al., 1992).

The primary effect of reversible lysine acetylation is the neutralization of its positive charge and consequently the modulation of the biochemical properties of lysine-containing proteins. Additionally, since lysine residues also provide specific sites for numerous other post-translational modifications, their acetylation would chemically lock the residue. The specific lysine locking activity of acetylation therefore confers to this modification a great regulatory potential with the ability to interfere with cellular functions relying on other lysine modifications, i.e., methylation, ubiquitination, sumoylation, neddylation, biotinylation. Lysine acetylation also either creates docking sites favouring protein-protein interactions or conversely interferes with the binding of specific partners (Sadoul et al., 2008).

1.5.2 Carboethoxylation of Histidine

Histidine is a relatively low-abundance amino acid residue in proteins. Nevertheless, it is an important residue due to its aromatic nature, moderate basicity, H-bonding capacity, and ability to bind divalent transition metals. Studies related to the histidine modification have shown great information about the mechanism of proteins (Díaz, de Castro, & Medina, 1993).

Findings show that histidine's reactivity is controlled by its solvent accessibility and protonation state (Mendoza & Vachet, 2009). Diethylpyrocarbonate (DEPC) is generally used to modify histidine residues. A single modification can be seen at low DEPC concentrations, but higher concentrations of the reagent can lead to a second carboethoxylation of histidine. Studies have shown that reactions of DEPC with lysine, tyrosine, cysteine, arginine, serine, and threonine residues can be common when increasing the pH value and the amount of lysine

modification increases over the pH range of 6-8 and leads to second modification. (Mendoza & Vachet, 2009)

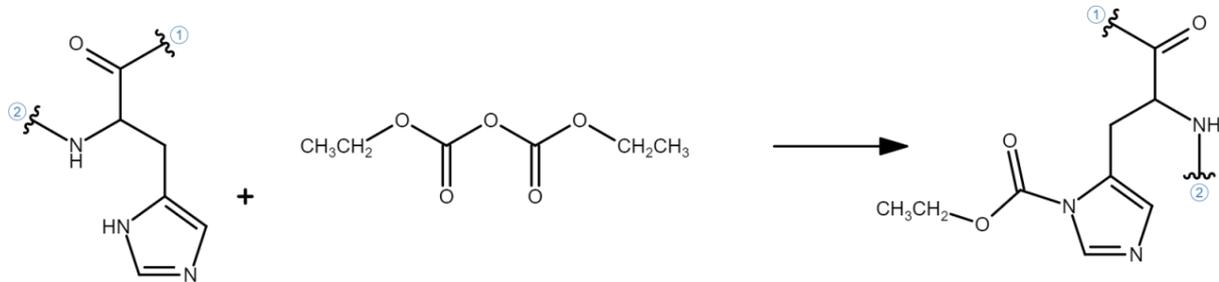


Figure 1-17. Modification reactions of histidine using DEPC, the addition of 72 Da to the protein.

1.5.3 Succinylation

Succinylation is a posttranslational modification where a succinyl group ($-\text{CO}-\text{CH}_2-\text{CH}_2-\text{CO}_2\text{H}$) is added to a lysine residue of a protein molecule. This modification can be seen in many proteins, including histones, and is a type of phosphorylation. The molecular weight of the peptide and its fragments, which may be determined by mass spectrometry, is used in this approach of mapping a PTM site to determine the location of the site.

The addition of a succinyl group changes the charge of lysine from +1 to -1 (at physiological pH) and introduces a structural moiety that is relatively substantial (100.0186 Da), much larger than that introduced by acetylation (42.01 Da). As a result, it is anticipated to result in more profound alterations in protein structure and function. (Zhang et al., 2011).

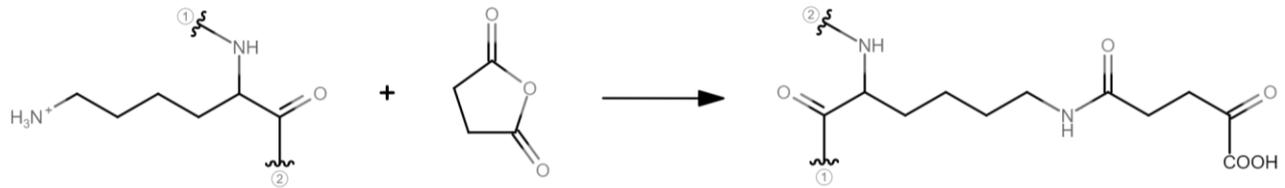


Figure 1-18. Succinylation reaction of lysine residue. Succinic anhydride reacts with lysine amino group of a protein or polypeptide by transferring succinyl group, since a positively charged -NH_3^+ from lysine is replaced by -COO^- of succinate half-amide.

A protein's chemical characteristics are significantly altered by lysine succinylation, as opposed to the chemical properties altered by lysine methylation or acetylation, two other PTMs that are known to have key physiological roles. A lysine residue will have its charge status changed from +1 to -1, from +1 to 0, and not at all when succinylation, acetylation, and monomethylation occur at physiological pH (7.4). As a result, the change in charge status at the lysine residue induced by succinylation is comparable to the change in charge status at the serine, threonine, or tyrosine residues caused by protein phosphorylation (from 0 to 2 charges). Additionally, succinylation contributes a larger structural moiety than either acetylation or methylation. Because of the more extreme structural mutation caused by lysine succinylation, as proven in the mutagenesis experiment, it is likely that the protein's structure and function will be altered to a greater extent. (Zhang et al., 2011)

1.6 Molecular dynamics simulation

The polypeptide chain that makes up a protein molecule has a great variety of different conformations. The native conformation of the protein spends the majority of its time in a microscopic fraction of the whole configuration space. As a result, a protein's amino acid sequence must meet two requirements, one thermodynamic and kinetic level (Martin Karplus & Šali, 1995). The sequence must have a unique folded conformation that is stable under physiological settings, according to the thermodynamic criterion. The denatured polypeptide chain must fold into this shape with reasonable speed to meet the kinetic requirement. The amount of interest in protein folding has risen dramatically in recent years.

Structural information concerning the folding and unfolding reaction comes from stopped-flow kinetics and NMR experiments (Crabtree & Shamma, 2018; Roder et al., 2006), mass spectrometry (Miranker, Robinson, Radford, Aplin, & Dobson, 1993), and mutation studies (Fersht, 1993). The majority of the kinetic data has been confined to milliseconds or longer time intervals. Therefore, new ways for fast activating folding reactions are vital for collecting data on critical events that occur during the first millisecond (Jones et al., 1993). Many people are contributing to the theory of protein folding, including physicists and mathematicians, as well as chemists and biologists. This opened a new pathway of theoretical techniques such as Molecular dynamics simulations (MD). The study of unfolding from the native state can be done using molecular dynamics simulations with an atomic model and explicit representation of the solvent.

Molecular dynamics simulations are excellent tools for exploring the conformational energy landscape that these molecules can access, and the rapid rise in computer capacity, combined with advancements in methodology, makes this an exciting time for simulation in structural biology. Molecular dynamics (MD) is a computer simulation method for studying atoms and molecules' physical motions. The atoms and molecules are allowed to interact for a set amount

of time, providing a perspective of the system's dynamic "evolution." The trajectories of atoms and molecules are determined in the most common version by numerically solving Newton's equations of motion for a system of interacting particles, with forces between the particles and their potential energies calculated using interatomic potentials or molecular mechanics force fields. Chemical physics, materials science, and biophysics are some of the fields where the approach is used (Schlick, 1996).

Molecular dynamics simulations allow atomic-level information to be observed in the conformational motion of a molecule such as a protein over time. Prior to launching a simulation, several decisions must be taken, including which software to use, which molecules to include in the simulation, and which force field to employ to explain their behaviour (M. Karplus & Kuriyan, 2005).

The first molecular dynamics simulation of a protein was reported by McCammon and consisted of a 9.2-ps trajectory for a small protein in vacuum (McCammon, Gelin, & Karplus, 1977). Since then, advances in computing power have enabled simulations of far larger proteins to be run that are 1,000–10,000 times longer than the original simulation (10–100 ns), in which the protein is surrounded by water and salt (Mackerell, 2004). Significant enhancements in the potential functions have also been made, resulting in significantly more reliable and accurate simulations (M. Karplus & Kuriyan, 2005).

The ability to construct simulations that approach the point where they can survive critical examination by the experimentalists who establish the structures of the proteins being simulated has resulted from a combination of increased computing power and enhanced potential functions (Lars Konermann, 2017; L. Konermann, McAllister, & Metwally, 2014).

1.7 Aims and objectives

The aim of this research is to develop a methodology to probe protein structure in the gas-phase using a combination of chemical modification and ion mobility mass spectrometry. If protein modification induces protein stabilization, or destabilization, then this will be reported by its unfolding pattern in CIU analysis.

The acetylation chemical modification of ubiquitin will be chosen due to its ability to turn a basic lysine site into a neutral amide. Then the Diethylpyrocarbonate (DEPC) modification will be explored to modify the histidine and lysine residues, followed by a succinylation modification to further modify basic residues.

The use of different proteins was essential to understand the impact of chemical modification and mutation on protein CIU. It was necessary to test various Ub mutants. So we acquired several mutants where lysine position was altered to have their CIU measured. Another protein, UBA2, was acquired for examination.

Finally, the role of CIU has been explored through a combination of CIU-MS studies and computational techniques. Molecular dynamic simulation of the Ub structure and its mutation were to be studied to better understand the molecular level of unfolding.

Chapter 2

Materials and methods

2 Materials and methods

2.1 Ubiquitin

Wild type recombinant ubiquitin was provided by Layfield Group (University of Nottingham), as expressed by Dr. Dan Scott. Samples were supplied as purified lyophilized stocks and were prepared as 20 μ M samples in 50 mM ammonium bicarbonate solution. Ammonium bicarbonate was purchased from Thermo Fisher Scientific (Loughborough, UK).

2.1.1 Ubiquitin mutants

Several ubiquitin mutants were provided by Layfield Group (University of Nottingham) and commercially from Boston BioChem that have been used in this experiment:

Ubiquitin mutant lysine 6 to alanine K6A were designed and expressed by Dr. Jed Long (University of Nottingham). All the ubiquitin mutants were lyophilized stocks and were prepared as 20 μ M samples in 50 mM ammonium bicarbonate solution. A summary of the Ub variants used can be found in Table 2-1.

Table 2-1. Mutant names and sequences.

Protein	MW	z	m/z	Ac m/z	Origin	Sequence
Ub	8565	5	1714	1722.4	Layfield Group	MQFVKLTGKTITLEVEPSDTIENVKAKIQDKEGIPP
(WT)	8565	6	1428.5	1435.5		DQQLIFAGKQLEDGRTLSDYNIQESTLHLVLRLLGG
Ub	8508	5	1702.6	1711	Layfield/Searle Group	MQFVATLTGKTITLEVEPSDTIENVKAKIQDKEGIPP
(K6A)	8508	6	1419	1426		DQQLIFAGKQLEDGRTLSDYNIQESTLHLVLRLLGG
Ub	8593	5	1719.6	1728	Boston BioChem	MQFVRTLTKKTITLEVEPSDTIENVKAKIQDKEGIPP
(K6R)	8593	6	1433.2	1440.2		DQQLIFAGKQLEDGRTLSDYNIQESTLHLVLRLLGG
Ub	8761	5	1753.2	1761.6	Boston BioChem	MQFVRTLTKRITLEVEPSDTIENVRARIQDREGIPP
(NoK)	8761	6	1461.2	1468.2		DQQLIFAGRQLEDGRTLSDYNIQRESTLHLVLRLLGG
Ub	8733	5	1747.6	1756	Boston BioChem	MQFVKLTGKTITLEVEPSDTIENVRARIQDREGIPP
(K6O)	8733	6	1456.5	1463.5		DQQLIFAGRQLEDGRTLSDYNIQRESTLHLVLRLLGG

2.2 Protein expression and purification

2.2.1 Ubiquitin binding domains

The UBA2 domains were expressed as GST fusion proteins in *E. coli* strain XL10-Gold, purified and thrombin cleaved. GST-fusion wild-type UBA2 plasmid was obtained from Dr. Dan Scott (University of Nottingham).

2.2.2 Preparation of ultra-competent cells

The Inoue method was used to produce highly efficient competent cell to transfer the mutant DNA plasmid (Im, 2011). the protocol reproducibly generates competent cultures of *E. coli* that yield 1×10^8 to 3×10^8 transformed colonies/ μg of plasmid DNA. The protocol works optimally when the bacterial culture is grown at 18°C , rather than the conventional 37°C , because the composition or the physical characteristics of bacterial membranes synthesized at 18°C are more favourable for uptake of the DNA plasmid.

Dimethylsulfoxide

High grade DMSO (dimethylsulfoxide, HPLC grade) obtained from Sigma-Aldrich (Poole, UK). Because of its membrane penetrating and water displacement capabilities, DMSO (Dimethyl Sulfoxide) is a polar, aprotic organic solvent that is often used as a cryoprotectant. It was added to cell culture media to avoid cell death during the freezing phase by reducing ice formation.

SOB media

A mixture of SOB medium was prepared adding NaCl (0.5 g), tryptone (20 g), deionized H₂O (950 ml), and yeast extract (5 g). After that a solution of KCl (250 mM) was added to the medium and the pH was adjusted to 7.0 with (5M) NaOH. Media were sterilized by autoclaving for 20 minutes. Finally, a sterile solution of (2 M) MgCl₂ was added.

SOC medium

SOC medium was prepared adding NaCl (0.5 g), tryptone (20 g), deionized H₂O (950 ml), and yeast extract (5 g), (20 mM) glucose. After the SOB medium was cooled to room temperature a sterile (1 M) solution of glucose has been added.

Inoue transformation buffer

Transformation buffer contained H₂O (800 ml), MnCl₂•4H₂O (10.88 g), CaCl₂•2H₂O (2.20g), KCl (18.65 g), 20 ml of PIPES (0.5 M, pH 6.7). The buffer was sterilized by filtration through a prerinsed 0.22- μ m Nalgene filter.

Preparation method of the competent cells

From a glycerol stock of XL-10 cells on chloramphenicol plates, bacterial cells were incubated for 16-20 hours at 37°C, A single bacterial colony (2-3 mm in diameter) was picked to incubate into 25 ml of SOB medium in a 250-ml flask. The culture was incubated for 6-8 hours at 37°C with vigorous shaking (250-300 rpm).

Several concentrations of the starting culture were prepared; first flask received 10 ml of starter culture, the second received 4 ml, and the third received 2 ml. all flasks were incubated overnight at 18-22°C with moderate shaking.

After 16 hours, the OD600 of all three cultures was measured and the flask that reached 0.55 was chosen to be used and transferred to an ice-water bath for 10 minutes. Cells were harvested by centrifugation at 2500 g (3900 rpm in a Sorvall GSA rotor) for 10 minutes at 4°C.

Cells were then resuspended gently in 80 ml in ice-cold Inoue transformation buffer and harvest by centrifugation at 2500 g (3900 rpm in a Sorvall GSA rotor) for 10 minutes at 4°C.

After pouring off the medium using a vacuum aspirator to remove any drops of remaining medium, cells were resuspended again in 20 ml of ice-cold Inoue transformation buffer. Finally, 1.5 ml of DMSO was added the bacterial suspension mixed by swirling and then store it in ice for 10 minutes. Aliquots of the ultra-competent cell were made by suspensions into chilled, sterile microcentrifuge tubes immersed in a bath of liquid nitrogen.

2.2.3 Preparation of solutions for protein overexpression and purification

Luria Broth (LB) was obtained from Sigma-Aldrich (Poole, UK). LB consisted of 40% tryptone, 40% sodium chloride and 20% yeast extract. The solution was prepared by adding 25 g LB to 1 L deionised water and autoclaved at 121°C, overnight.

Ampicillin stock 100 mg/mL stock prepared by dissolving 1 mg in 10 mL of deionised water and was filter sterilised.

Isopropyl- β -D-thio-galactopyranoside (IPTG) obtained from Roche Diagnostics Ltd (Burgess Hill, UK) and prepared as a solution to a final concentration of 200 mM.

Tri-buffered Saline with 0.1 % Triton (TBS-T) Solution consisted of 10 mM tris- HCl, 150 mM NaCl and 0.1% Triton X-100 at pH 7.5.

Thrombin cleavage buffer (TCB) Solution consisted of 20 mM tris-HCl, 150 mM NaCl and 2.5 mM CaCl at pH 8.4.

Thrombin was obtained from Sigma-Aldrich (Poole, UK).

2.2.4 Protein expression and sample preparation

Centrifugation for cell collection were performed in an Avanti JA-25-50 Centrifuge. Protein concentration was performed in a CS-15R (Beckman) centrifuge.

Protein samples were desalted in ammonium acetate and buffer exchanged for mass spectrometry analysis using ultracentrifugation spin filters with 3 kDa molecular weight cut-off (Sartorius, Germany) and an Eppendorf 5417C centrifuge.

2.3 Expression and purification UBA2s

The UBA2 domains (wild type and mutants) were expressed as glutathione S-Transferase (GST) fusion proteins in *E. coli* strain XL10-Gold. All constructs were verified by DNA sequencing. *E. coli* bacterial cultures were prepared by inoculating glycerol stock cells in 10 mL LB solution containing 10 μ L ampicillin stock (100 μ g/mL) and incubating overnight at 37°C with constant agitation (220 rpm). The solution was transferred to 1L of sterilised LB solution (25 mg/mL) that contained 1 mL of ampicillin stock (100 μ g/mL).

Bacteria were grown at 37°C for 4 h and then induced with 0.2 mM IPTG at OD₆₀₀ - 0.6 and incubated overnight at 20°C with agitation (200 rpm). Cells were pelleted by centrifugation (20 min at 4000 rpm) and were frozen for at least 2 h. Then cells were re-suspended with 10 mL of TBS-T solution and lysed by sonication (three times for 30 Sec at 6-7 microns with 1 min relaxation between sonication cycles).

Lysate was clarified by centrifugation (20 min at 15000 rpm) and the supernatant was applied to a 5 mL gravity polypropylene column (Qiagen Ltd, UK) containing 130 μ L glutathione beads (glutathione Sepharose 4B, GE Healthcare, UK) pre-equilibrated with TBS-T solution.

The lysate-glutathione beads mixture was incubated for 1 h at 4 °C with continuous mixing by rotation. After the binding of GST-UBD fusion protein, the beads were washed with three column volumes of TBS-T solution followed by one washing cycle with TCB solution and finally incubation with 5 U of thrombin (diluted in 1 mL TEB) at 4 °C overnight. Released UBA2 were eluted, and the beads were washed with TCB (3 mL).

2.4 Site-directed mutagenesis

The QuikChange II site-directed mutagenesis kit was used to make point mutations. Mutants of UBA2 were prepared by mutating the lysine 14 site (K14A, K14H) by site directed mutagenesis. Primers were designed using the QuikChange Primer Design Program and were ordered from Sigma-Aldrich. Using the QuikChange II method primers were inserted into pGEX-4T1 plasmid using QuikChange II Site-Directed Mutagenesis Kit the and PCR was performed to amplify the mutated DNA sequence. All Mutagenesis step were performed in PCR Gradient Thermal Cycler, and DNA sequences were verified by DNA sequencing.

The F (forward) and R (reversed) primer were initially diluted into 0.2 μ M and 0.5 μ L were added to the PCR tube, 1 μ l of the 50 ng dsDNA templet, 2.5 μ l pfu reaction buffer and 1 μ l of 200 μ M dNTP mixture (dNTP Mix is a solution containing sodium salts of dATP, dCTP, dGTP and dTTP), the 0.5 μ l of *PfuII* DNA polymerase (3 U/ μ l) was added at the end, prior to the thermal cycler reaction. Each Cycle reaction was done using the cycling parameters outlined in table 2-2.

Table 2-2. The cycling parameters for the QuikChange site-directed mutagenesis.

Segment	Temperature	Time	Cycles
1	95°C	2 minutes	1
2	95°C	30 seconds	18
3	60°C	1 minute	
4	68°C	6 minutes	
5	68°C	5 minutes	1

Following temperature cycling, the product amplification reaction was treated with 1 μ l of the *Dpn I* restriction enzyme (10 U/ μ l). The *Dpn I* endonuclease is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for

mutation-containing synthesized DNA. The nicked vector DNA containing the desired mutations is then transformed into XL1-10 super competent cells. Later was transferred into BL-21 and expressed the mutant protein.

2.5 Acetylation reaction

Acetic anhydride was obtained from Sigma-Aldrich (Poole, UK). Several equivalent acetic anhydride conditions were tested to acetylate the proteins, and then the MS spectra were generated, to obtain both the acetylated peak and the non-acetylated in the same MS spectrum. The method was optimized, and the ubiquitin sample was diluted to concentration of 10 μM into ammonium bicarbonate 50mM, an equivalent of 15 acetic anhydrides was added to acetylate the sample for 15 min.

2.6 Diethylpyrocarbonate reaction

Diethylpyrocarbonate (DEPC) 97% was purchased from AMRESCO, Inc. Several concentrations of diethylpyrocarbonate in acetonitrile (DEPC is insoluble in water, but it does dissolve in acetonitrile) was used to react with the ubiquitin to modify the protein, and then the MS spectra were generated, to obtain both the modified peak and the unmodified in the same MS spectrum, the method was optimized and the ubiquitin sample were diluted to concentration of 10 μM into ammonium bicarbonate 50 mM, an equivalent of 10 diethylpyrocarbonate was added to the sample for 4 min, the reactions are mostly carried out in -4 ice path.

2.7 Succinylation reaction

Succinic anhydride $\geq 99\%$ was obtained from Sigma-Aldrich (Poole, UK).

Several equivalent succinic anhydride condition was tested to modify the protein, and then the MS spectra were generated, to obtain both the succinylation modified peak and the non-modified peak in the same MS spectrum, the method was optimized, and the ubiquitin sample was diluted to a concentration of 10 μM into ammonium bicarbonate 50 mM, an equivalent of 100 succinic anhydride was added to react with the sample for 15 min.

2.8 Fourier Transform Ion Cyclotron Resonance (FT-ICR)

The FT-ICR (Thermo Fisher, UK), a Fourier Transform Ion Cyclotron Resonance MS apparatus with CID and ECD capabilities, was used in top-down proteomics experiments to characterise ubiquitin peptide fragment chains. The nanospray ion source (Thermo Scientific, UK) was installed in the apparatus. The Xcalibur 2.0 SR2 data system was used for instrument control and data processing. Ion positive mode was used to acquire all spectra.

2.9 Ion mobility-mass spectrometry

The Waters Synapt HDMS (Waters, Altrincham, UK) uses a traveling wave ion guide to perform IM separation ($t/\Delta t \approx 10\text{--}15$), coupled to a high-performance time-of-flight MS ($m/\Delta m \approx 20,000$) and was released as the first commercial IM-MS platform in 2006 (Pringle et al., 2007)

2.9.1 Mobility calibration standard

Lyophilised myoglobin, human heart cytochrome c, bovine red blood cell ubiquitin and bradykinin were purchased from Sigma-Aldrich (Poole, UK) and were prepared as 1 μM samples in $\text{CH}_3\text{OH}/\text{H}_2\text{O}/\text{CH}_3\text{COOH}$ (50:50:1 (vol/vol)) solution. Acetic acid was obtained from Thermo Fisher Scientific (Loughborough, UK).

2.9.2 Calibration and CCS calculation

Calibrations were performed under the same instrument conditions; ESI capillary voltage, 2.5 kV; sampling cone voltage, 20 V; extraction cone voltage, 3 V; source temperature, 40 °C; desolvation temperature, 50 °C; cone flow gas, 30 $\text{L}\cdot\text{h}^{-1}$, desolvation flow gas, 100 $\text{L}\cdot\text{h}^{-1}$. The ion mobility cell contained N_2 gas and operated at a pressure of 5.0×10^{-1} mbar; travelling wave height and velocity, 7 V and 280 $\text{m}\cdot\text{s}^{-1}$, respectively; transfer wave height and velocity, 3.0 V and 248 $\text{m}\cdot\text{s}^{-1}$. Respectively; trap collision voltage ramp from 6 to 22 V and constant transfer collision voltage 6 V, respectively; trap DC bias, 10 V; backing pressure, 3.8 mbar; trap pressure, 2×10^{-2} mbar; TOF region pressure, 1.3×10^{-6} mbar. Quad profile: Mass 1,700 m/z ; Dwell 1, 20%; Ramp 20%; Mass 2, 1700 m/z ; Dwell 2, 40%; Ramp 20%; Mass 32,000 m/z .

Calibration was carried-out using the ions from the following proteins: myoglobin, human heart cytochrome c, ubiquitin and bradykinin. These proteins were chosen to cover the ATD range of folded and unfolded ubiquitin, since their CCS values is known in the literature. The measured drift time (t_D) of each charge state of the three calibrant proteins was recorded and then converted to a corrected drift time (t_D'), using equation (2.1):

$$t_D' = t_D - 33 \left(\frac{\text{frequency}(\text{pusher})}{6} \right) \left(\frac{1}{1000} \right) - 1.41 \sqrt{m/z} \left(\frac{1}{1000} \right) \quad 2.1$$

where frequency (Pusher) = $90 \mu\text{s}^{-1}$ (Experimental Mass Spectrometer Parameter). The CCS values (Ω) from the databases, they were converted to corrected CCS values (Ω') using equation (4)

$$\Omega' = \frac{\Omega}{\left(z \sqrt{\frac{1}{\mu}} \right)} \quad 2.2$$

where, Ω is the CCS value as given in the databases and μ is the reduced mass that is calculated by equation (5)

$$\left(\frac{1}{\mu} \right) = \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \quad 2.3$$

A plot of $\ln \Omega'$ against $\ln(t_D')$ is constructed and its slope, m , was used to obtain (t_D'') Equation (6) values that were further used to construct the linear plot $\Omega = f(t_D'')$. The latter was used to calculate the experimental CCS values of protein ions.

$$t_D'' = t_D' m z \sqrt{\frac{1}{\mu}} \quad 2.4$$

2.9.3 Software

Magiplot : the gaussian fitting analysis was performed using (scientific plotting and data analysis software, 2021)

CIUSuite2 : The CIU data were plotted using either *CIUSuite2* (Eschweiler, Rabuck-Gibbons, Tian, & Ruotolo, 2015; Polasky, Dixit, Fantin, & Ruotolo, 2019) or custom Python scripts *CiuGui* .

MassLynx : All IMS–MS experiments were performed on a Synapt HDMS. These instruments were operated with the MassLynx (v4.0 and v4.1, Waters).

*Xcalibur*TM : All FTICR MS measurements were performed on a FTICR MS (Thermo Scientific) These instruments were operated with Xcalibur (v2.0, Thermo).

2.10 Molecular dynamics simulation

All simulations were performed by Dr. David Robinson using CHARMM (Brooks et al., 2009), with the Minerva high performance computing (HPC) facility at the University of Nottingham and the Hamilton HPC facility at Nottingham Trent University.

1UBQ was used as a starting point for the coordinates in the MD simulations using the PDB structure. First, the simulation was performed on WT and K6A Ub. The $[M+5H]^{5+}$ and $[M+6H]^{6+}$ charge states were produced in the following way:

WT $[M+5H]^{5+}$ ion – neutralisation of residues E24, D32, D39, E51, D58 and E64, and K63;

WT $[M+6H]^{6+}$ ion – neutralisation of residues E18, E24, D32, D39, E51, D58, E64, and K63;

K6A $[M+5H]^{5+}$ ion – neutralisation of residues E18, E24, D32, D39, E51, D58, E64, and K63,

WT neutral K6 $[M+5H]^{5+}$ ion – neutralisation of residues E18, E24, D32, D39, E51, D58, E64, K6 and K63.

Simulations were run at either 298 K or 750 K in two steps: a 0.5 ns equilibration step, whereby the temperature was steadily increased until it reached the desired degree, and a ca. 9 ns production step using the CHARMM force field (Huang & MacKerell Jr, 2013).

The timestep was set at 2 fs. Each system was subjected to three separate simulations. VMD was used to visualise the results (Humphrey, Dalke, & Schulten, 1996). which were then converted to PDB files for viewing in Pymol (*The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.*)

Chapter 3

Chemical modification

3 Chemical modification

3.1 Introduction

The use of chemical modification can alter protein structure by changing the charge on amino acid residues, changing the mass and disrupt other interactions, as detectable by IM-MS. Nucleophilic amino acid residues, such as lysine and cysteine, are susceptible to acetylation in the presence of a suitable acyl donor.(Mendoza & Vachet, 2009)

The importance of post-translational modification is shown by the modifications of the side chain of lysine, one of the three basic residues vital for protein structure and function (Zhang et al., 2011). The primary effect of reversible lysine acetylation is the neutralization of its positive charge at pH 7. As a result of shifting the site of the chemical reaction, the biochemical and structural properties of proteins are altered.(Mendoza & Vachet, 2009)

Using chemical modification, it is possible to make rapid changes to a large number of residues, and when paired with site-directed mutagenesis, it is possible to induce structural changes in even bigger proteins. This chapter investigates synthetic chemical modification in preparation for use with CIU IM-MS, which will be discussed in further detail in later chapters.

In this chapter, we will look at how to optimize the conditions for protein modification, as well as how to analyze the ECD and CID MS/MS spectra of changed proteins in order to identify the most favourable sites for chemical modification.

3.1.1 The optimisation of acetylation on wild type ubiquitin (WT-Ub)

To produce acetylated ubiquitin, wild type ubiquitin protein (WT-Ub), was incubated with a range of concentrations of acetic anhydride at room temperature. The modified Ub was tested after different periods of time using MS. The acetylation of WT-Ub resulted in a mass shift corresponding to the addition of C_2H_3O and the loss of H (+42.01 Da).

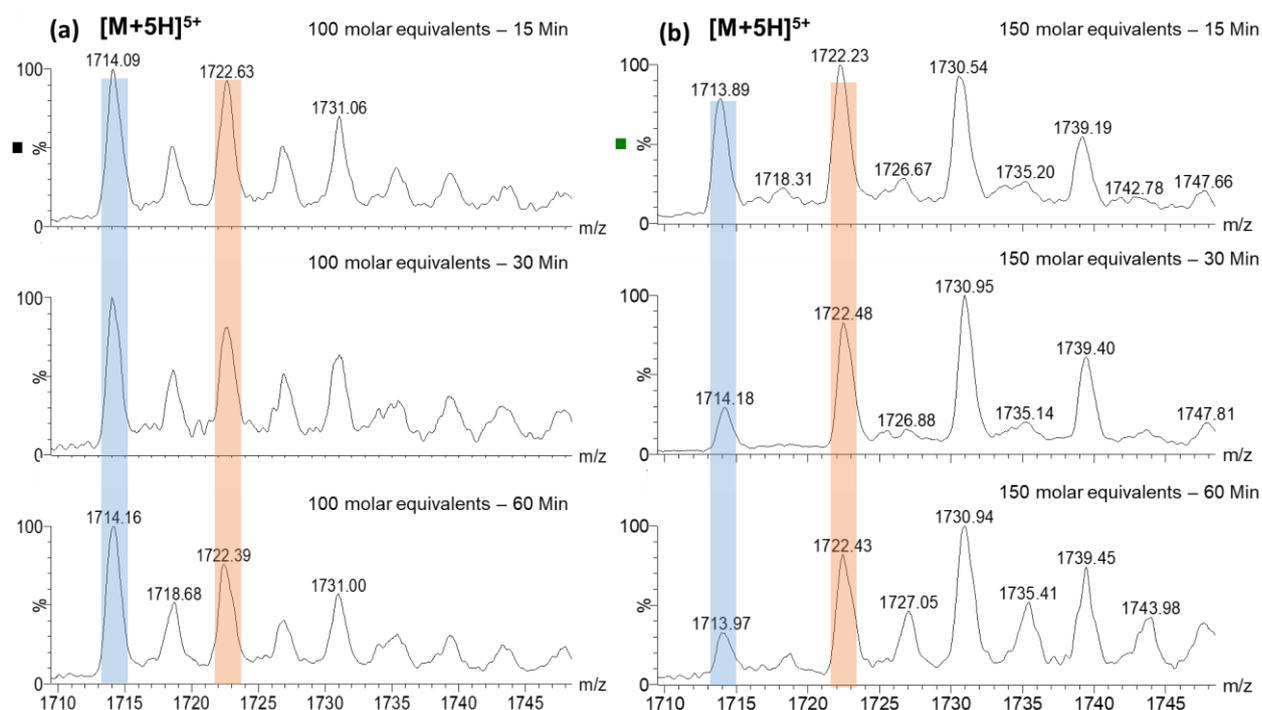


Figure 3-1. ESI-MS spectra showing $[M+5H]^{5+}$ ion (a) the first reaction of 100 molar equivalents of acetic anhydride to acetylate WT-Ub in ammonium bicarbonate for 15, 30 and 60 minutes. In contrast (b) the reaction of Ub with 150 Eq molar of acetic anhydride after 15, 30, and 60 minutes.

To obtain a spectrum with showing clear signals for both unmodified and acetylated Ub, a range of acetylation conditions were explored. It was found that the optimum condition for acetylation of Ub was to use 150 molar equivalents of acetic anhydride to acetylate WT-Ub in ammonium bicarbonate for 15 minutes at room temperature and then the reaction was quenched by adding ammonium bicarbonate (50 mM), see Figure 3.1.

To maintain the native structure of both ions, the sample was injected in an ESI source under conditions able to obtain the lower charge state ions ($[M+6H]^{6+}$ and $[M+5H]^{5+}$). Figure 3-2 shows the optimum spectrum of acetylation reaction for both charge states. Figure 3.2 (a)

shows the spectrum for WT-Ub before acetylation with acetic anhydride, while Figure 3.2 (b) shows the mass spectrum of Ub after acetylation and the location of acetylated ions.

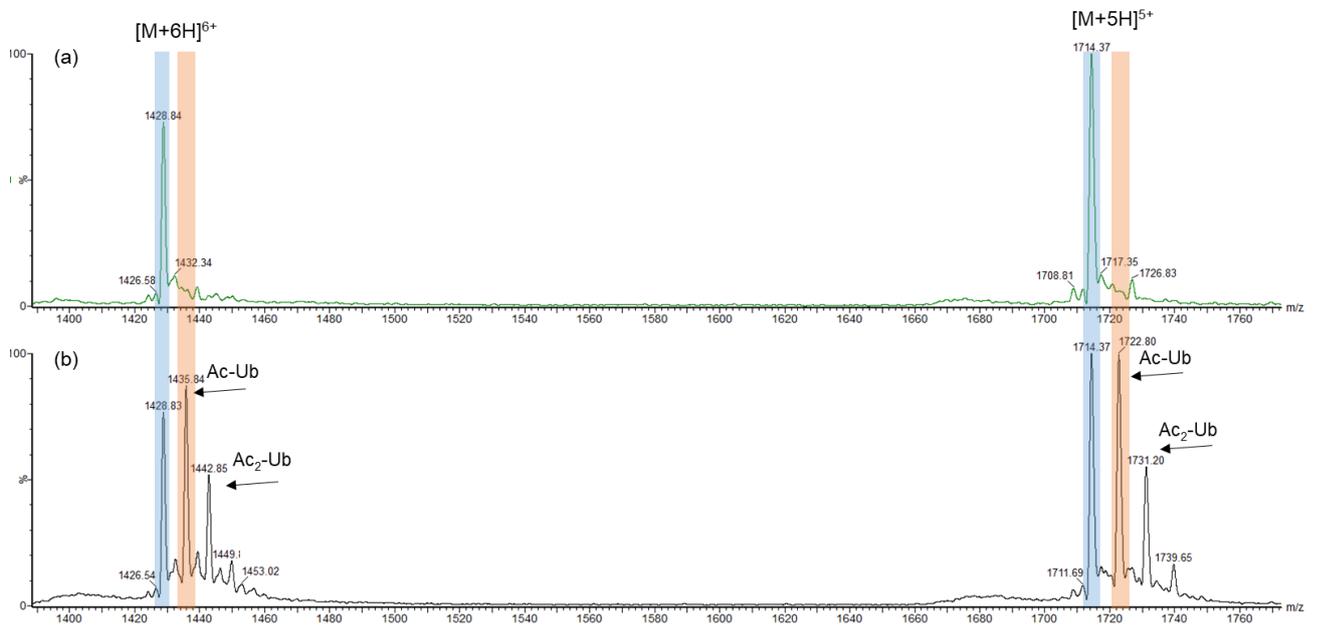


Figure 3-2. (a) ESI-MS spectra of WT-Ub $[M+6H]^{6+}$ m/z 1428.5, and WT-Ub $[M+5H]^{5+}$ m/z 1714.0. (b) Shows the acetylation position on WT-Ub $[M+6H]^{6+}$ at m/z 1435.6, and WT-Ub $[M+5H]^{5+}$ m/z 1722.3.

3.1.2 Tandem mass spectrometry

In order to identify the acetylation site(s), fragmentation of the protein ion is required. Electron capture dissociation (ECD) was chosen for this purpose. The ECD process occurs along the peptide backbone and is able to preserve modifications such as acetylation. ECD provides higher coverage than CID and it is more efficient in backbone fragmentation of intact proteins, for this reason the ECD technique combined with an FT-ICR mass spectrometer, has been selected for the analysis of acetylated Ub.

A targeted acquisition of ECD was performed in the cell of a Fourier transform FT-ICR mass spectrometer of intact monoacetylated Ub on the 11+ charge state of a denatured sample (see Materials and Methods for experimental details). Ub contains seven lysine residues and an N-terminal amino group that can be modified by the acetyl group so the fragments that contain lysine were investigated in details.

First, unmodified WT-Ub was denatured and fragmented to serve as control to compare it to the modified sample. Figure 3-3 shows the ECD MS/MS spectrum of unmodified WT-Ub $[M+11H]^{11+}$, showing the fragmented peptides of c-ions and z-ions. There was clear fragmentation showing the c-ions and z-ions such as C_{11} m/z 1263.75 that contains the N-terminus and K6, C_{24} m/z 2692.47 that C24 contains the N-terminus, K6 and K11 and K6. While when compared with Figure 3-4 that shows the same fragment with addition of 42.01 Da such as $(C_{11}+Ac)$ m/z 1305.75.

Using the relative intensities of the acetylated ions as a measure, valuable information about the extent of acetylation of each lysine residue was obtained. This information was obtained by comparing the relative acetylation for each acetylated ion and matching it to the corresponding peptide fragment.

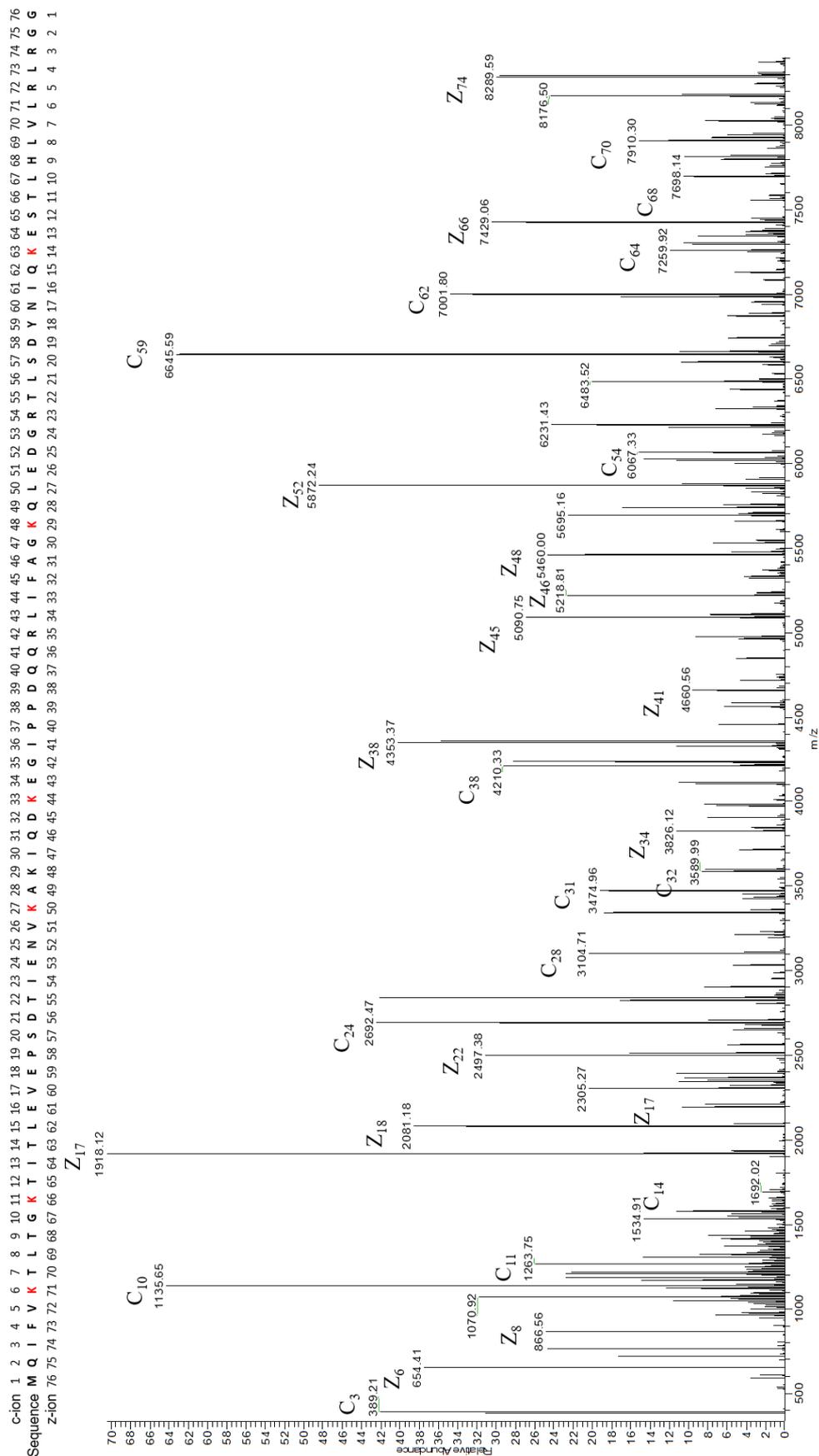


Figure 3-3 Deconvoluted ECD MS/MS spectrum of unmodified WT Ub [M+11H]¹⁺ showing c-ions and z-ions.

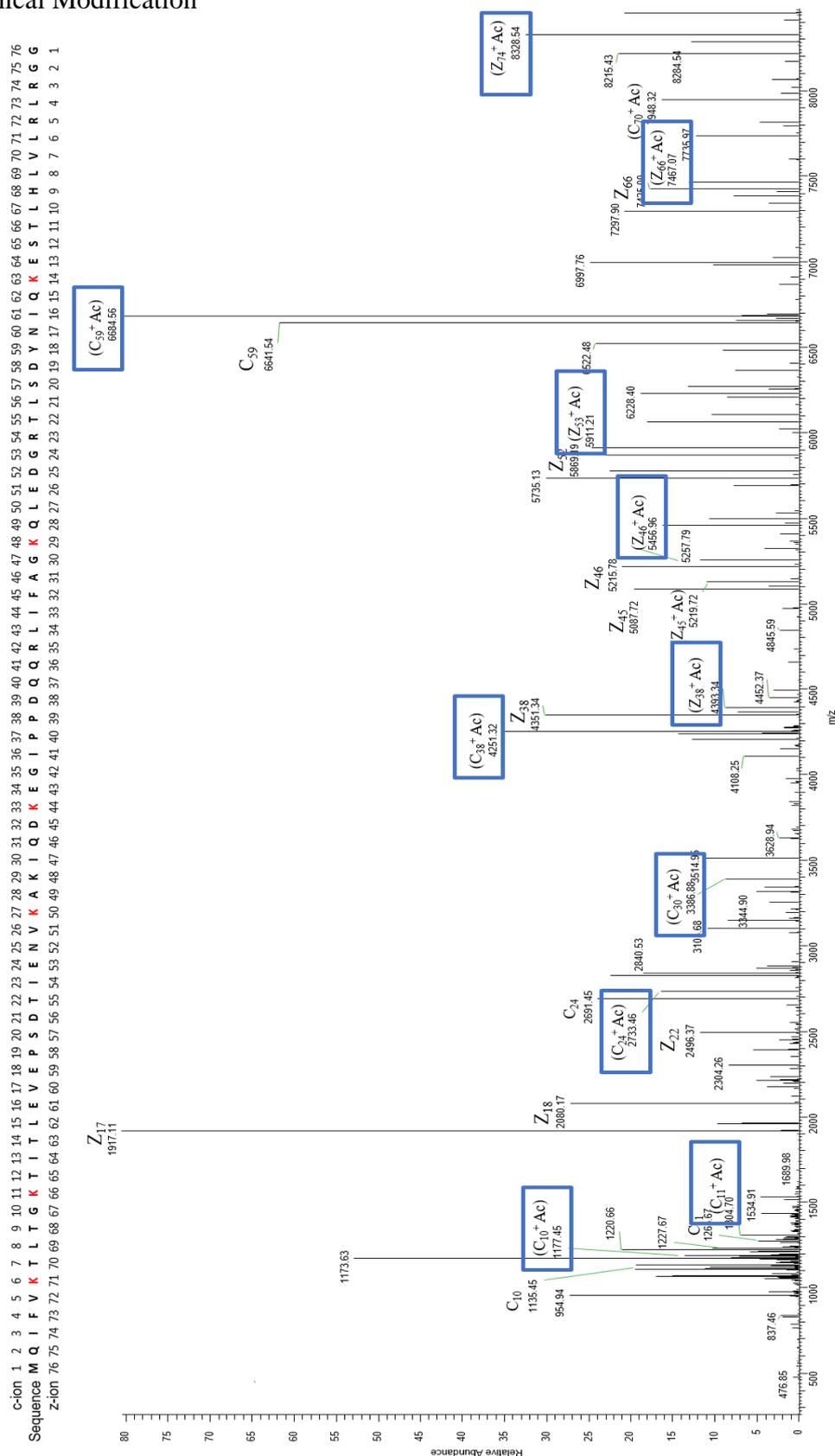


Figure 3-4 Deconvoluted ECD MS/MS spectrum of monoacetylated WT Ub [M+11H]¹¹⁺ showing c-ions and z-ions.

Based on the spectra produced by the ECD fragmentation, the locations and approximate proportions of modification were identified as shown in Figure 3-5. The majority of acetylation goes on K6 with value of 30%, while the N-terminus, K63, K27 and K29 show 15% acetylation, K48 have only 10% and no acetylation found on K11. Given these values, K6 shows the most prominent site for modification, which may or may not indicate importance for structural stability. In solution, this residue points out into bulk solvent and does not show strong interactions with other amino acids in the protein, which may make it more available for chemical modification. The effect of acetylation at this residue on the conformational stability of compact Ub ions in the gas phase will be examined in Chapter 4. Figure 3-5 show the percentage of acetylation on each residue on the native Ub structure.

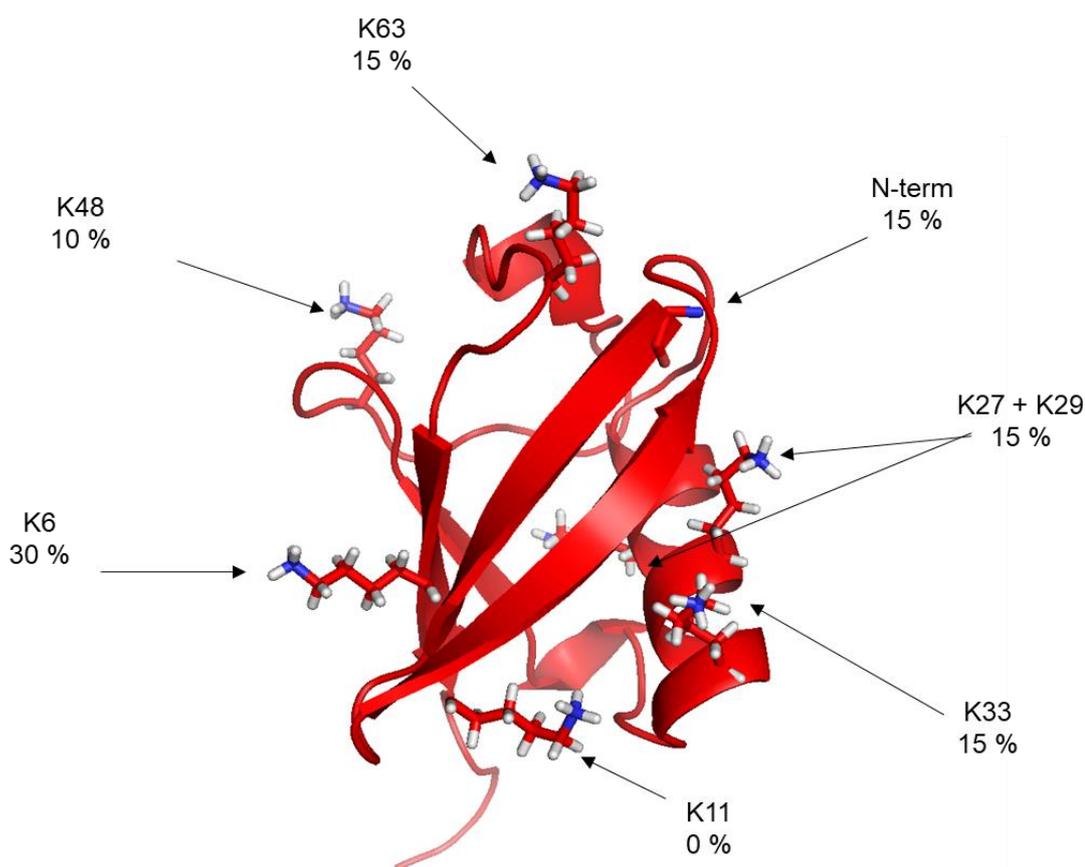


Figure 3-5. Extent of acetylation at each lysine residue of Ub as determined by ECD fragmentation.

3.2 Diethylpyrocarbonate modification

3.2.1 Modification of histidine and lysine residues in proteins by diethylpyrocarbonate

Diethylpyrocarbonate (DEPC) has been widely used to modify histidine residues (Mendoza & Vachet, 2009). When using low concentration of DEPC, carboethoxylation of histidine can occur and single modification would result. However, using higher concentrations can cause a second modification of histidine. Studies have shown that the pH range to 6-8 leads to modification on lysine, tyrosine, cysteine, arginine, serine, and threonine residues. However, several conditions would cause these sites to react such as increasing the temperature and time (Dage, Sun, & Halsall, 1998). Monitoring the reaction time and buffer pH range of 6-8 to achieve the required results to modify the lysine residue only.

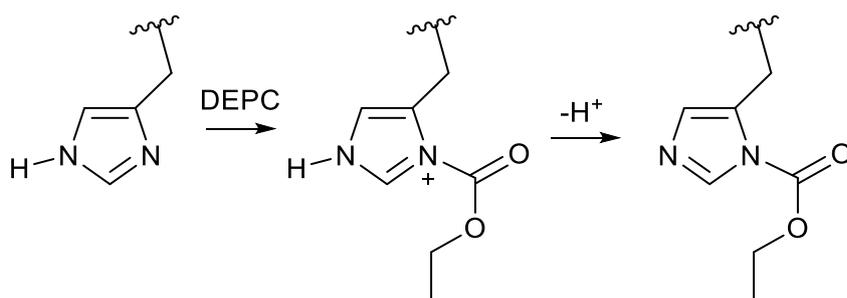


Figure 3-6. DEPC modification of the histidine imidazole ring, which leads to N-acylation but retention of an imino nitrogen atom capable of acting as an H-bond acceptor.

Initially, the method was optimized to achieve a single modification on histidine. The final reagent ratio of DEPC to Ub 10:1 was used, and the reaction times increased from 4 to 10 min. These conditions are suitable to study proteins under native conditions; however, because DEPC is not water-soluble, the addition of organic solvent is required. The amount of organic solvent should be as low as possible, so as to not disrupt the protein structure, the final concentration of solvent should be no more than 1% acetonitrile (Zhou & Vachet, 2012).

To maintain the native structure of the WT-Ub and DEPC, Ub was injected in the ESI source using ammonium bicarbonate solution 50 mM to obtain the lower charge state ion $[M+6H]^{6+}$ and $[M+5H]^{5+}$. The increase in mass shift (72 Da) for the modified protein is shown in Figure 3-7, which shows the carboethoxylation of WT-Ub $[M+6H]^{6+}$ at m/z 1440.38 and the second carboethoxylation at m/z 1452.39, for the $[M+5H]^{5+}$ the first modification is at m/z 1728.39 and the second modification at m/z 1742.81.

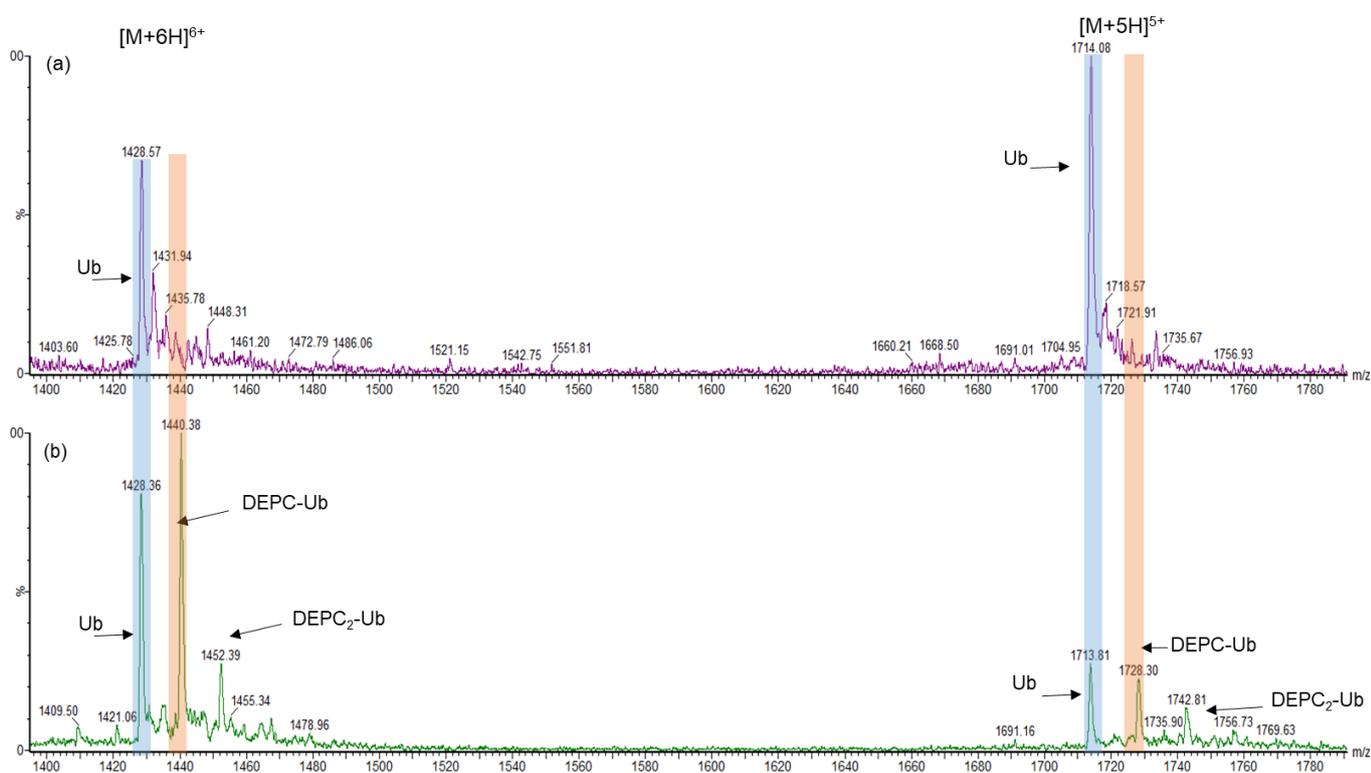


Figure 3-7 (a) ESI-MS spectra of WT-Ub $[M+6H]^{6+}$ 1428.5 m/z . and WT-Ub $[M+5H]^{5+}$ 1714.0 m/z . (b) Shows the DEPC Modification position on WT-Ub $[M+6H]^{6+}$ at 1440.38 m/z and the second DEPC position going on 1452.39 m/z , DEPC-Ub ion $[M+5H]^{5+}$ 1728.39 m/z and the second DEPC position going on 1742.81 m/z .

3.2.2 Tandem mass spectrometry ECD fragmentation for DEPC-based modification

In order to determine the location of modification, a top-down method was employed in conjunction with the ECD technique mixed with FTICR mass spectrometry, as was done before with acetylation. Ub with carboethoxylation was investigated, the charge state of the $[M+9H]^{9+}$ ion (which has the most significant peak) of a denatured sample indicated that solvent-accessible histidine (H68) was preferentially changed at low DEPC molar ratios, and that this was the most prominent peak.

There were two ions of interest created by the DEPC modification, with the first modification expected to be on a single histidine residue and the second modification expected to be either on the histidine or on additional nucleophilic residues, as predicted. Figure 3-8 shows the ECD MS/MS spectrum of singly modified Ub following DEPC treatment, the modified ions shown were calculated and measured against the peptide sequence to localise the modification position. The annotated peaks shows the majority of modification on the z-ions, so the search for the position had to be by narrowing down the unmodified peaks against c-ions. Figure 3-9 summarise these findings and localise the modification.

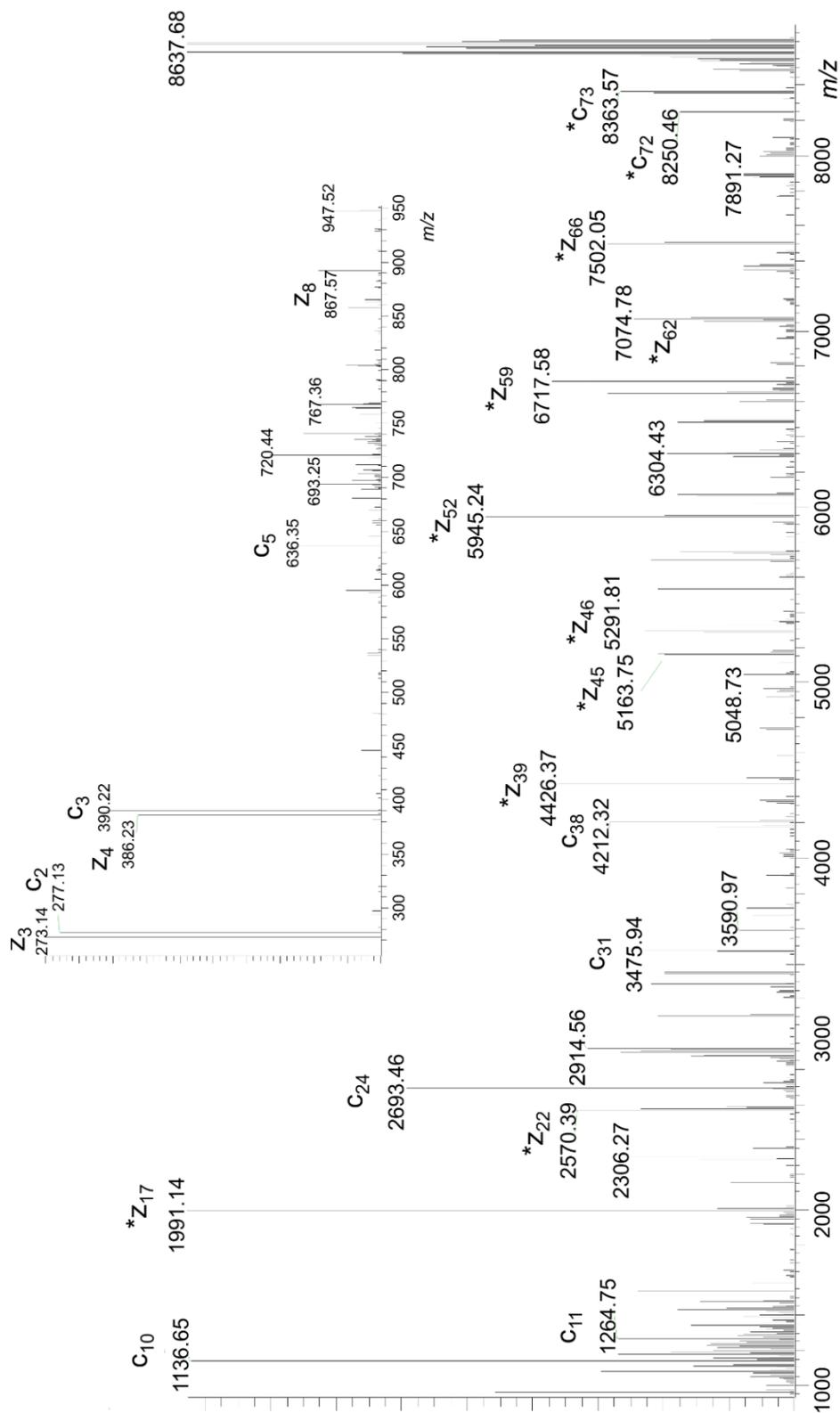


Figure 3-8 Deconvoluted ECD MS/MS spectrum of singly modified Ub following DEPC treatment ($[M+9H]^{9+}$) showing c-ions and z-ions. * indicates DEPC modification. The split of m/z axis and that the inset is a zoom from 250-950 m/z.

3.3 Succinylation modification

3.3.1 The use of succinic anhydride to modify lysine residues in proteins

Succinylation is one specific type of acylation reaction that may provide special advantages in some circumstances, it can lead to a change in the protein charge from positively charged residue into a negatively charged residue. The reagent required, succinic anhydride, being non-volatile, is easy to handle experimentally.

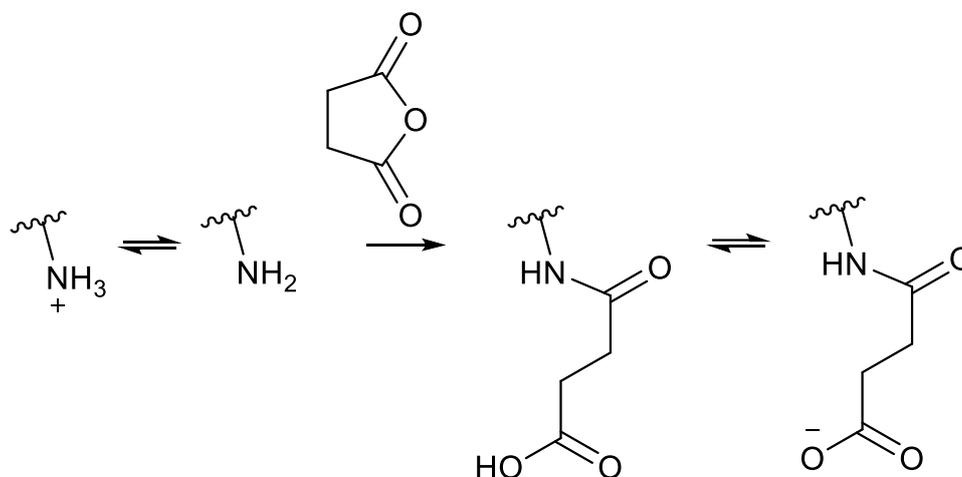


Figure 3-10. Succinylation of amino groups (K or N-terminus) by succinic anhydride, which converts basic residues into acidic ones.

At neutral pH, the succinylation reaction of lysine residues with succinic anhydride results in a charge reversal due to the presence of succinic anhydride. This reaction might result in a change in the overall charge distribution of the protein, which could provide important information regarding the unfolding behavior of the protein in the gas phase.

To examine the effect of succinylation on the structure and functions of protein, the modified WT-Ub tested by reacting the protein with succinic anhydride (100 mol. equivalent) to obtain MS spectra showing the modified peaks.

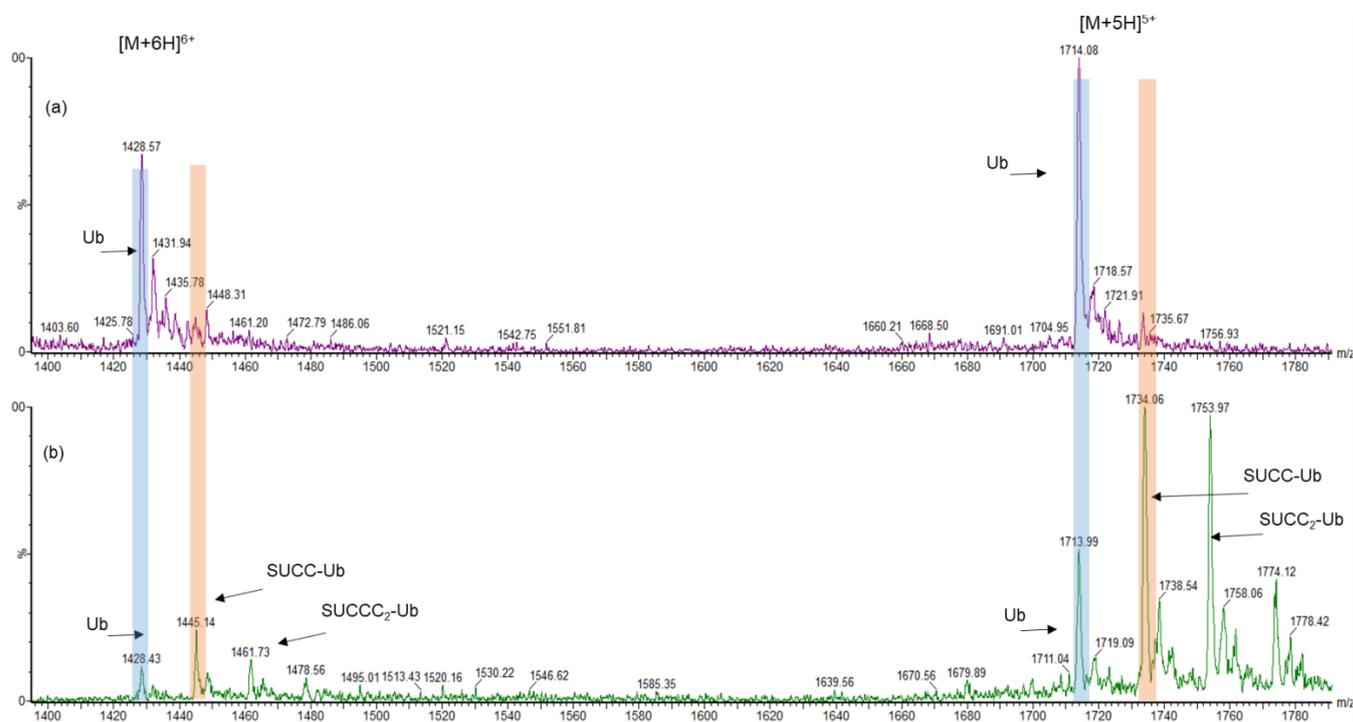


Figure 3-11. (a) ESI-MS spectra of WT-Ub $[M+6H]^{6+}$ 1428.5 m/z . and WT-Ub $[M+5H]^{5+}$ 1714.0 m/z . (b) Shows the first succinylation position on WT-Ub $[M+6H]^{6+}$ at m/z 1445.14. and WT-Ub $[M+5H]^{5+}$ m/z 1734.06.3 and second succinylation position going on second ion m/z 1753.97.

Figure 3-11 shows the spectra of the unmodified Ub and succinylated ubiquitin (SUCC-Ub), the spectrum in (b) shows number of modifications, which indicates that the modification takes place on several location on the protein, The $[M+5H]^{5+}$ charge state at m/z 1734 displays mass shift of 100.35 Da.

3.3.2 Tandem mass spectrometry ECD fragmentation for mapping succinylation sites

It was essential to use an MS/MS fragmentation analysis to find the mass addition in Ub since it comprises many lysine residues and a free N-terminus in order to determine its location. The $[M+9H]^{9+}$ ion was fractured with the application of ECD. In Figure 3-12, a fragment b_2 MQ at m/z 360.12 was identified as having a mass shift of 100.35 Da (unmodified ion mass at m/z 260.106). It was found to be the primary site of change and was located at the N-terminus. However, no other segment demonstrates any alteration, indicating that all of the succinylation occurs at the N-terminus and none at the K residues in the core of the protein molecule.

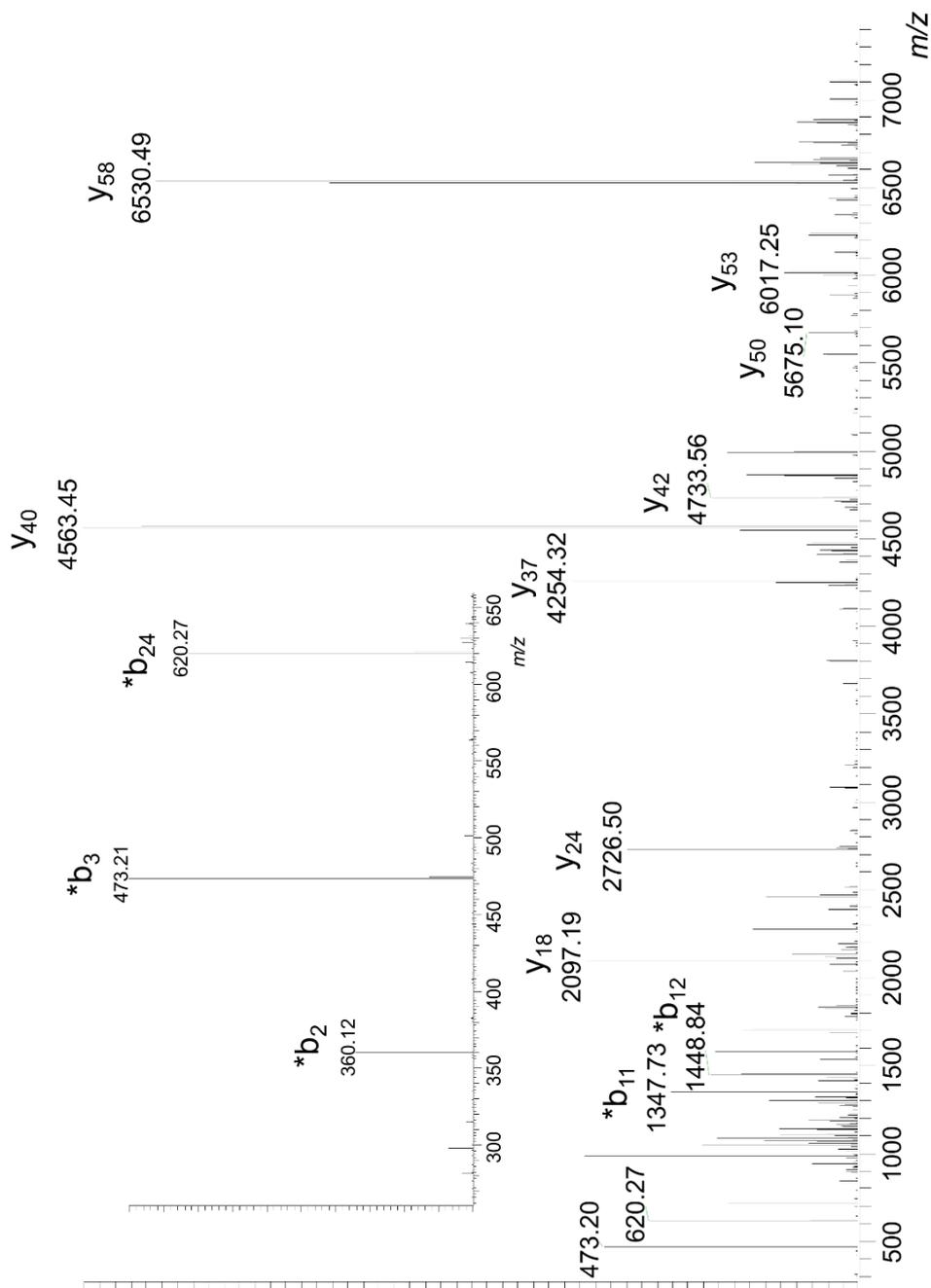


Figure 3-12 Deconvoluted CID MS/MS spectrum of singly modified Ub following succinylation ($[M+9H]^{9+}$) showing b-ions and y-ions. * indicates succinylation.

3.4 Conclusions

As demonstrated in this chapter, chemical modification can be used to probe proteins for mass spectrometry studies and protein characterisation. A variety of conditions for chemical modification have been investigated, and the most effective reagent concentrations have been determined.

The incorporation of CID and ECD fragmentation techniques has enabled the mapping of the amino acid residue location where the modification occurred, demonstrating the importance of linking future unfolding measurements by IM-MS to modification on specific residues in the protein (see later Chapters).

The results of the diethylpyrocarbonate modification reaction revealed that it had preferentially modified the histidine residue; however, further investigation is required to determine the location of the second modified residue. When the succinylation modification reaction was investigated by ECD, it was found that it was primarily concentrated on the N-terminus of the protein and that it had no effect on the lysine residues.

Chapter 4

Collision-induced unfolding of chemically modified ubiquitin

4 Collision-induced unfolding of chemically modified ubiquitin

4.1 Collision-induced unfolding of acetylated ubiquitin

Molecular structure of proteins is critical to understanding how they function, and studies in molecular biology have focused on determining the molecular structure of proteins. (Mendoza & Vachet, 2009).

Developing and employing instrumentation in the study of a protein structure like X-ray crystallography and NMR have been able to provide detailed information about a protein's three-dimensional arrangement. Yet these techniques have disadvantages such as cost and speed (Woods, Radford, & Ashcroft, 2013). The development of mass spectrometry (MS) in proteomic structure studies have overcome some of these disadvantages related to these techniques.

Mass spectrometry coupled with ion mobility (IMS) has been developed to investigate higher-order protein structure, and it has shown excellent results in terms of revealing protein structure rather than simply measuring mass to charge ratios. (Clemmer & Jarrold, 1997).

This chapter introduces the result achieved by spraying native protein into the IM-MS and comparing the results of collision induced unfolding (CIU) between the wild type and chemically modified ubiquitin to understand how modification at particular residues might affect the stability of compact protein structure in the gas-phase.

4.1.1 CCS for WT-Ub vs Acetylated WT-Ub for $[M+5H]^{5+}$

Collision cross section CCS calibration measurements were performed as in materials and methods (see Chapter 2). These measurements were based on the work of Ruotolo et al. The CCS was then calculated against known standard calibration CCS database from literature (Bush et al., 2010), the calibration curve in Figure 4-1 shows the resulting TWIMS calibration plots for calibrants bradykinin, ubiquitin, cytochrome C and myoglobin.

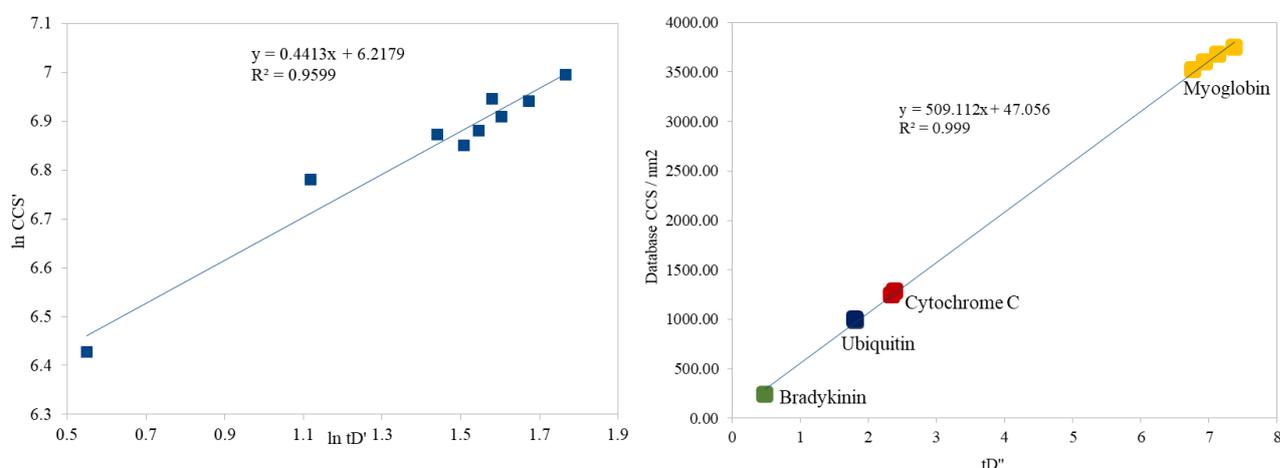


Figure 4-1 TWIMS calibration plots for calibrants bradykinin, ubiquitin, cytochrome C and myoglobin

CIU experiments require optimization of the acquisition method to achieve compact native structure of the protein and to show the unfolding pattern as the CE energy increases, it was revealed that there was shifting in the population of the compact structure to a more unfolded one, by acquiring the result-based ramp function where the trap collision energy gradually increased while the transfer collision energy was constant. This method is beneficial to illustrate the population shifting by increasing the energy.

The acquisition started at the lowest energy ($E_{lab} = 30$ eV) where the drift traces of acetylated and unacetylated $[M+5H]^{5+}$ ions as shown in Figure 4-2 are identical, which means the protein is still in its compact state with a CCS of 975 \AA^2 . By increasing the trap energy to $E_{lab} = 60$ eV the acetylated $[M+5H]^{5+}$ in red showed changes in the drift trace, with a pronounced shoulder

that become even more significant at $E_{lab} = 70$ eV, demonstrating clear unfolding of the acetylated form compared to Ub, to give a CCS at 1110 \AA^2 . By increasing the trap energy, this difference in the unfolding remains constant. Three experimental replicates were done, and the same result were shown. These results emphasize that the unfolded shoulder have high confidence.

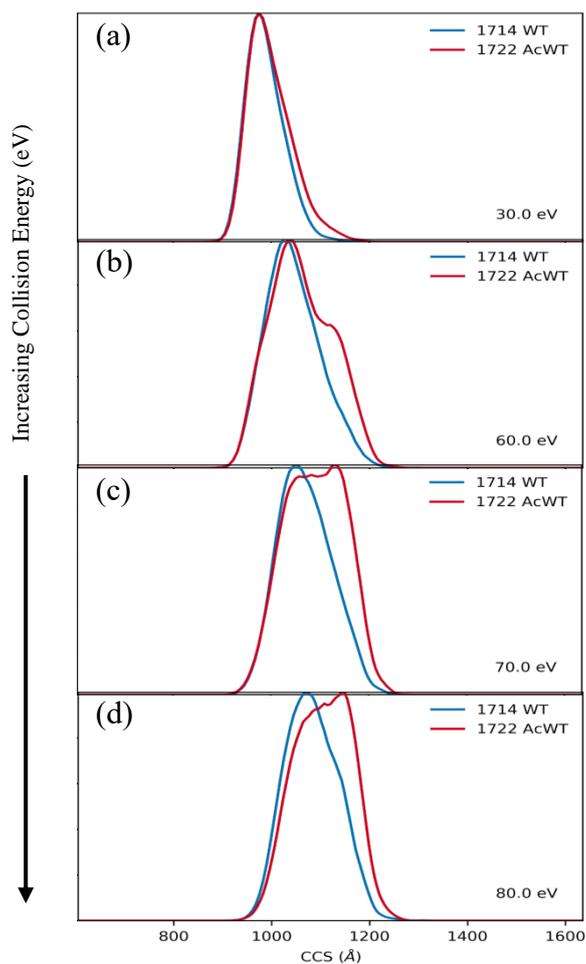


Figure 4-2. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for WT-Ub in blue vs Ac-Ub in red (a) the more compact structure at trap energy 30 eV where the first population is shown. Increasing the voltage to from 60 to 70 eV as shown in (b), (c) respectively indicate the first unfolding shifting in the population. (d) shows that at higher energy the second population is more prominent.

The sample was partially acetylated using acetic anhydride, as described in Chapter 3 (see also Materials and Methods) which facilitated the testing of CIU profile for the unmodified and

modified ion in the same sample, at the same time, therefore eliminating the impact of sample preparation and testing from the result of the acquisition while guaranteeing the same instrumental conditions was applied during the CIU application.

The Ac Ub ion m/z 1722 was isolated using the quadrupole analyser of the Synapt HDMS instrument and activated in the trap collision cell at (70 eV) collision energy. Partially unfolded protein ions were then separated by IMS and then fragmented in the transfer cell by CID fragmentation (90 eV).

The CIU of the b5 ion MQIFV was extracted first, with no or little variation in the unfolding conformation. When the CIU of the b6 fragment MQIFVKT, which contains the K6 location, was compared to the unmodified conformation, the CIU of the b6 fragment MQIFVKT revealed a larger unfolding pattern. Figure 4-3 shows the drift time of the first acetylated and non-acetylated b5 ion, this identifies K6 as the site of modification causing destabilisation.

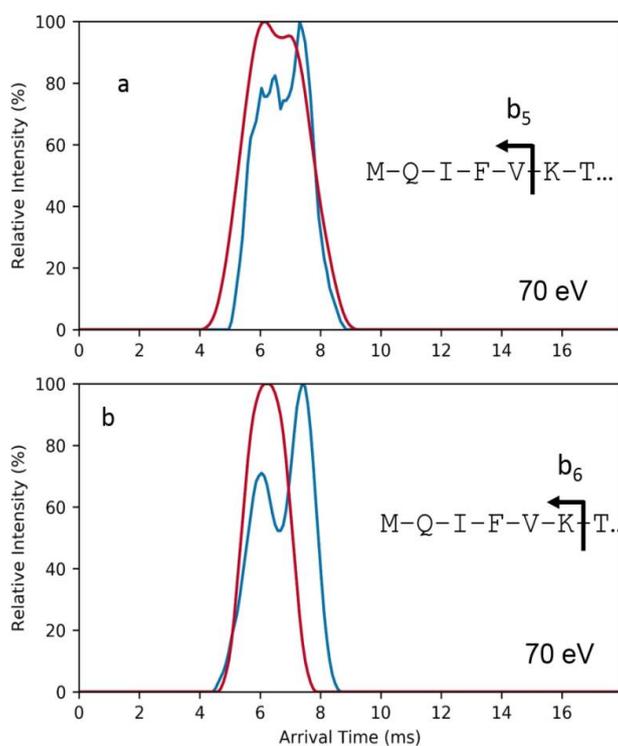


Figure 4-3. (a) ESI-MS spectra of $[M+6H]^{6+}$ ion (WT-Ub in red vs Ac-Ub in blue) shows the extracted IM drift traces at trap energy 36 V, (b) ESI-MS spectra of $[M+5H]^{5+}$ ion shows the extracted IM drift traces at trap energy 36 V.

Data analysis methods and techniques for handling the huge amount of data produced by CIU fingerprints must be created and applied in order to move forward in the use of CIU as a tool for general structural biology, as well as for high-throughput pharmaceutical applications especially.

To better represent the difference in the unfolding profile between the WT Ub and the Ac Ub *CIUSuite2* was used to generate Figure 4-4 which shows the plotted collision cross section (CCS) of the 5+ ions for each species as a function of collision voltage.

CIUSuite2 is a software developed by Ruotolo research group for the analysis and interpretation of native mass spectrometry data (Allison et al., 2020). *CIUSuite2* was used to compare the stabilities of proteins against collision-induced unfolding.

The ‘*CIUSuite_compare*’ function in the software was used to show the simplest comparison of CIU fingerprints by matrix subtraction and visualization of the two ions in different colours. The RMSD parameter is used by *CIUSuite compare* to display the absolute difference between two matrices and displays the RMSD on the difference plot.

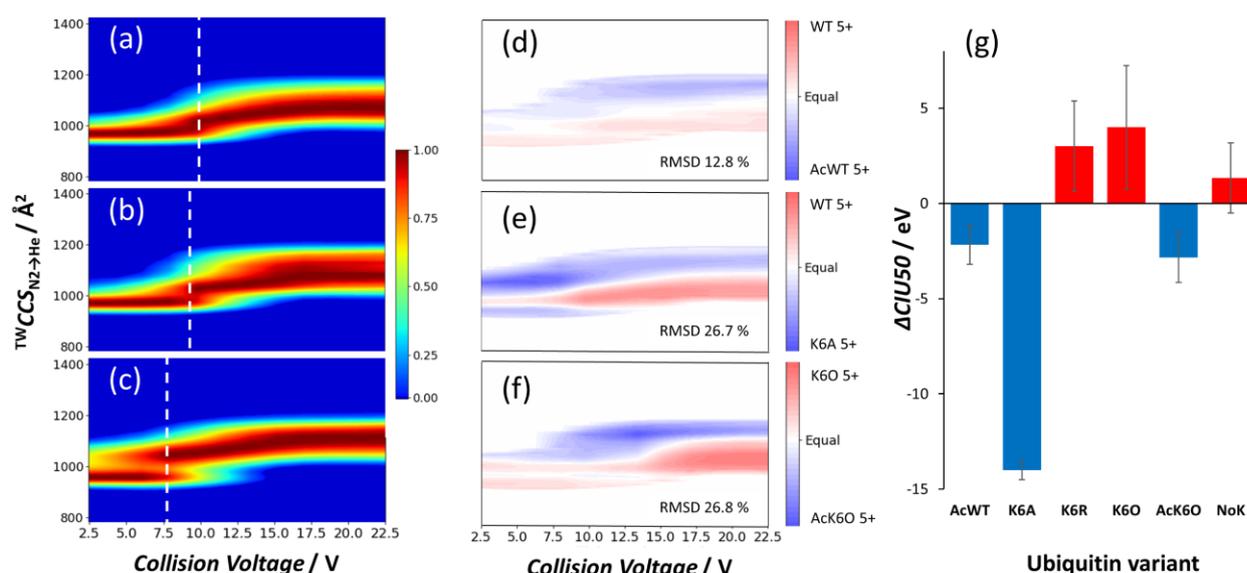


Figure 4-4. CIU data showing the unfolding of 5+ Ub ions (a) WT-Ub, (b) AcWT-Ub, and (c) K6A mutant Ub. Difference plots for unacetylated vs acetylated WT Ub, unacetylated WT vs K6A mutant Ub, and unacetylated vs

acetylated K6O Ub are shown in (d), (e) and (f) respectively. ΔCIU_{50} values in (g) reveal increased unfolding due to destabilisation of acetylated WT Ub, K6A Ub and acetylated K6 only Ub. Values are relative to unacetylated WT Ub except for acetylated K6O, which is relative to unacetylated K6O.

The E_{lab} values required for 50 % unfolding, termed CIU_{50} , were extracted and ΔCIU_{50} , the difference between CIU_{50} s for modified and unmodified Ub, calculated. To produce quantitative value to measure the destabilisation of unmodified wild type WT-Ub and compare it to the acetylated WT with value of -2.2 (± 1.0) eV after acetylation. To report the absolute difference between the modified and unmodified ions, the root-mean-square deviation (RMSD) parameter was measured and plotted against each other. Figure 4-4. show a value of (RMSD 12.8 %) for acetylated Ub (blue) distributed at higher CCS values relative to Ub (red) during the unfolding process, data of K6A, K6O and NoK will be discussed in Chapter 5.

4.1.2 CCS for WT-Ub vs Acetylated WT-Ub for $[M+6H]^{6+}$

The same ramping method was applied on the $[M+6H]^{6+}$ ion, the results show that the unfolding of the native structure starting at $E_{lab} = 12$ eV and reaching the final unfolding population at $E_{lab} = 144$ eV. However, the $[M+6H]^{6+}$ ion was more sensitive to the trap collision energy since the charge was higher than the $[M+5H]^{5+}$ ion, so several optimizations to the method were required to maintain the native compact structure to be monitored at the lowest energy.

The $[M+6H]^{6+}$ ion signal was small and easily affected by the adjacent drift trace of the adduct it caused disturbance to the CCS of the ion, after comprehensive search by filtering the sample from all sodium we found that the Na^+ adduct caused shift in the mass of the ion, by using Zebra Spin Desalting Columns we were able to purify the sample. After that we adjusted the profiling method by mass isolating the ion of importance in the quadrupole analyser of the Synapt HDMS instrument and activated in the trap collision cell. The correct mass, and ion, was confidently isolated according to this method.

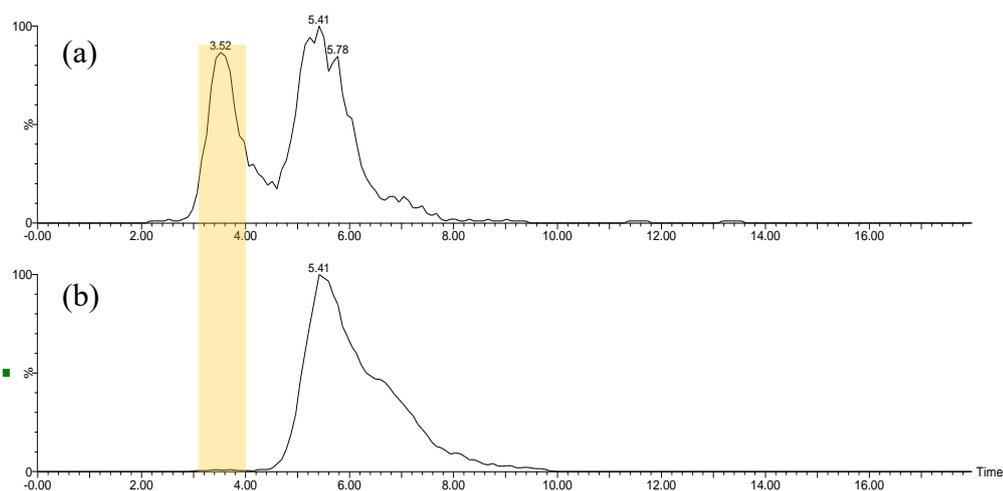


Figure 4-5. IMS-MS of WT-Ub $[M+6H]^{6+}$ m/z 1435.6 drift trace (a) ESI-MS spectra before treatment and mass isolation. (b) IMS-MS after desalting and mass isolation.

A CIU profile for the $[M+6H]^{6+}$ ion was performed and the m/z 1428 Ub ion compared to the m/z 1435 Ac Ub ion. CIU profiling at multiple energies were formed by starting to use the lowest trap collision voltage possible to maintain the native structure yet able to detect the ions. Trap energy was then used at $E_{lab} = 12$ eV as first acquisition parameter, then increased up to $E_{lab} = 132$ eV as the final energy where maximum unfolding occurred. The transfer energy was kept at constant value of ($E_{lab} = 36$ eV) to prevent the further fragmentation unfolding of the protein.

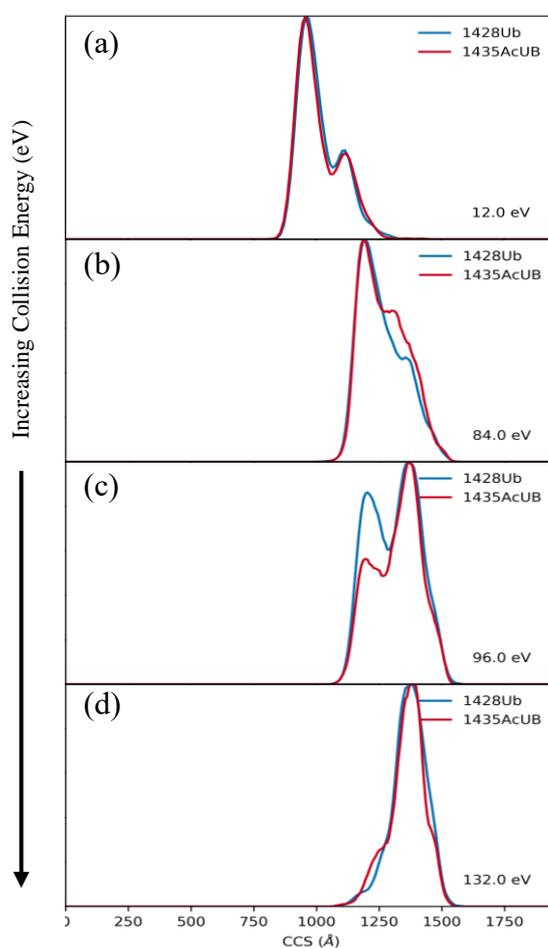


Figure 4-6. A detailed collision voltage ramp is conducted with a single charge state Ion (6+) mobility drift traces obtained for WT-Ub in blue vs Ac-Ub in red. (a) the more compact structure at trap energy 12 V where the first population is shown. Increasing the voltage to from 84 to 96 V as shown in (b), (c) respectively indicate the first unfolding shifting in the population. (d) shows that at higher energy the second population is more prominent.

4.2 Collision-induced unfolding for DEPC and SUCC modification

The previous section discussed the acetylation reaction and showed a clear result regarding the importance of lysine residue K6. To understand better the consequence of modifying amino acid residues, various chemical reactions were exploited to modify different residues. Diethylpyrocarbonate (DEPC)-based modification and succinylation SUCC by succinic anhydride (see Chapter 3) were performed on the WT-Ub and the result were complementary to the acetylation reaction results.

4.2.1 $[M+5H]^{5+}$ ion CCS comparison for WT-Ub vs. DEPC-modified-Ub

The two ions for of DEPC-Ub were compared against the unmodified Ub and the difference in population shape and unfolding started to show at 60 eV. This indicated that modification affected conformational stability of the protein leading to destabilisation as the result.

Figure 4-7 (a-d) shows the singly modified ion $[M+5H]^{5+}$ m/z 1728 having a different unfolding population than the doubly modified ion ion $[M+5H]^{5+}$ m/z 1743, which shows more destabilisation at intermediate collision energy, but reaches the same unfolding population at the higher energy. As described in Chapter 3, the first modification is associated with His modification, whereas the second occurs at lysine residues. Hence, these findings indicted further that modification on lysines cause the destabilization on protein CIU profile.

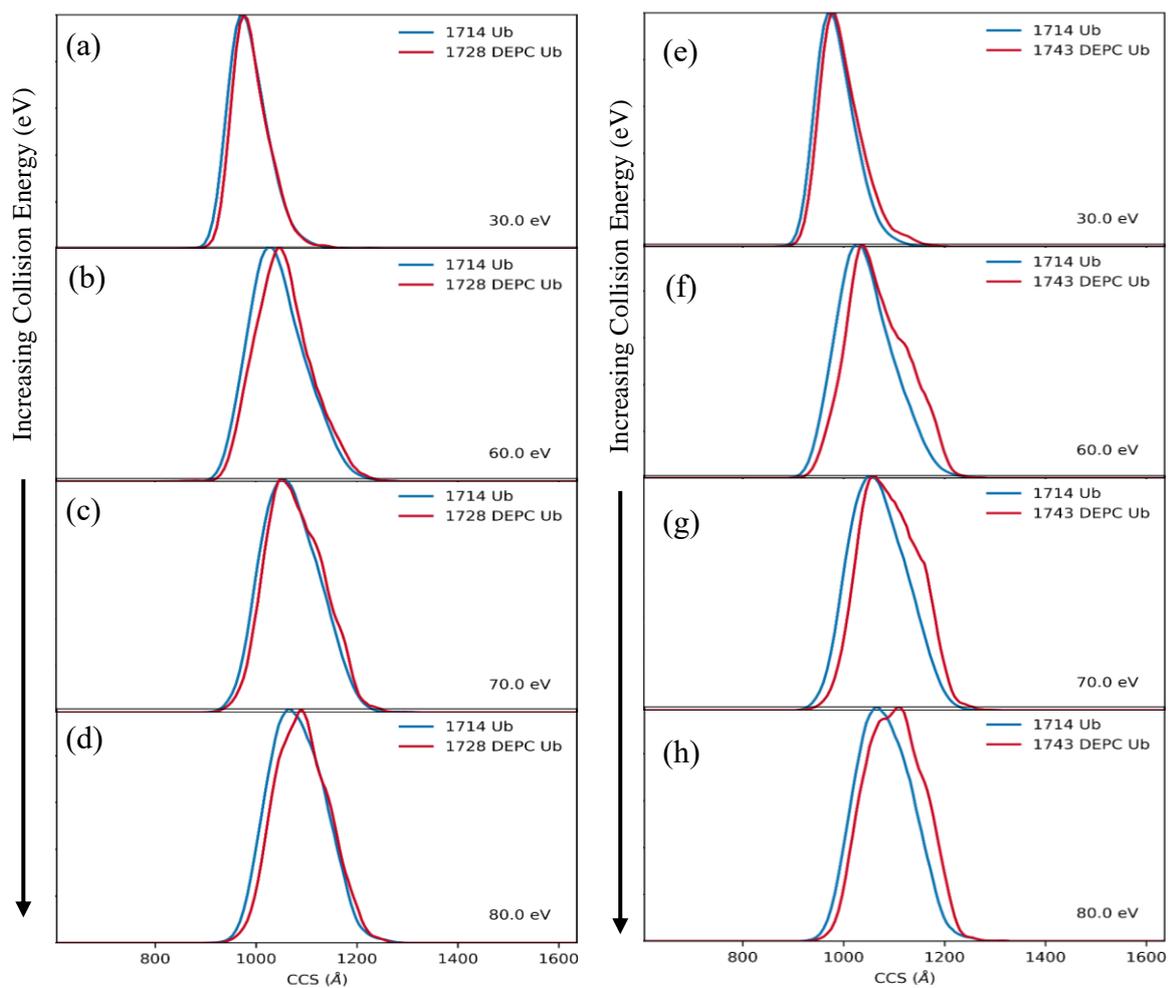


Figure 4-7. A detailed collision voltage ramp is conducted with a single charge state $[M+5H]^{5+}$ Ion mobility drift traces obtained for WT-Ub in blue vs DEPC-Ub in red the first column shows the first m/z 1728 ion while the second column shows the second ion m/z 1743. (a,e) the more compact structure at trap energy (10 eV) where the first population is shown in the green line. Increasing the voltage to from 40 to 80 V as shown in (b,f), (c,g) respectively indicate the first unfolding shifting in the population. (d,h) shows that at higher energy.

4.2.2 [M+5H]⁵⁺ ion CCS comparison for the WT-Ub vs. SUCC-Ub

The acetylation and DEPC modification induced significant variation in unfolding behaviour, to achieve a comprehensive picture there was a need for another chemical modification, so the succinylation reaction was the last to be performed. The result for this modification is shown in Figure 4-8 for the two ions that have been modified, the first ion m/z 1734 shown from (a-d) displays no difference in the unfolding conformation, it is possible that this is due to the fact that succinylation adds 100 Da on Ub, making it possible that the effect the presence could be attributed to the large mass added which cause the population shift to the first ion m/z 1734.

Another possibility could be that the first modification occurs at the N-terminus, which could be the reason that the protein structure is not affected and no shows no difference in the CIU comparing to the un modified protein.

The second ion m/z 1754 have 200 Da added mass and shows to some extent a shift in the population unfolding shown in figure (e-h), however looking at the CIU and comparing to the previous chemical modification and mutant there is no disruption to unfolding behaviour, at mid energy where 80 V the modified and unmodified ions show similar unfolding but with reasonable shift attributed to the addition of large mass or the fact that the second modification occurs on unmodified lysine residues, resulting in moderate destabilization.

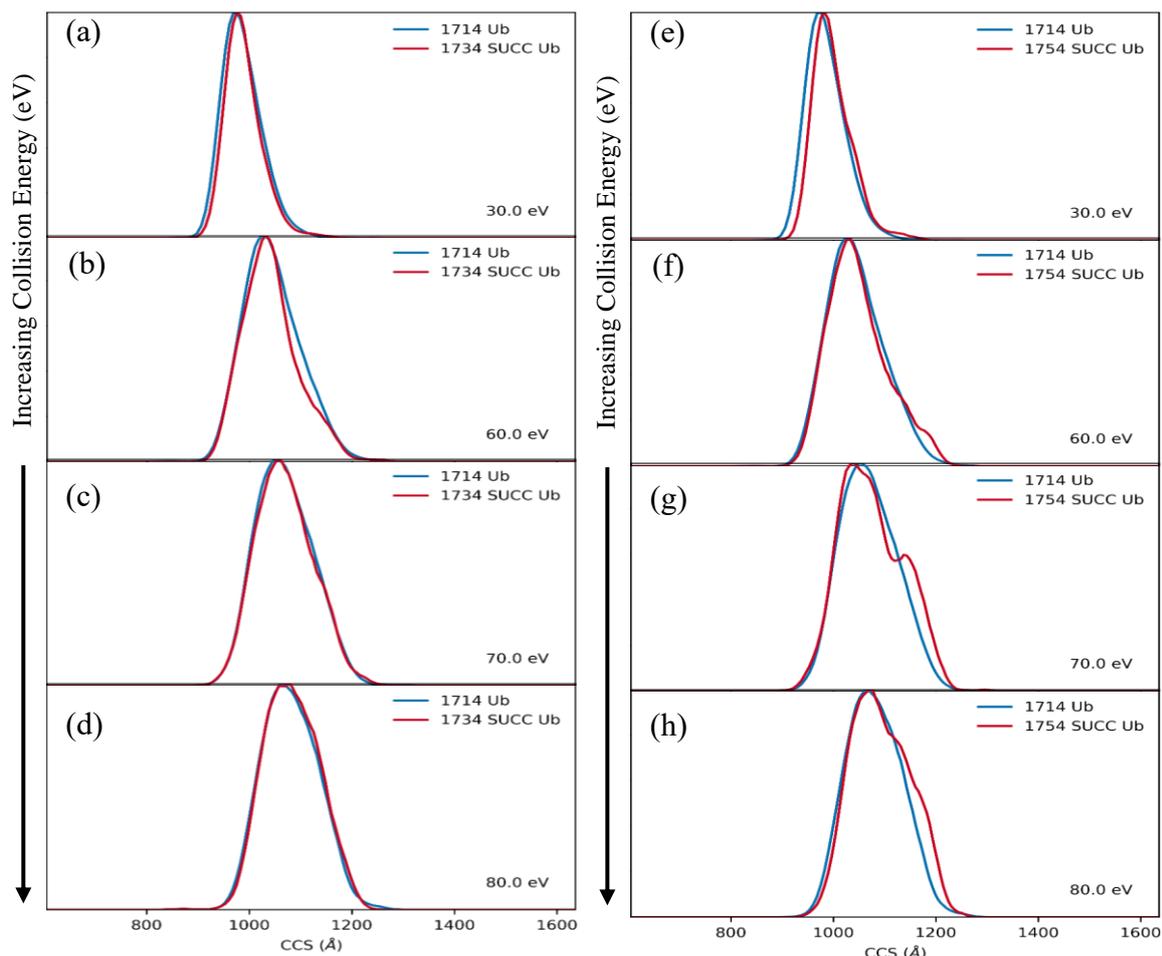


Figure 4-8. A detailed collision voltage ramp is conducted with a single charge state $[M+5H]^{5+}$ ion mobility drift traces obtained for WT-Ub vs SUCC-Ub. (a,e) the more compact structure at trap energy (10 eV) where the first population is shown in the green line. Increasing the voltage to from 40 to 80 V as shown in (bf), (cg) respectively indicate the first unfolding shifting in the population. (d,h) shows that at higher energy the second population is more prominent.

Succinylation with unsubstituted succinic anhydride places substantial negative charge on the protein, generally an increment of -2 per succinyl group, since a positively charged $-NH_3^+$ from lysine is replaced by $-COO^-$ of succinate half-amide. With a protein consisting of a single polypeptide chain, this increase in electrostatic repulsion can produce an expansion of the macromolecule.

4.3 Conclusions

The collision-induced unfolding (CIU) tests were carried out on the WT-Ub first, and the resulting profile was measured and computed. The resulting *CCS* was utilised as the control conformation profile, which was then compared to the changed Ubs in the following studies.

When the *CCS* of the WT-Ub was compared to the *CCS* of the acetylated WT-Ub, a significant difference was observed, indicating that the acetylated Ub had been destabilised, which can be explained by preferential unfolding of the compact conformation due to the addition of an acetyl group on the lysine 6 residue.

Histidine modification of the protein was performed to assess its potential impact on stability. DEPC is an electrophile that is able to acylate the imine nitrogen atom of histidine's imidazole ring with an ethoxycarbonyl group. Subsequent deprotonation of the pyrrole-like nitrogen neutralises the iminium and regenerates the tautomeric imidazole ring, with the ethoxycarbonyl group on the newly generated pyrrole-like nitrogen.

The succinylation result suggested that modification of the N-terminus, even by a group that changes the polarity of the position, is not detrimental to the stability of the compact gas-phase structure of Ub. In contrast, a second succinylation of the protein was found to destabilise the structure significantly. As with DEPC, this was attributed as being probably due to modification at K6.

Chapter 5

Collision-induced unfolding of genetically modified ubiquitin

5 Collision-induced unfolding of ubiquitin mutants

5.1 Introduction

The result obtained from the previous chapter indicate the importance of K6 to the overall structure and charge distribution of the Ub in the gas-phase. Therefore, to obtain complete picture, the analysis of different type of Ub mutants were studied. As described in the experimental section. The Ub mutants are Ub protein that have the amino acid position K6 (and elsewhere) altered from the native sequence by changing the original DNA sequence using the site-directed mutagenesis method.

Site-directed mutagenesis is a molecular biology method that is used to make specific and calculated changes to the DNA sequence. It can include single and combinational mutations to produce mutant protein.

A mutated protein can have a single amino acid altered or several positions. Four types of mutant were analysed. First was the K6A mutant where only the K6 had been modified to A. Second, was the NoK mutant where K6, K11, K33, K27, K29, K48 and K63 were been modified to R. Third, K6R where only K6 was modified to R, finally the K6-only mutant (K6O) in which only the K6 had not been modified to R.

5.2 [M+5H]⁵⁺ ion CCS comparison for mutants vs. WT-Ub

For the objective of understanding the mechanism of unfolding of each mutant the CIU drift trace was measured as in Chapter 4, and the CCS of the conformers was compared for each mutant against WT-Ub. The [M+5H]⁵⁺ charge state was first studied.

5.2.1 Ubiquitin mutant (K6A)

The first Ub mutant to be studied was K6A, where lysine 6 was altered to alanine. This changes position 6 from protonated charged residue to a neutral residue. The change is one single chargeable site, which made a huge effect on the unfolding profile of the protein Figure 5-1. In order to obtain clean spectra for K6A the mutant, several cycles of desalting, to remove any adducts, were required. The presence of any contamination from sodium ions had a major effect on the unfolding, since the protein became more unstable after changing the amino acid residue.

Figure 5-1 shows the CIU of K6A compared to WT-Ub at a range of energies, the first plot shows the unfolding at low energy $E_{lab} = 30$ eV, and there is already a clear difference between the mutant and WT Ub, with K6A-Ub showing a shoulder due to unfolding. This indicates that the mutant is extremely unstable and may have already unfolded in the source, which indicates how much changing the charge on this position could lead to destabilization of the structure.

At a higher energy $E_{lab} = 70-80$ eV, Figure 5.1 showed that the mutant has completely unfolded. K6A-Ub showed a larger CCS value than WT-Ub during the unfolding process. The alteration of the lysine at position 6 to alanine causes less stabilisation, this result reinforces the importance of K at this position for stabilisation of the global protein structure.

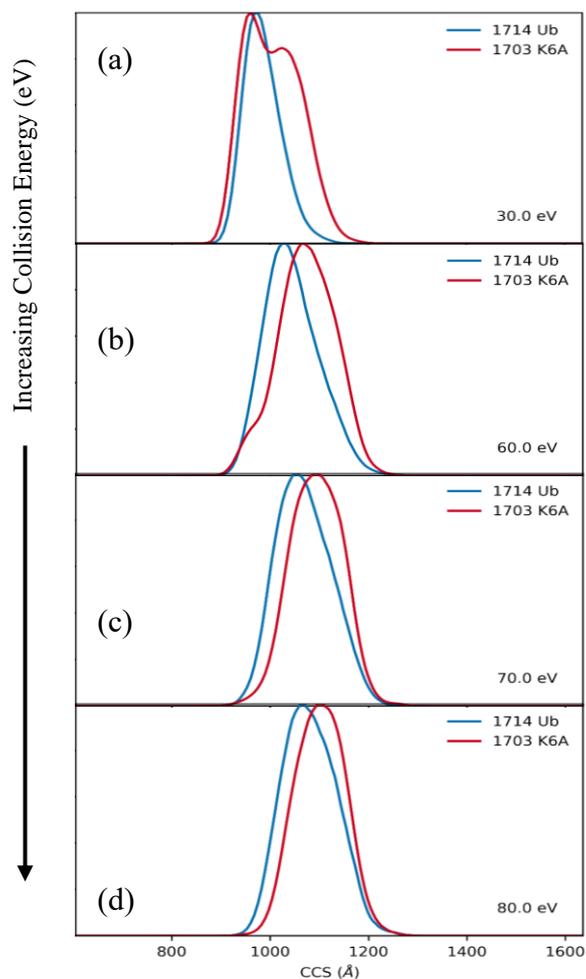


Figure 5-1. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for 1714 WT-Ub (blue) vs 1703 K6A (red). (a) the more compact structure at trap energy 30 eV where the first conformation is shown. Increasing the voltage from 60 to 70 eV as shown in (b), (c) respectively indicate the first unfolding shifting in the conformation. (d) shows that at higher energy the second conformation is more prominent.

5.2.2 Ubiquitin mutant (K6O)

The K6O mutant, where all the lysines in the Ub have been altered to arginine except K6, showed similar conformation to the WT-Ub, the drift trace shows that the unfolding of the mutant protein exhibit somehow similar *CCS* to the WT-Ub with little difference due to the mutation of lysine residue to arginine. Figure 5-2 shows the plot of K6O unfolding compared to the WT-Ub at a variable number of energies.

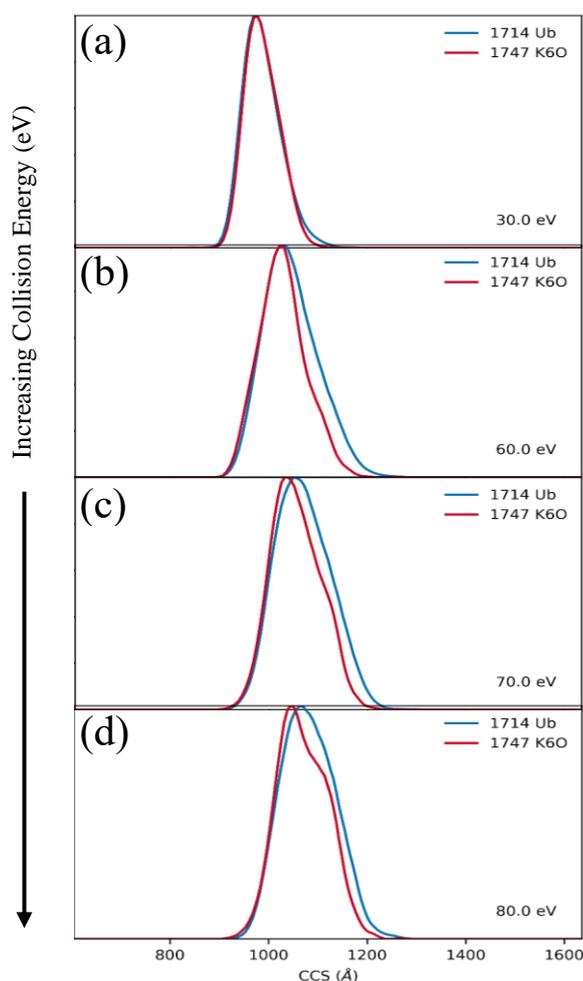


Figure 5-2. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for 1714 WT-Ub (blue) vs 1747 K6O (red). (a) the more compact structure at trap energy 30 eV where the first conformation is shown. (b), (c),(d) shows that at higher energy increasing the voltage to from 60 to 70 eV shows the final conformation.

The first plot shows the lowest energy $E_{lab} = 30$ eV where the two protein have identical *CCS*, increasing the energy up to 70 eV promote slight change in the unfolding but not as prominent as with the K6A mutant, which all can be explained by the significance of the K6 position, altering all the other K in the protein caused little or no different in the *CCS*.

The K6O Ub mutant contains one lysine position, which is interesting for it to be chemically modified by acetylation, since only the K6 can be modified it can emphasise on the important role of K6 for stabilization of the protein. Comparison of the acetylated and unacetylated K6O in Figure 5-3 showed similar behaviour to WT-Ub when acetylated, at lower energy the drift trace exhibited a similar profile but when the energy increased to 60 eV a shoulder started to appear in the acetylated protein mutant just as we have observed in the acetylated WT-Ub.

At higher energy $E_{lab} = 70-80$ eV, the plot of the acetylated mutant appeared to have entirely unfolded and different *CCS* values than the unmodified protein, as this protein contains only positions where the acetyl group could have been placed, and exhibits the same unfolding profile as the WT-Ub. This result strongly supported the finding that chemical modification at K6, specifically, has a detrimental effect on the stability of the compact conformation of Ub in the gas phase.

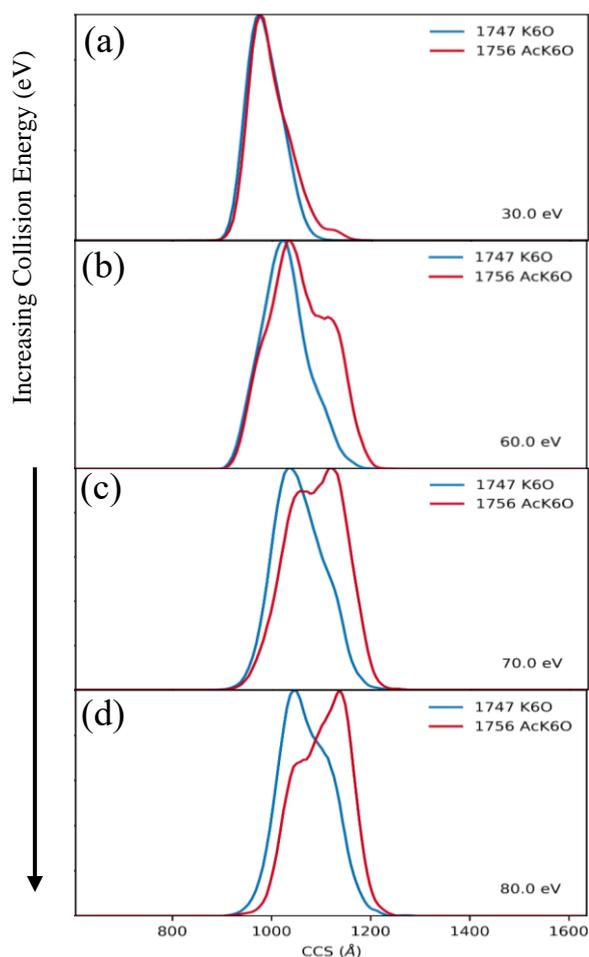


Figure 5-3. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for 1747 K6O (blue) vs 1756 AcK6O (red). (a) the more compact structure at trap energy 30 eV where the first conformation is shown. Increasing the voltage to from 60 to 70 eV as shown in (b), (c),(d) shows that at higher energy.

5.2.3 Ubiquitin mutant (K6R)

The K6R mutant (where only position K6 was altered by mutagenesis to arginine) in contrast to the K6A mutant, appeared to be more stable than the WT-Ub. When the CIU experiment was conducted on this mutant, the conformation of K6R was stable to the energy variation from $E_{lab} = 30-80$ eV Figure 5-4.

This result of changing K6 to R caused stabilisation due to the nature of the R sidechain that applying higher energy did not cause unfolding comparing to the WT-Ub. Arginine is a more basic residue than lysine, making its protonation, or ability to take part in stabilising intramolecular salt bridges, more likely, which may explain the additional stability seen by this variant.

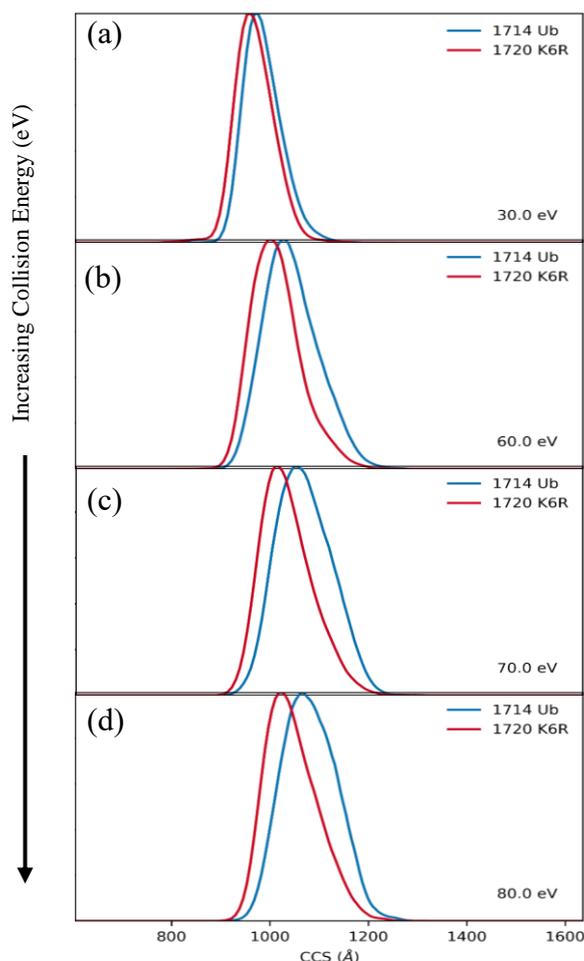


Figure 5-4. detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for 1714 WT-Ub (blue) vs 1720 K6R (red). (a) the more compact structure at trap energy 30 eV where the first conformation is shown. Increasing the voltage to from 60 to 70 eV as shown in (b), (c) respectively indicate the first unfolding shifting in the conformation. (d) shows that at higher energy the second conformation is more prominent.

The findings from K6R were supported by the evidence of acetylated K6R, by modifying K6R mutant by chemical modification did not cause any change to the conformation, the CCS in the Figure 5-5 had no distinction difference between the unmodified and acetylated mutant.

Acetylation of K6R has no effect on the *CCS*, consequently the destabilising effect of acetylation is dependent on the presence of K at position 6. This means that acetylation of K residues, other than K6, did not lead to significant destabilisation of the protein, which further emphasises the key role of position 6 in Ub.

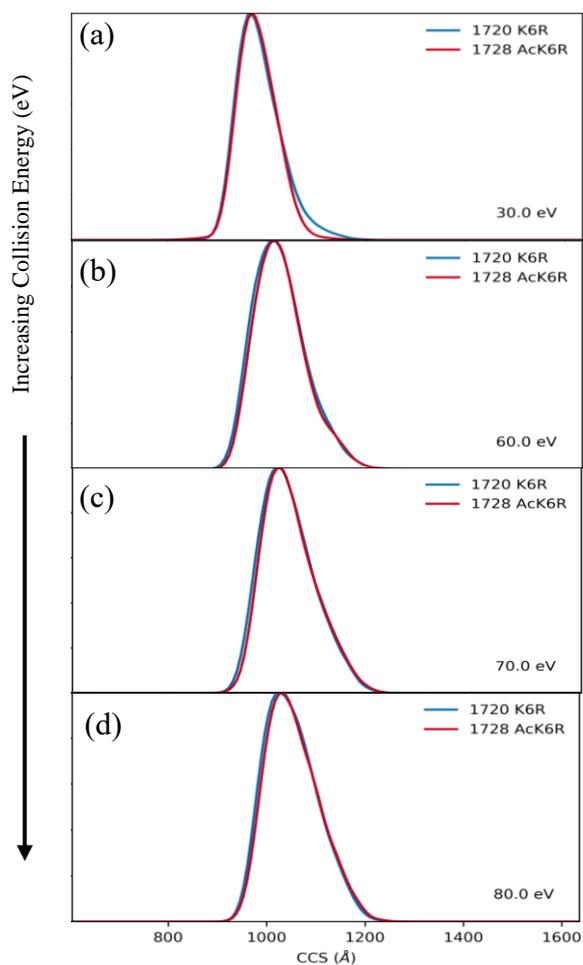


Figure 5-5. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for 1720 K6R (blue) vs 1728 AcK6R (red). (a) the more compact structure at trap energy 30 eV where the first conformation is shown. Increasing the voltage to from 60 to 70 eV as shown in (b), (c), and (d) show no change in the unfolding.

5.2.4 Ubiquitin mutant (NoK)

The NoK mutant contained no lysine residues, since all the lysines were altered to arginine. This mutation causes the protein to be more stable than any other mutant. By comparing the drift trace of NoK in Figure 5-6 to that of the WT-Ub, it appears that applying collision energy causes the WT-Ub to unfold while the NoK mutant does not show any unfolding. This shows that the presence of arginine residues lead to stabilisation of the protein. This result may be interpreted as arginine possessing stronger intramolecular interactions with other regions of the protein. Because there was no lysine in this mutant, there was no point in conducting the acetylation reaction.

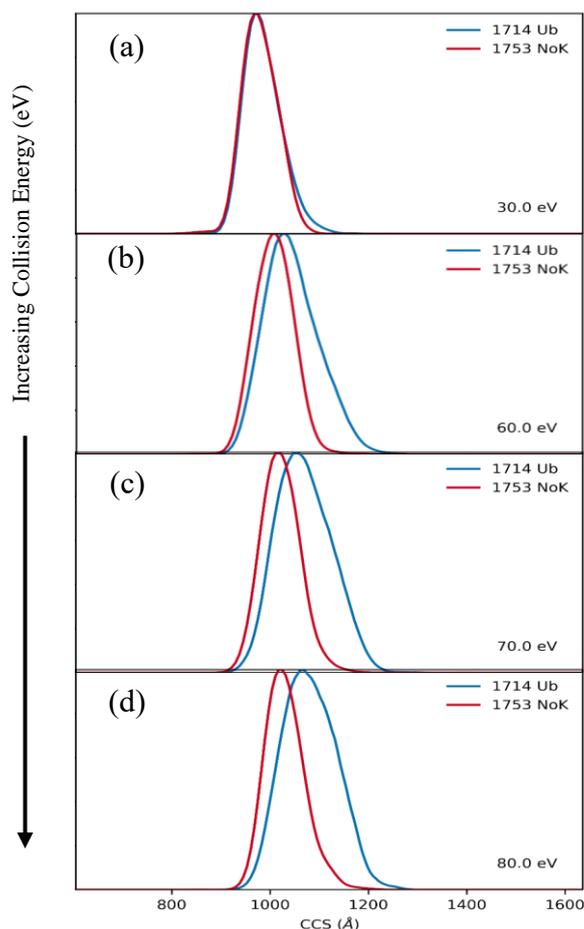


Figure 5-6. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for 1714 WT-Ub (blue) vs 1753 NoK (red). (a), (b), (c), and (d) showed the unfolding of WT-Ub vs. the NoK where it shows a stable drift trace.

5.3 $[M+6H]^{6+}$ ion CCS comparison for the mutants vs. WT-Ub

The second ion that was analysed was the $[M+6H]^{6+}$ charge state of Ub. Since it had lower signal intensity, it was challenging to acquire acetylated data for K6A mutants, nevertheless the full CIU profile was obtained for each mutant and compared to the WT-Ub. The resulting data showed clear resolved conformers that could be investigated.

The data from $[M+6H]^{6+}$ ion represent important complementary information to the $[M+5H]^{5+}$ ion since the presence of charge on the K6 position is considered vital to the structure of Ub, and the analysis of second charge state provides a comprehensive picture. Figure 5-7 shows the general CIU fingerprints for all the Mutants, this shows the overall picture of the difference between the mutants.

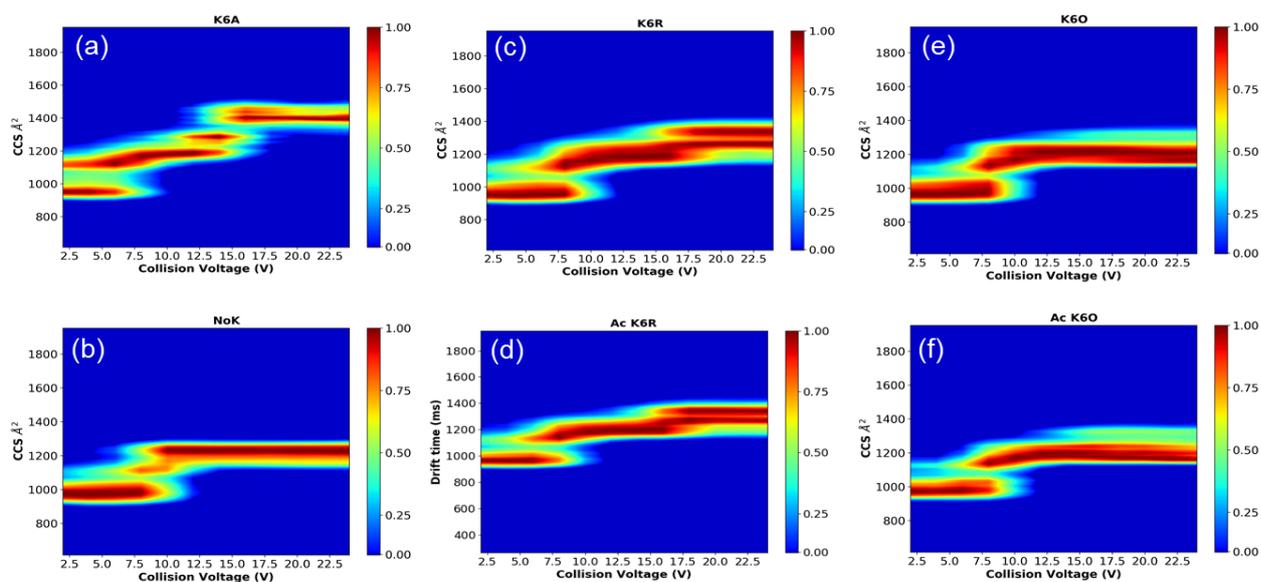


Figure 5-7. Overview of CIU fingerprints for (a) K6A, (b) NoK, (c) K6R, (d) Ac K6R, (e) K6O and (f) Ac K6O showing the differences in their profiles. Data plotted using *CIUSuite2*.

5.3.1 Ubiquitin mutant (K6A)

The K6A mutant was tested using the same ramping method applied for the $[M+5H]^{5+}$ ion. But this time less energy was applied for the $[M+6H]^{6+}$ ion, the CCS shown in Figure 5-8 for the drift traces obtained from applying energy from ($E_{lab} = 12-144$ eV) with set of increase 12 energy interval, the resulted plots revealed similar data as the $[M+5H]^{5+}$ ion, with the K6A mutant unfolding more than WT-Ub, but with slightly more resolved conformers, at $E_{lab} = 12$ eV the drift trace for K6A was less stable than the WT-Ub.

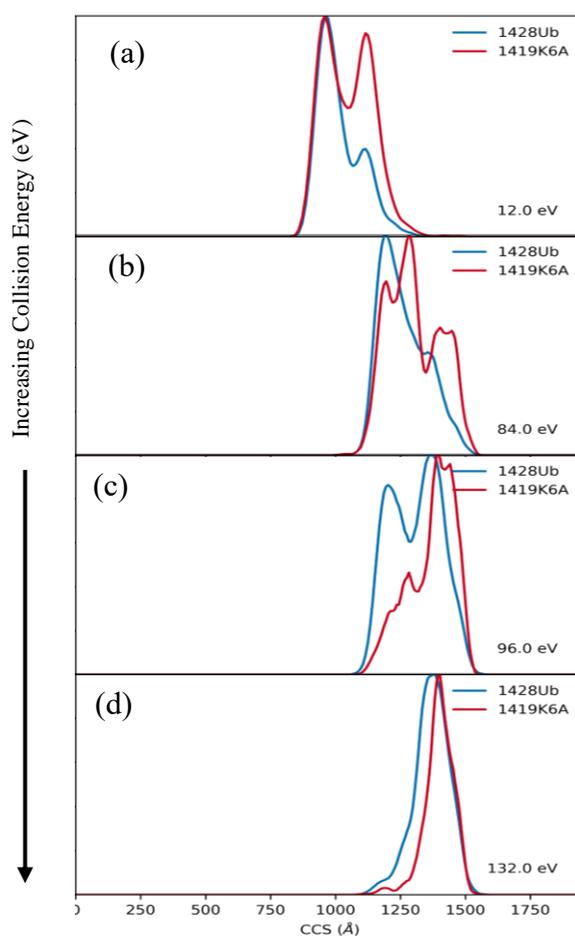


Figure 5-8. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+6H]^{6+}$ mobility drift traces obtained for 1428 WT-Ub (blue) vs K6A(red). (a) the more compact structure at trap energy 12 eV where the first conformation is shown. (b),(c) Increasing the voltage to from 84 to 96 eV. (d) show the final conformation.

5.3.2 Ubiquitin mutant (K6O)

The K6O mutant was found to be more stable than the WT-Ub, as was the case with the $[M+5H]^{5+}$ ion. When K6O was acetylated and compared to the K6O mutant it showed insignificant difference in CCS. Figure 5-9 shows detailed collision voltage for the $[M+6H]^{6+}$ charge state, with ion mobility drift traces obtained for K6O vs Ac K6O. This result is in contrast to that obtained for the $[M+5H]^{5+}$ where acetylation led to clear destabilisation. It implies that the nature of position K6 in Ub is less important when the charge changes from 5+ to 6+, this indicates the overall charge on the protein is vital to the structure.

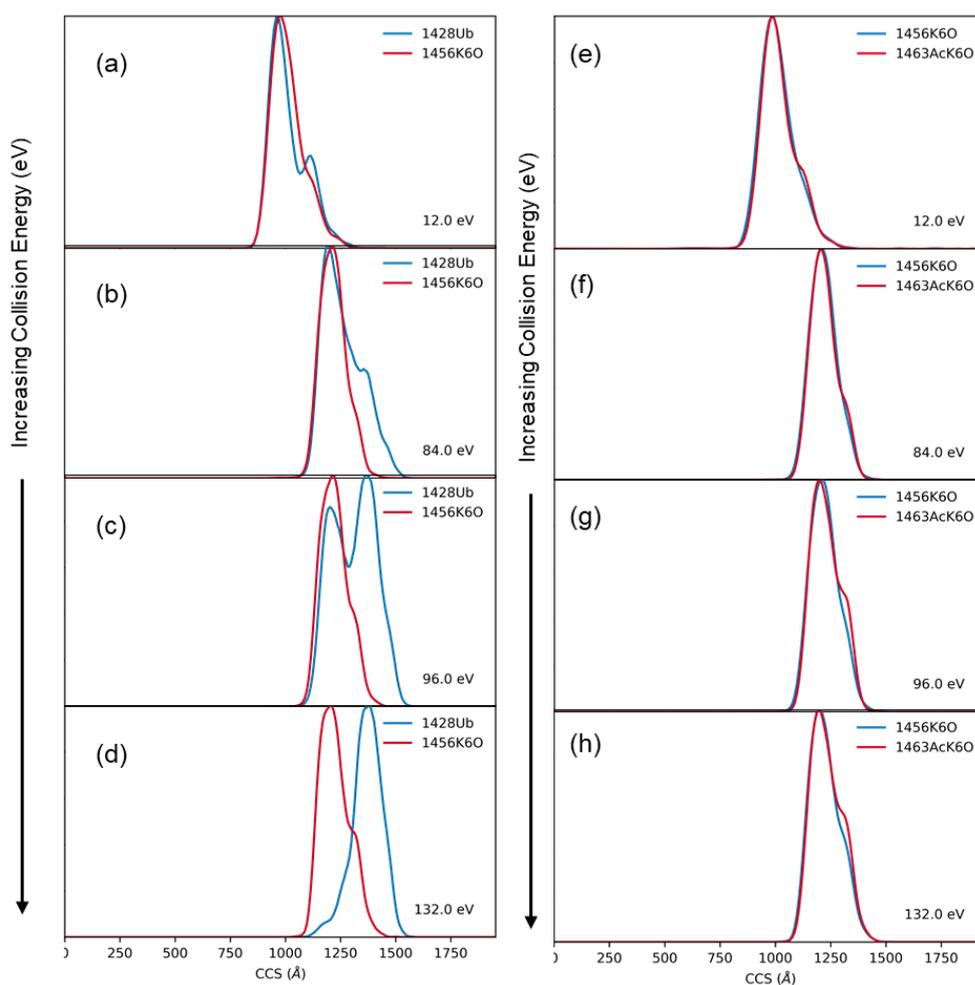


Figure 5-9. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+6H]^{6+}$ mobility drift traces obtained for 1428 WT-Ub (blue) vs 1456 K6O (red). (a) the more compact structure at trap energy 12 eV where the first conformation is shown. (b), (c) Increasing the voltage to from 84 to 96 eV. (d) show the final conformation at 132 eV.

5.3.3 Ubiquitin mutant (K6R)

The K6R mutant was shown to be more stable than the WT-Ub which matches the result from $[M+5H]^{5+}$ ion. A detailed collision voltage ramp was conducted with a single charge state ion $[M+6H]^{6+}$ and mobility drift traces (see Figure 5-10) obtained for 1432 K6R (blue) vs 1439 Ac K6R (red). The first figure shows the CIU of K6R compared to WT-Ub at a range of energies, the first plot shows the unfolding at lowest energy $E_{lab} = 12$ eV, by increasing the unfolding energy the two conformers start to unfold differently, at $E_{lab} = 96$ eV the WT-Ub appear to be less stable than the K6R mutant.

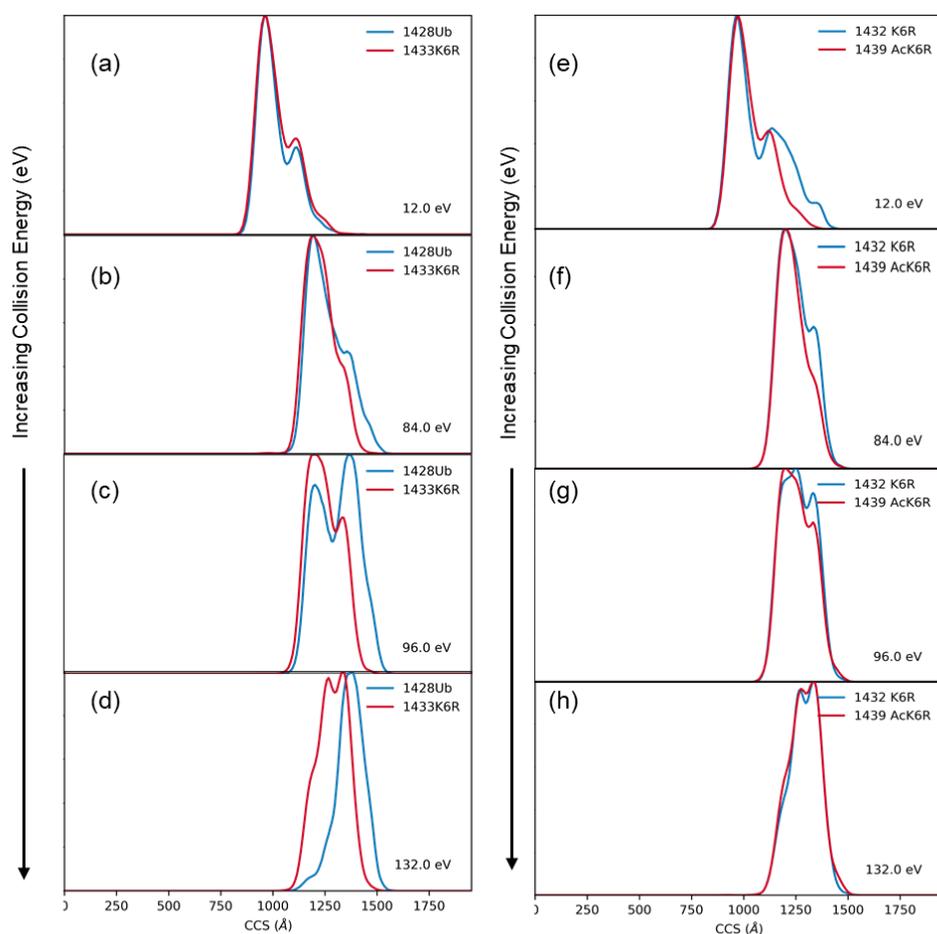


Figure 5-10. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+6H]^{6+}$ mobility drift traces obtained for 1428 WT-Ub (blue) vs 1433 K6R (red). (a) the more compact structure at trap energy 12 eV where the first conformation is shown. (b), (c) increasing the voltage to from 84 to 96 eV. (d) show the final conformation at 132 eV.

5.3.4 Ubiquitin mutant (NoK)

The NoK mutant had no lysine residues since all of the lysines had been changed to arginine. As a result, this protein is more stable than in any other mutant. When the *CCS* in Figure 5-11 is compared to the WT-Ub confirmation *CCS*, it appears that providing energy causes the WT-Ub to unfold whereas the NoK mutant does not. Because arginine has greater intramolecular interactions with other areas of the protein, the inclusion of an arginine residue may lead to protein stabilisation. This result was in line with that for the $[M+5H]^{5+}$ ion.

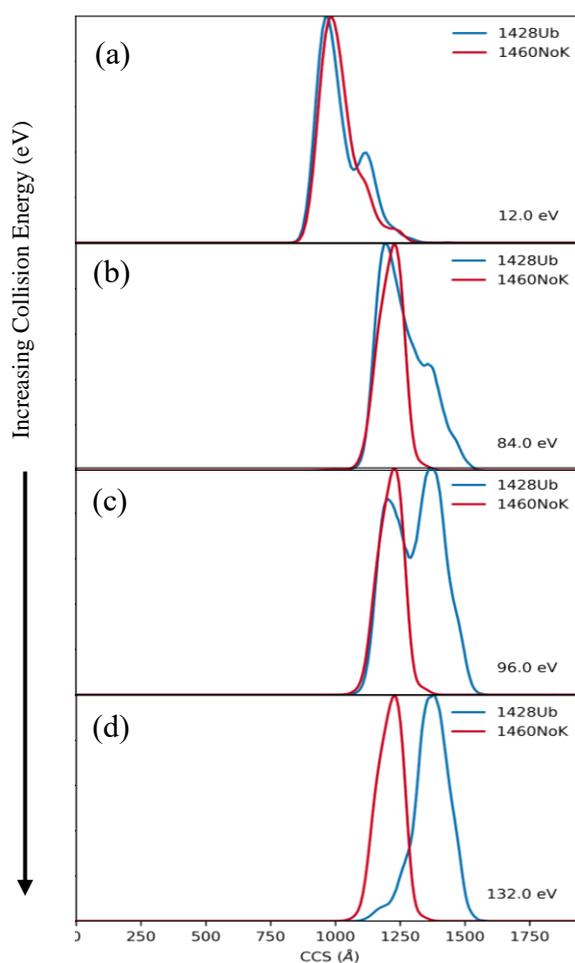


Figure 5-11. A detailed collision voltage ramp is conducted with a single charge state Ion $[M+6H]^{6+}$ mobility drift traces obtained for 1428 WT-Ub (blue) vs 1460 NoK (red). (a) the more compact structure at trap energy 12 eV where the first conformation is shown. (b), (c) increasing the voltage to from 84 to 96 eV. (d) show the final conformation at 132 eV.

5.4 Conclusions

A CIU profiling method was performed on a range of mutants designed to probe the effect of modification at K6 on the stability of compact, folded Ub ions in the gas-phase. More than a hundred drift trace chromatograms were obtained, to simplify and show a visual movement of the unfolding profile an alternative plot was used to illustrate the similarity and different between all the sample. *Magicplot* was used to identify and plot Gaussians to measure the area for each population in the CIU drift traces.

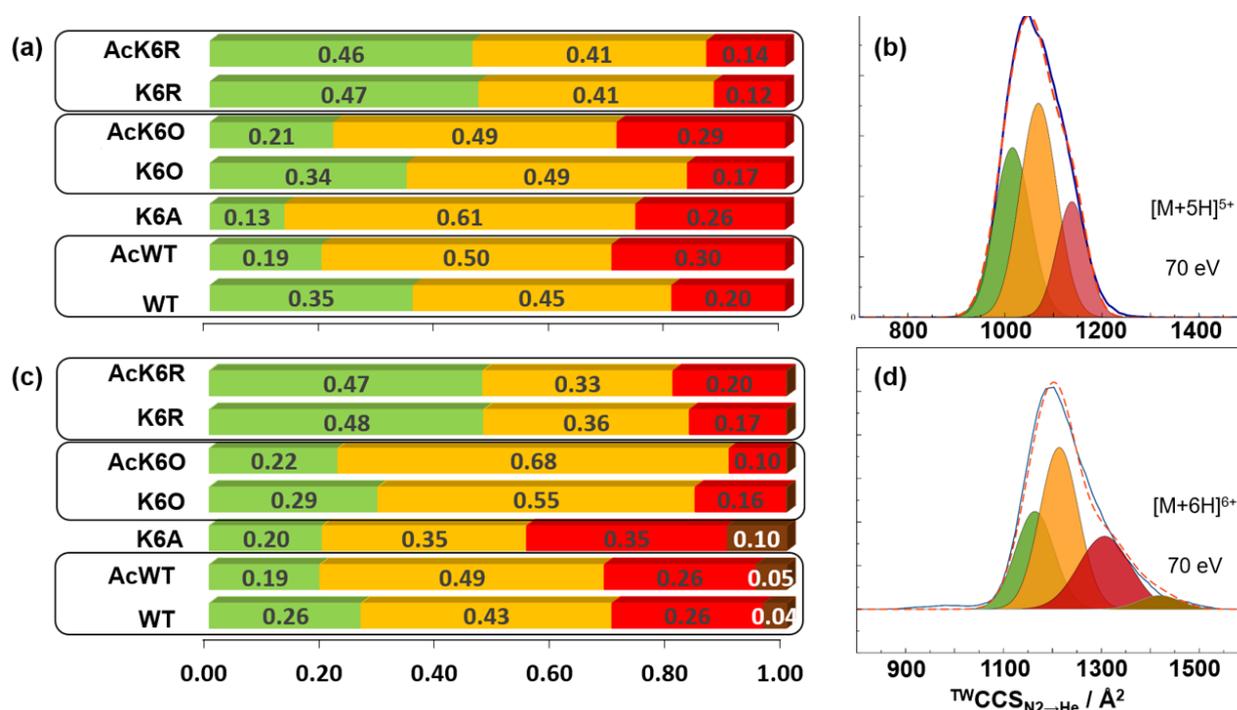


Figure 5-12. Summary of CIU unfolding of Ub ions using Gaussian fitting to IM drift traces. Ub variants are: (WT), (AcWT), (K6A), (K6O), (AcK6O), (K6R) and (AcK6R). The relative areas of the Gaussians for $[M+5H]^{5+}$ ions at 70 eV and $[M+6H]^{6+}$ ions at 72 eV are shown in (a) and (c) respectively. The Gaussians used to fit each IM drift trace at those energies are illustrated in (b) and (d) respectively using WT Ub in each case. The experimental trace is shown as a solid blue line and the sum of the Gaussians as a dotted red line.

Three major population were observed along the drift trace, and their area was calculated by gaussian fitting and plotted as shown in Figure 5-12 , The top section summarizes the $[M+5H]^{5+}$ ion result, which includes the WT-Ub, mutant, and chemically modified protein, while the bottom section displays the $[M+6H]^{6+}$ ion data. The green bar represents the initial conformers, the orange bar represents intermediate conformers, and the red bar represents the final unfolded structure.

The acetylated K6R and K6R show similar unfolding confirmation in the $[M+5H]^{5+}$ and $[M+6H]^{6+}$ ions. The total area under the gaussian is similar for the folded conformers between 0.45 and 0.46, the intermediate conformers as well as the unfolded conformers is showing similarity for both of the ion, this could be because the lysine in K6 position has been modified to arginine, which means the acetylation goes to another lysine location and causes no difference in the unfolding.

The acetylated mutant of the AcK6O mutant appears to be less stable than the unmodified protein, possibly because the acetylation occurs on the K6 position and causes the protein to be disrupted. When looking at the data in the middle row we can see that the total area of AcK6O by measuring the gaussian is 0.21 for the most compact conformers. The area of 0.61 for the intermediate conformers and 0.26 for the final unfolded conformers. However, for the K6O there is increase in the intensity of the compact conformer to 0.34, while the intermediate conformers increase to 0.49, and substantial reduction in the unfolded conformers from 0.23 to 0.17. for the $[M+5H]^{5+}$ at $E_{lab}=70$ eV.

For the final row, the result from the $[M+5H]^{5+}$ and the $[M+6H]^{6+}$ ion have fairly similar pattern in term of the unfolding. The total area of WT-Ub by measuring the gaussian is 0.35 for the most compact conformers. The area of 0.45 for the intermediate conformers and 0.20 for the final unfolded conformers. However, for the acetylated Ub there is reduction in the

intensity of the compact conformer to 0.19, while the intermediate conformers increase to 0.50, and significant increase in the unfolded conformers from 0.20 to 0.30. therefore there is more unfolding in the acetylated Ub than the WT-Ub at $E_{lab}=70$ eV.

The data for K6A mutant is shown great match to the AcUb in the first conformer between 0.20 to 0.19, the the intermediate conformers seem to be greater for the AcUb, while the unfolded conformers have higher value for the K6A $E_{lab}=70$ eV. Taken together, these results confirm the importance of K6 for the stabilisation of Ub's compact structure in the gas-phase, but do show some subtle differences between the $[M+5H]^{5+}$ and $[M+6H]^{6+}$ ions. That a single residue within a 76 residue protein can have such a dramatic effect on structural stability is extremely interesting, and will be the subject of further investigation in Chapter 6.

Chapter 6

Molecular dynamics and charge distribution

6 Molecular dynamics and charge distribution

6.1 Introduction

Molecular modelling and molecular dynamics simulations have become valuable computational approaches for studying the structure–function relationship of proteins. To gain atomic level resolution insights into ubiquitin and the ubiquitin signalling system, researchers used molecular modelling and simulations as a computational microscope (McCammon et al., 1977). These simulations allow researchers to explore the conformational energy landscape accessible to biomolecules and to connect the dots between a protein's three-dimensional structure and its dynamics (Fox & Kollman, 1996).

Large molecules like protein are sustained by many intramolecular forces. These forces are the result of a complicated system of numerous hydrogen bonds, salt bridges, attractions between polar groups, and other interactions to reach an optimum conformation to achieve protein stability and intended function (Fox & Kollman, 1996).

The objective of the first MD simulations of Ub was to investigate its dynamic character. These simulations were only a few nanoseconds long, used a combination explicit-implicit water treatment, or even concentrated on just one short peptide (Fox & Kollman, 1996). These research, as well as others, looked at the protein's natural state dynamics, examining its conformational space and accounting for its inherent oscillations (Kony, Hünenberger, & van Gunsteren, 2007; Marianayagam & Jackson, 2005; Wlodarski & Zagrovic, 2009).

These experiments paved the door for far more detailed and comprehensive MD studies that looked at the folding process from numerous angles, with the goal of generating the folding (and unfolding) trajectory.

Previous simulation studies on ubiquitin folding have investigated at the folding mechanism of human ubiquitin combining a physically realistic model and extensive MD simulations. Simulations suggest a folding mechanism that matches a significant body of experimental data (Piana, Lindorff-Larsen, & Shaw, 2013).

The fast-collapse phase preceding the folding process has been analysed by rapid low-temperature quenching of unfolded-state structures generated in high-temperature simulations using atomistic force fields and simulations on the nanosecond timescale (Alonso & Daggett, 1998).

Experiments and simulations on ubiquitin and other well-engineered fast-folding proteins appear to be able to provide key details regarding protein folding mechanisms in general, including for naturally occurring proteins (Lindorff-Larsen, Piana, Dror, & Shaw, 2011).

Solvent molecules (such as water) do not interact with the hydrophobic amino acid sidechains of a protein, resulting in an unfavourable decrease in entropy, which causes the protein to fold to its globally native form while hydrophilic and polar side chains tend to be on the surface. However, during ESI proteins experience major change in the surrounding because in water the hydrophobic effect stabilizes the structure but in the absence of water the protein bonds are no longer stable (T. Wyttbach et al., 2014).

In addition, the ESI undoubtedly leads to changes to the protonation states of polar sites, and ESI-MS work on proteins indicated a close correlation between the protein charge state distribution in the mass spectrum and the ESI solution conditions (Bakhtiari & Konermann, 2019). To further understand the unfolding of protein and to eliminate any limitation of the IMS, molecular dynamics theoretical simulations were implemented and compared to the experimental *CCS* result, thereby presenting a comprehensive picture of the mechanism of unfolding and the effect of chemical modification to the structure of protein.

6.2 Experimental CCS equated to theoretical CCS

Prior work by Konermann et al. (Bakhtiari & Konermann, 2019) discussed the mechanism of MD simulation and the charge distribution on the Ub model. These findings were used by Dr David Robinson, of Nottingham Trent University, to perform MD simulations on Ub reported in this chapter.

The Ub model that was used in the simulation was obtained from the Protein Data Bank [1UBQ] and was overall neutral in charge, so for the purpose of simulation a charge was added to the structure to achieve the $[M+5H]^{5+}$ ion. This was done based on Konermann et al. work by deprotonating the E18 while leaving all K and R side chain protonated, also all the D and E sidechain and c-terminus deprotonated. In the neutral Ub structure all K and R sidechains were protonated in addition to the N-terminus and all D and E sidechains as well as the C-terminus were deprotonated.

The $[M+6H]^{6+}$ ion was created by deprotonating K63 and reprotonating the sidechains of residues E18, E24, D32, D39, E51, D58, and E64. The K6A Ub mutant includes alanine instead of lysine, which means that the position is not protonated, to achieve a $[M+5H]^{5+}$ charged MD structure the charge was added by neutralising E18.

The MD simulation was replicated three times to produce a repeatable result. The simulation ran first to mimic the low gas-phase structure by applying 298 K and then produced again at 750 K at (0.5 ns and 9 ns) to mimic the unfolding heated structure. This clearly demonstrates that the timescale chosen has an impact, and that longer runs are not necessary. It is conceivable for the interaction between K6 and H68 to become ineffective for relatively short periods of time if the conditions are right (within 0.5 ns).

The theoretical *CCS*s of output trajectories were calculated and compared to the experimental *CCS*s, MD run was produced in two stages: an equilibration run was carried out for approximately 0.5 ns, followed by a production run for approximately 8 ns, as previously stated. All simulations were run in triplicate to ensure accuracy. For the resulting structures, the collisional cross sections *CCS* were calculated using the corrected PA method, in which the *CCS* determined by projection approximation (PA) is converted into the apparent trajectory model (TM) value using a constant scaling factor (Benesch & Ruotolo, 2011; Sokratous et al., 2012).

Figure 6-1 shows the experimental drift traces against the corresponding theoretical drift trace, the first column represents the experimental drift trace at 30 eV at peak maximum value of 975 Å² which was similar to the simulated drift trace value at 950 Å², the gas-phase structure of Ub shown at the plot was taken from the MD at 950 Å², it shows the protein is globally folded into a basic secondary and tertiary structure.

The second column shows the plots for the 70 eV CIU drift trace and the MD simulation at 750 K, the *CCS* peak maximum for the experimental trace 1038 Å² which is almost similar to the simulation drift trace at equilibration phase at 1039 Å², the structure shows the bond started to break and the native secondary structure started to unfold.

The third column shows the drift trace at 110 eV experimental versus 750 K MD simulation drift trace at production step, by comparing the initial line to the end, the peak of the trace had reached maximum of 1083 Å² *CCS* while 1099 Å² for the theoretical spectra. looking at the final structure, the Ub has lost all its native folding, the hydrophobic residues have migrated to the surface and the polar sidechain inverted inward, the bonds that hold the protein have all been broken.

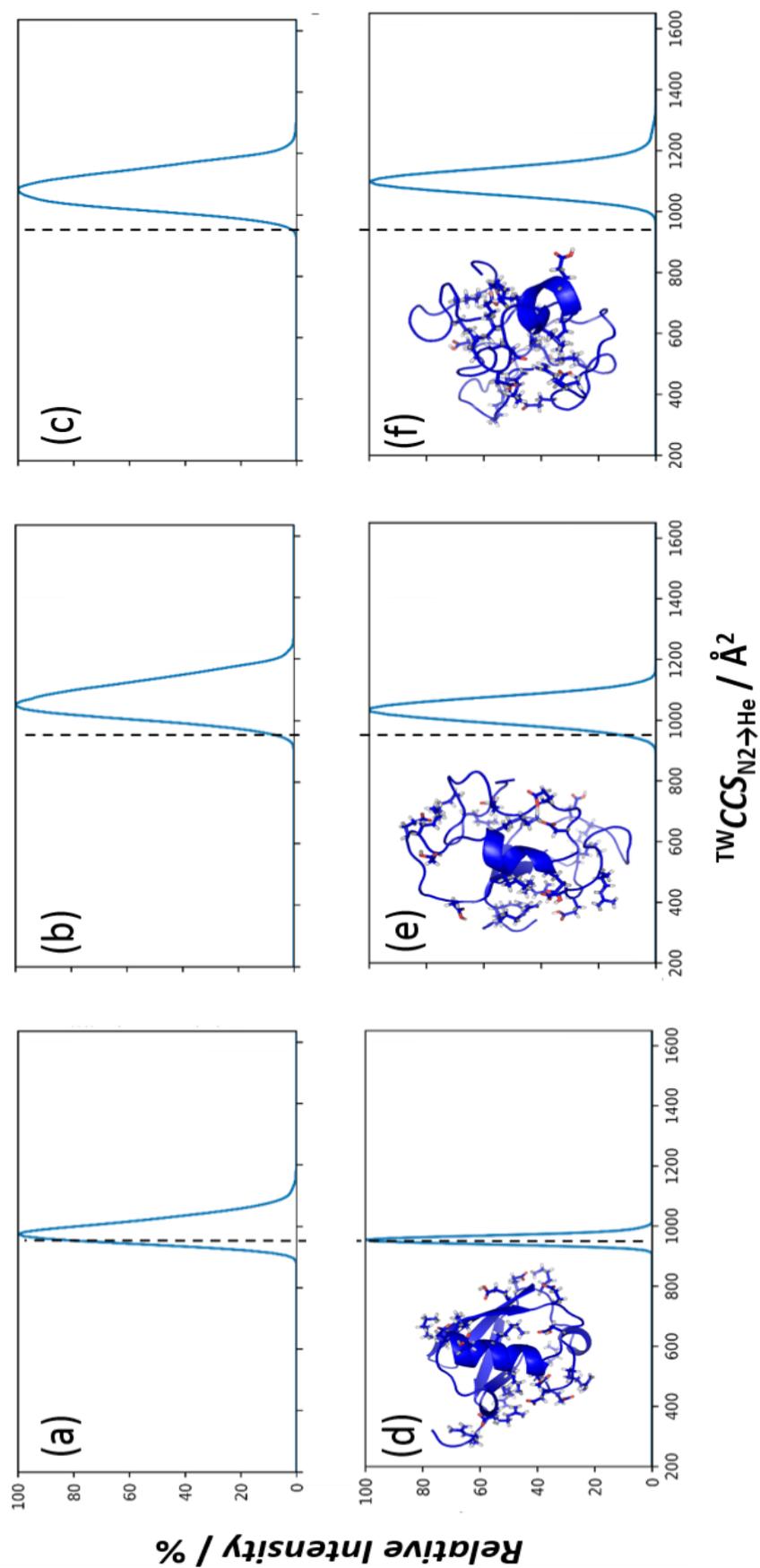


Figure 6-1. Experimental CIU drift traces for WT-Ub $[M+5H]^{5+}$ ion at (a) 30 V, (b) 70 V and (c) 110 V, and MD simulations of the WT-Ub $[M+5H]^{5+}$ charge state at (d) 298 K (0.5 ns), (e) 750 K (0.5 ns) and (f) 750 K (9 ns). Inserts in (d)-(f) show MD structures parallel to the most abundant CCS values, charged residues K, R, D, and E shown as sticks.

6.3 Lysine and histidine interaction

RMSD (root mean square deviation) was used to measure the scalar distance for the proteins throughout the trajectory, in which all protein frames are first aligned on the reference frame backbone and then the RMSD is calculated based on the C alpha or sidechain. Figure 6-2 represents the RMSD plot of comparison between analogous $[M+5H]^{5+}$ ion, $[M+6H]^{6+}$ ion, neutral WT-Ub and $[M+5H]^{5+}$ K6A. The plot shows the three structures as they unfold at 750 K in the MD simulation at the equilibration step, as the difference in fluctuation of the coordination can be explained by the charged transposition on K6.

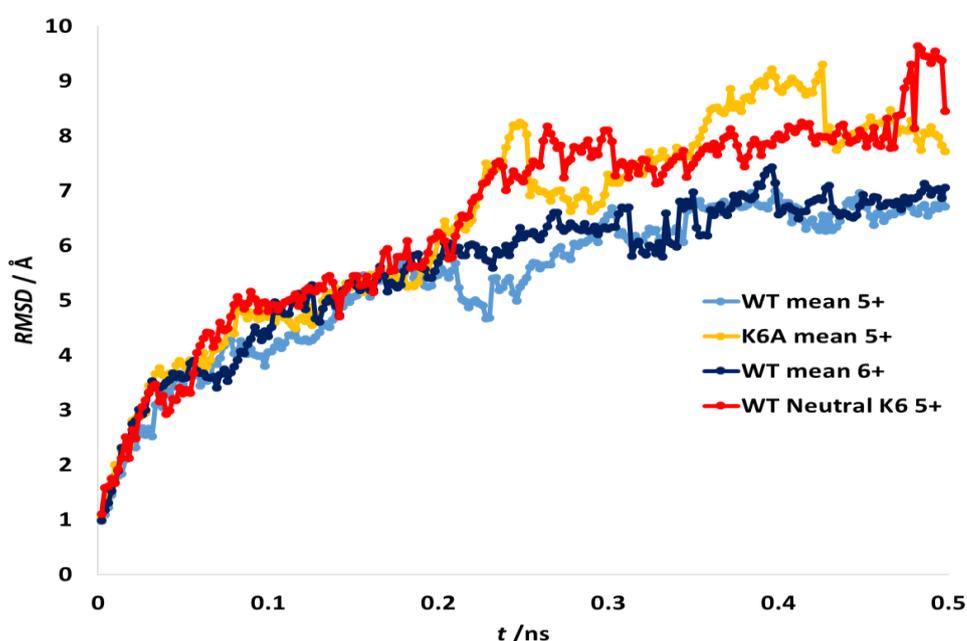


Figure 6-2. Plot of all atom RMSD from initial structures over 0.5 ns MD simulation. Data are the mean values from 3 independent simulations.

When the WT-Ub K6 position was neutralized, the structure became more unstable and unfolded significantly more than the native structure. This was shown as well in K6A $[M+5H]^{5+}$ ion, where the K6 position was neutralized by mutation. The WT-Ub $[M+6H]^{6+}$ ion shows slight deviation from the original structure, but not as strong as the changes done on the K6. This further shows the importance of the K6 position for the stabilization of the Ub structure.

Monitoring the protein's RMSD throughout the simulation can reveal the extent of structural change, as well as indicate if the simulation has equilibrated and its fluctuations are around some thermal average changes on the order of 1 to 3 angstroms, which are perfectly acceptable for small globular proteins like Ub. However, it is critical that the simulation converges or that the RMSD values stable at or around a fixed value, as is the case above, where the simulation is beginning to reach a plateau. Figure 6-3 show the mean RMSD values for each protein type. We can see from these values that the runs are concordant with one another.

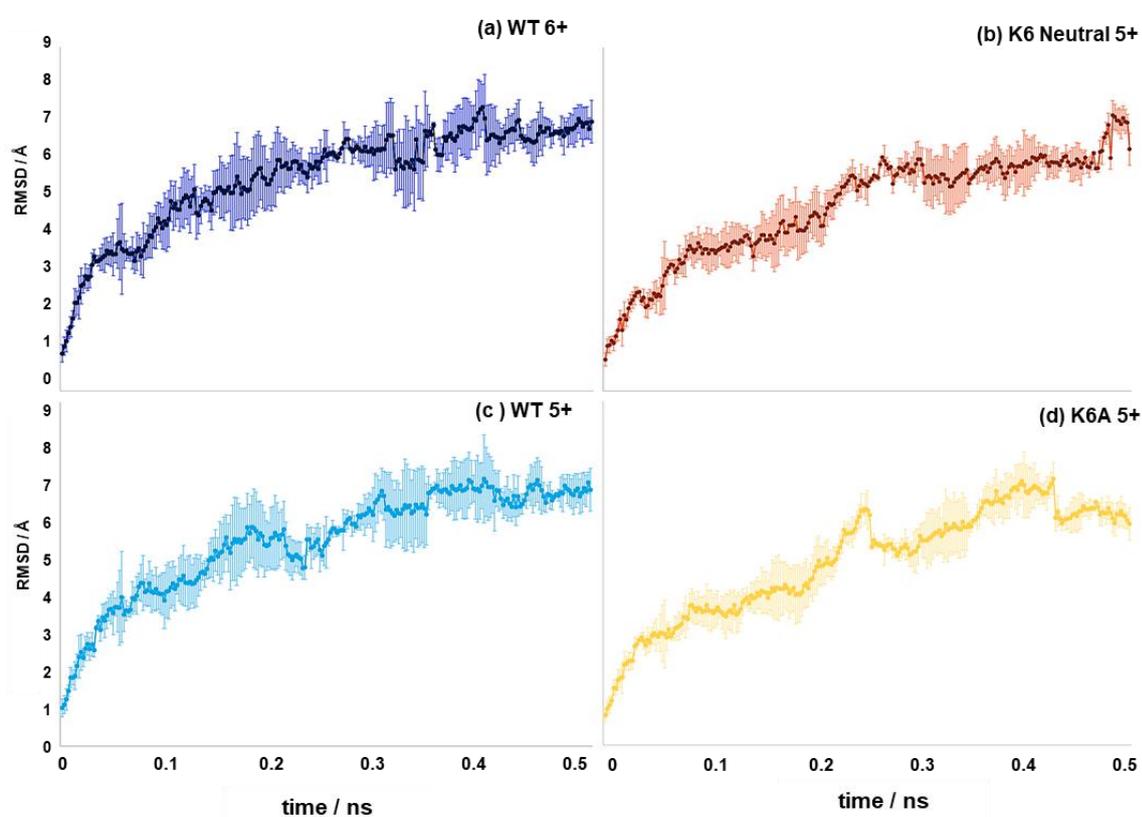


Figure 6-3 Plot of all atom RMSD from initial structures over 0.5 ns MD simulation. Data are the mean values from 3 independent simulations, (a) Shows the RMSD variation between the three runs for WT 6+ , (b) Shows the RMSD variation between the three runs for K6 5+, (c)) Shows the RMSD variation between the three runs for WT 5+ and (d)) Shows the RMSD variation between the three runs for K6A 5+.

The MD simulation results in an MD structure for each protein every 5 ps. By analysing the equilibration stage, it was clear that the difference in the distance between K6 and H68 in case of the WT-Ub and A6 and H68 in case of the K6A mutant could be observed. In both cases, the residue at position 6 is close to H68 in the starting structure for the simulations, but as the simulation proceeded, differences became apparent.

The structure shown in the Figure 6-4 taken from the MD simulation reveals the distance between the protonated K6 and H68 at several time frames, the structure obtained from the equilibrium step of the MD simulation for the WT-Ub [M+6H]⁶⁺ ion and [M+5H]⁵⁺ ion, K6A [M+5H]⁵⁺ ion and K6 neutral [M+5H]⁵⁺ ion. WT-Ub were extracted and the distance between the residue was measured.

First the distance between K6 and H68 in WT-Ub [M+5H]⁵⁺ ion is shown to be 6.4 Å and after 0.25 ns simulation the distance show no rapid change but the structure unfolded, however after 0.5 ns the distance grow to be 11.1 Å and more unfolded resulted. For the WT-Ub [M+6H]⁶⁺ The initial structure had a distance of 7.4 Å, but after 0.25 ns, the structure became even more destabilized, and the distance increased. And the final structure had 13.0 Å, This indicates that the 6+ ion is less stable than the 5+ ion, but not significantly so.

For the next two proteins K6A [M+5H]⁵⁺ and K6 neutral [M+5H]⁵⁺, there was a noticeable difference compared the previous structures, Primarily, it appears that the distance between the two proteins initial structures is the same for both, with the first distance being 13.8 Å, which is significantly bigger than the first distance for WT-Ub [M+6H]⁶⁺ and [M+5H]⁵⁺. After 0.25 ns of simulation, the unfolding of K6A has increased significantly, reaching 27.1 Å while 23.4 Å for the K6 neutral, this drastic increase in the distance between the residues results in a significant loss of structural stability, which is consistent with the experimental results.

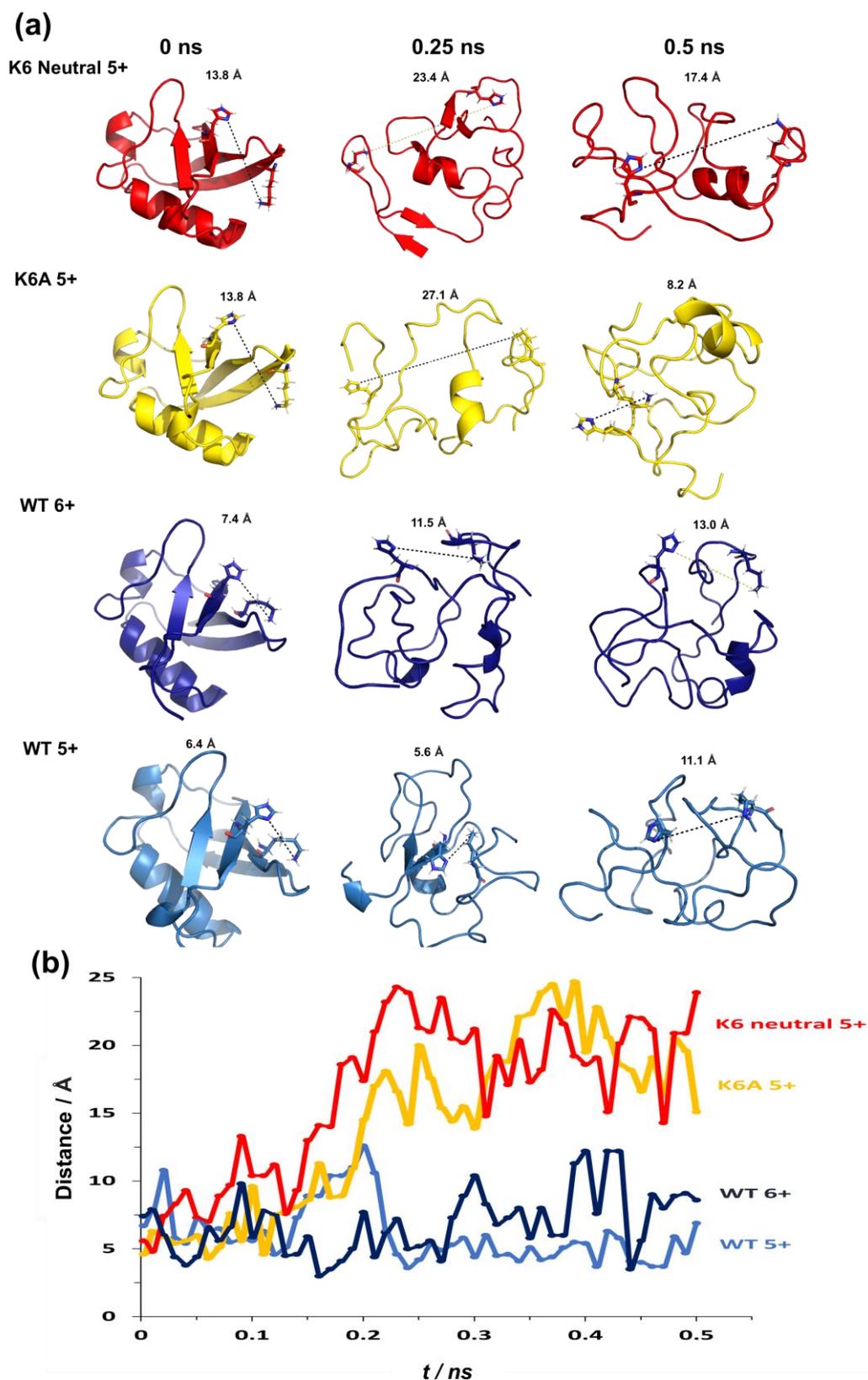


Figure 6-4. (a) Structures for WT-Ub [M+5H]⁵⁺ and WT-Ub [M+6H]⁶⁺ Ub, WT-Ub [M+5H]⁵⁺ with neutral K6, and K6A [M+5H]⁵⁺ showing increased unfolding of the N-terminal β -hairpin at 0 ns, 0.25 ns and 0.5 ns. (b) Plot of K6-H68 distance over the 0.5 ns equilibration stage of MD simulations showing increased unfolding for protein. Data represent the mean of 3 repeat simulations.

Figure 6-4 (b) shows the average inter-residue K6 or A6-H68 distance \AA against the time in ns for the four different charge states of these Ub variants. This MD simulation equilibrium stage was chosen because it represents the early stages of compact structure unfolding. It shows similar behaviour to the RMSD plots in Figure 6-2.

These results demonstrate that the WT-Ub $[\text{M}+5\text{H}]^{5+}$ ion and WT-Ub $[\text{M}+6\text{H}]^{6+}$ ion, where the K6 position is protonated, have the inter-residue distance of the simulation of about 5-8 \AA . However, when neutralizing the position for the K6A mutant $[\text{M}+5\text{H}]^{5+}$ ion, and WT-ub $[\text{M}+5\text{H}]^{5+}$ ion the distances is shown to be larger and more unstable. This indicates that the intermolecular force between the two residues is important for both function and structure, and the significance of this residue is in the unfolding process.

These findings suggest to the importance of an interaction between protonated K6 and H68 in the gas-phase maintenance of the compact fold of Ub. K6 modification or direct neutralization appears to decrease this connection, allowing Ub's N-terminal-hairpin to unfold.

6.4 Conclusions

For researching the structure–function relationship of Ub proteins, molecular modelling and molecular dynamics simulations have proven useful. Molecular dynamics and charge distribution studies has provided comprehensive picture of the structure and dynamics of gas-phase Ub ions, such as those generated by ESI. In this chapter molecular dynamics simulations were used to investigate protein unfolding and the influence of chemical modifications on protein structure.

Using information from previous chapters, we were able to obtain the unfolding profiles for several proteins. In order to compare the experimental results to the theoretical results, MD simulations were first performed at 298 K to replicate the low gas-phase structure, and then at 750 K for (0.25 seconds and 9 seconds) to replicate the unfolding heated structure.

In the gas phase the interaction from hydrophobic amino acid sidechains, entropy decreases, causing the protein to fold to its globally natural form, with hydrophilic and polar side chains on the surface. The experimental drift traces were compared to the matching theoretical drift trace, and the first result showed that the gas-phase structure of Ub protein is globally folded into a basic secondary and tertiary structure, which is identical to the simulated drift trace value. The structure of the second result demonstrates that the native secondary structure began to unfold.

Finally, The Ub has lost all of its original folding, the hydrophobic residues have moved to the surface, the polar sidechain has reversed inward, and all of the protein's linkages have been broken.

The hydrophobic effect stabilizes the structure in water, but without water this effect is removed. Despite this, ESI-MS study on proteins revealed a close correlation between the protein charge state distribution in the mass spectrum and the ESI solution conditions.

The results presented here indicate the importance of interaction between protonated K6 and H68 in stabilisation of the compact fold of Ub in the gas-phase, when changes affect the K6 position in WT-Ub, for example, chemical modification or mutagenesis of the sequence it destabilises the structure by changing the charge the intermolecular forces affected which leads to unfolding of the N-terminal β -hairpin of Ub.

The combination of CIU data from modified Ub and MD modelling has allowed structural insights into gas-phase proteins to be resolved down to the level of single amino acid interactions, which was previously more difficult and time-consuming, as well as costing more.

Chapter 7

Ubiquitin-binding domain UBA2

7 Ubiquitin-binding domains UBA2

7.1 Introduction

The UBA domains of hHR23A is a good example to illustrate unfolding of small protein domains, The UBA domain was the first reported and well characterized ubiquitin binding domain from Rad23 and R23A (Chen et al., 2001).

UBA domains structure is made of three compact helix bundles, two of the helical, $\alpha 1$ and $\alpha 3$ helices (Yang, 2020). The structure of UBA is well understood and it contains several position accessible for chemical modification, the DNA sequence of UBA was available for site-directed mutagenesis so the study of several mutant was accessible.

7.2 Acetylation of UBA2

UBA2 protein was acetylated using acetic anhydride with molar equivalent of 150 to obtain modified ions with mass addition of 42 Da. The acetylation of UBA2 resulted in modified $[M+5H]^{5+}$ ion m/z 1153, it was fragmented using CID to show the b-ion and y-ion. The UBA2 sequence contain many lysine that can carry the acetylation QVTPQEKEAIERLKALGFPELVIAQYFACEKNENLAANFLLSQNFDDE , so full MS/MS spectrum was searched for any mass addition to be able to locate the position of acetylation.

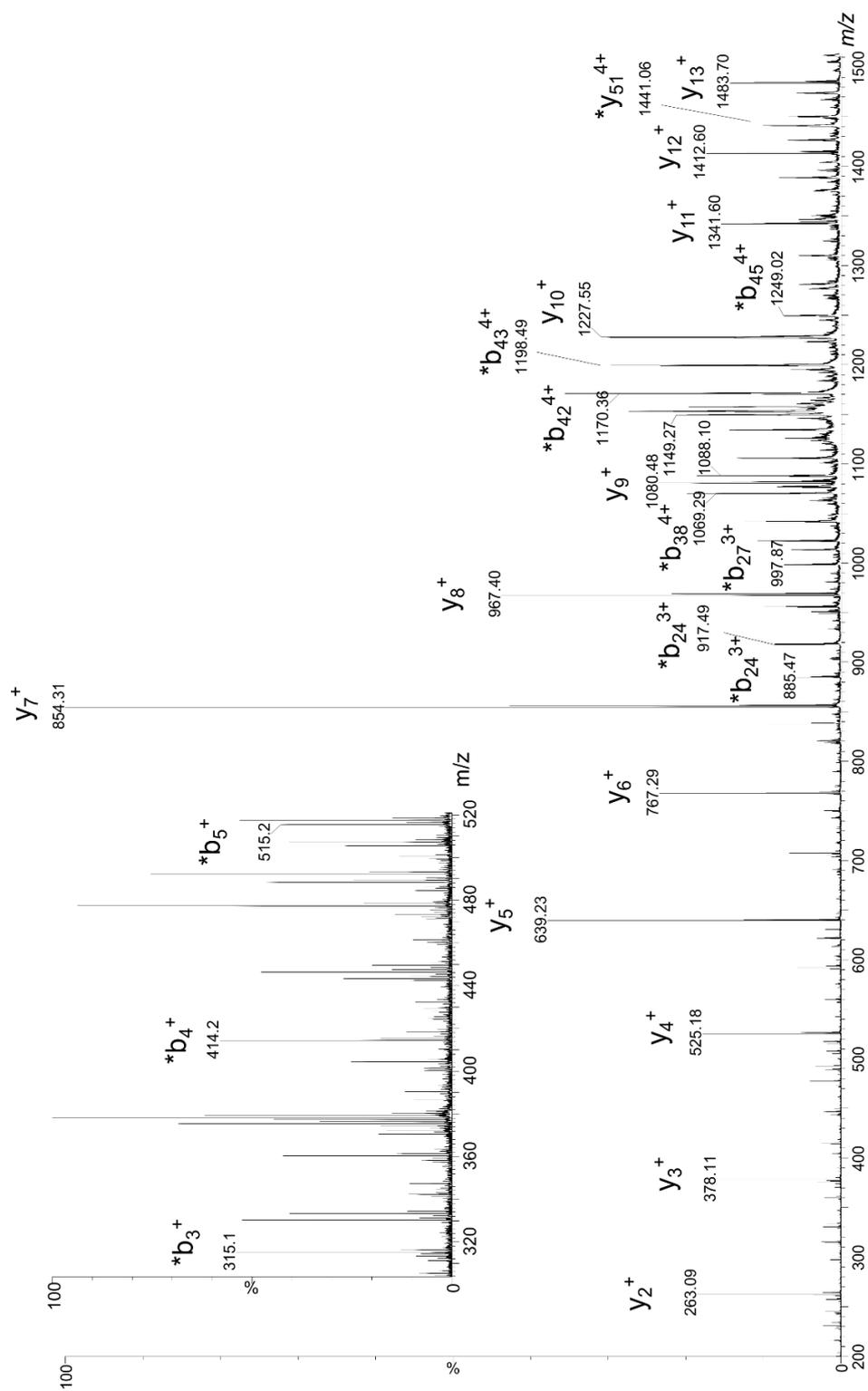


Figure 7-1 CID MS/MS spectrum of acetylated UBA following modification ($[M+5H]^{5+} = m/z 1153$) showing b-ions and y-ions, with the(*) indicates acetylated ions).

Studying the spectrum showed that b-ion at the N-terminus have been modified with the addition on 42 Da, Figure 7-1 shows the MS/MS spectra (insert) of acetylated b-ions $b3^+$ m/z 315.1, $b4^+$ m/z 414.2 and $b5^+$ m/z 515.2 suggesting that the location of acetylation could be on the N-terminus. However scanning through the spectrum acetylated b-ions $b24^{+3}$, $b25^{+3}$ at m/z 884.80 and 917.82 were found (below). Together with $b13^{+2}$, which showed no acetylation on K9, these data demonstrate that high percentage of acetylation is at K14. So next we use the Site-directed mutagenesis to produce K14A and K14H mutant to further understand the importance of K14.

7.2.1 CCS of Acetylated of UBA2

Collision induced unfolding of the UBA2 was obtained for the two ion $[M+4H]^{4+}$ the most abundant charge state and $[M+5H]^{5+}$. The conformation of the unmodified UBA2 for $[M+5H]^{5+}$ shows great difference in the CCS than the modified CCS. The unmodified conformation shows two species that unfold as the energy increases from 30 eV to 70 eV. At 110 eV ions reached their final conformation where no further unfolding was noticed. On the other hand, lower energy volt was enough to drive acetylated ions to their final unfolding conformation. This result indicates that the intramolecular interaction is destabilised by acetylation.

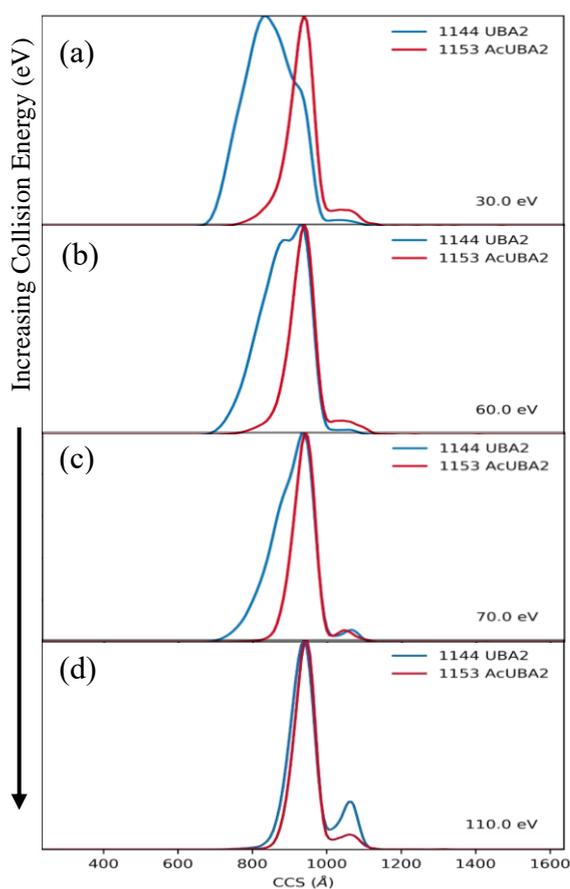


Figure 7-2 A detailed collision voltage ramp is conducted with a single charge state Ion $[M+5H]^{5+}$ mobility drift traces obtained for WT-UBA2 (blue) vs Ac-UBA2 (red).

It is noteworthy to indicate that the performance of CIU on the $[M+4H]^{4+}$ ion showed no significant change in the conformation of both modified and unmodified ions even by applying higher energy. Greater stabilization was achieved when removing one charge from the gas-phase structure of UBA2 compared to $[M+4H]^{4+}$. Consequently emphasizing on the importance of charge placement on protein's structure and function.

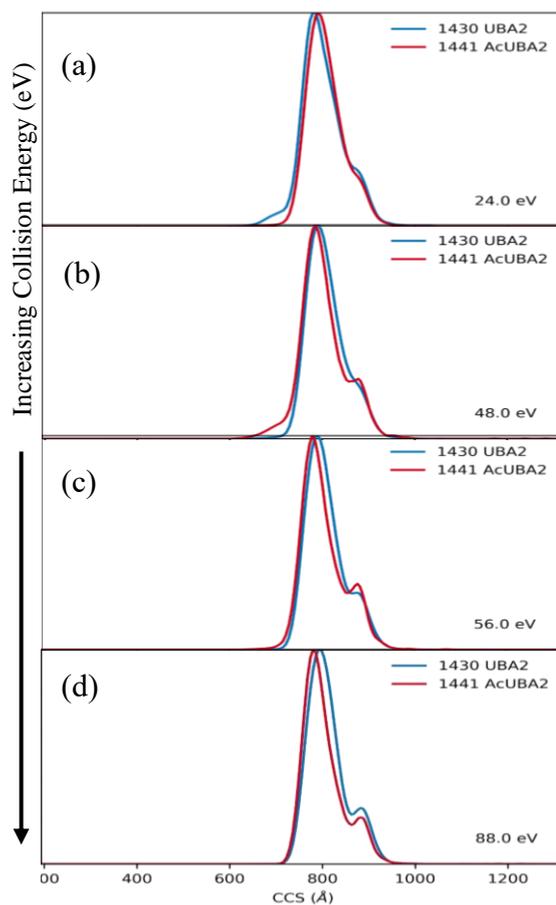


Figure 7-3 A detailed collision voltage ramp is conducted with a single charge state Ion $[M+4H]^{4+}$ mobility drift traces obtained for WT-UBA2 (blue) vs Ac-UBA2 (red).

7.3 Collision Induced Unfolding of UBA2 Mutant

7.3.1 CCS for UBA2 mutant K14A

The K14A mutant possessed an alanine residue instead of a lysine residue, which caused protein destabilization due to the change in single amino acid residue from protonated to neutral location. Figure 7-4 show the CIU of the two ion under several energies. The presence of an extra charge on the $[M+5H]^{5+}$ produced CIU with second species that unfold further than the wild type UBA2, however when the charge is taken the unfolding profile doesn't show any difference between the conformations.

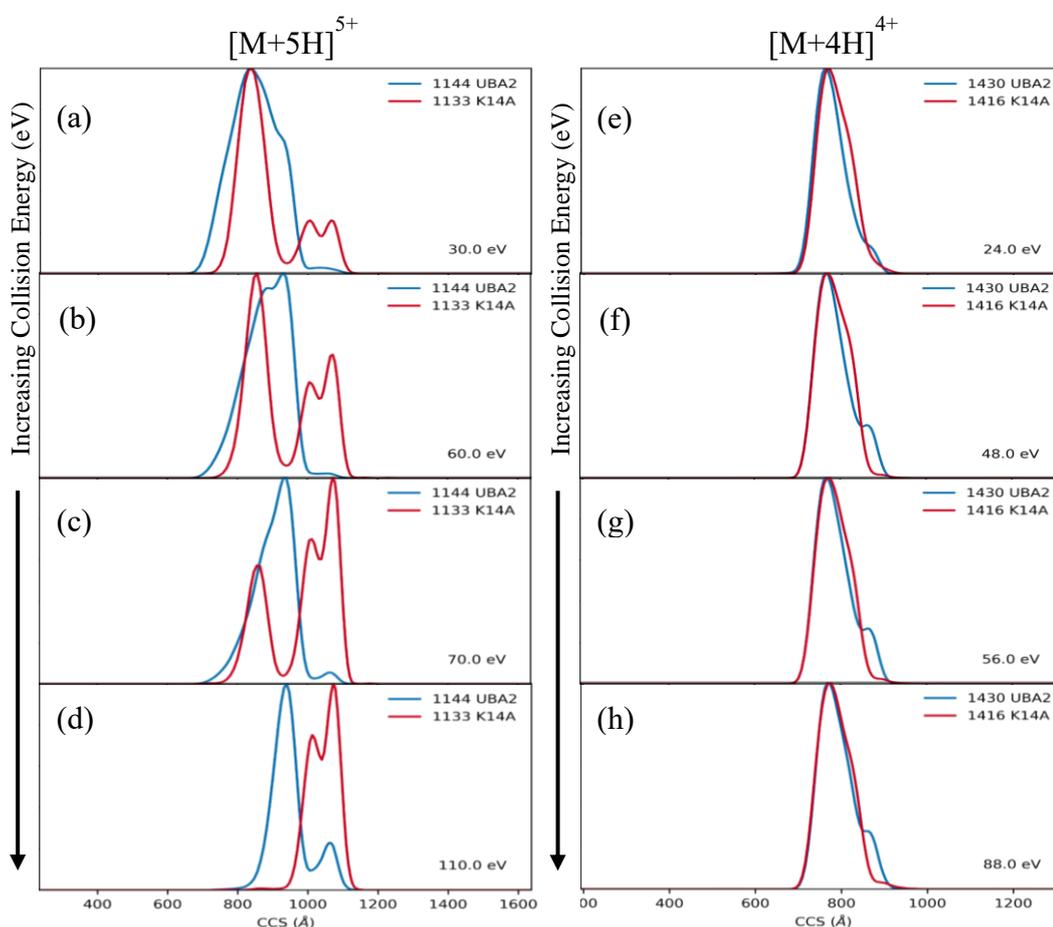


Figure 7-4 A detailed collision voltage ramp of the two ion $[M+5H]^{5+}$ and $[M+4H]^{4+}$ mobility drift traces obtained for WT-UBA2 (blue) vs K14A (red).

7.3.2 CCS for UBA2 mutant K14H

The CIU of K14H mutant did not affect its structure compared to the UBA2 as demonstrated in Figure 7-5. The CCS of both ions $[M+5H]^{5+}$ and $[M+4H]^{4+}$ showed no differences in conformational folding compared to the wild type UBA2. Even though $[M+5H]^{5+}$ unfolds at high energy, the destabilization of $[M+4H]^{4+}$ has not been affected. This could be due to the absence of an additional charge on the overall structure of protein.

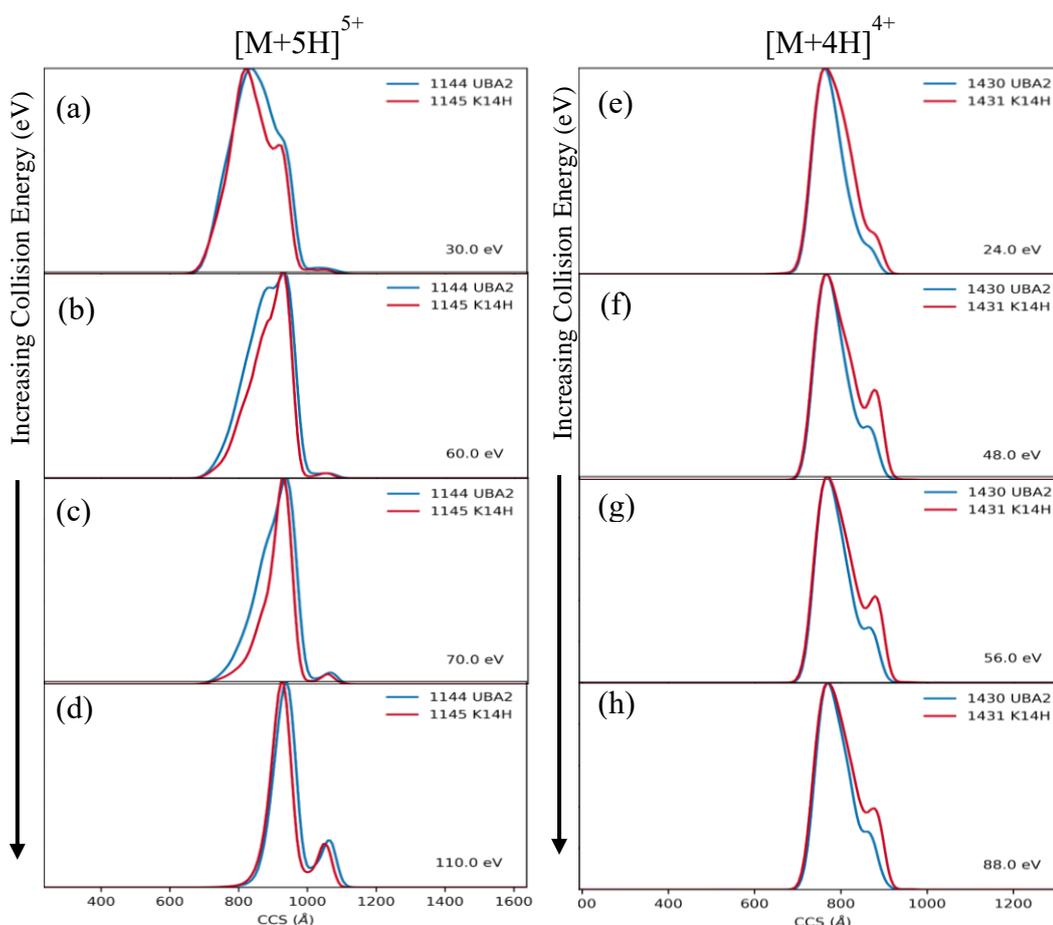


Figure 7-5 A detailed collision voltage ramp of the two ion $[M+5H]^{5+}$ and $[M+4H]^{4+}$ mobility drift traces obtained for WT-UBA2 (blue) vs K14H (red).

7.4 Conclusions

UBA2 protein confers a different structure due to the ionic interaction between the N-Terminus and different amino acids. By adding an acetyl group to the N-terminus, the intermolecular

interaction breaks and the protein unfolds, forming a stable conformation that does not unfold further with increasing energy.

According to the MS/MS results, there is a high chance that the acetylation occurred on K14; thus, by creating K14A mutants with the K replaced with A, the position of the K was shifted to neutral rather than protonated, resulting in increased destabilisation of the overall structure. However, the second mutant K14H, in which the position K has been altered to H, has no effect on the structure in comparison to wild type UBA2.

The unfolding of the $[M+5H]^{5+}$ ion varies significantly as the energy increases, however there is no difference between the variants of ~~in~~ the $[M+4H]^{4+}$ data. This implies that additional charge, acetylation and mutations alter the gas-phase structure. Thus, the excess proton most likely goes to lysine at position K14.

Conclusions and Outlook

8 Conclusions and Outlook

This work demonstrated the importance of applying ion mobility-mass spectrometry (IM-MS) methods to investigate protein structure with the support of chemical and genetic modifications. IM-MS and CIU were used to determine the individual stability of ubiquitin and chemically modified ubiquitin. These findings confirm the collaborative destabilising effect, as demonstrated by the CIU profile and CCS measurement.

Chemical modification can be used to probe proteins for mass spectrometry and protein characterisation. The most effective reagent concentrations for chemical modification have been determined under different experimental set-ups

In Chapter 3, The amino acid residue site where the modification occurred was identified, emphasizing the importance of establishing a link between unfolding detected by IM-MS and modification on specific residues. It is critical to link IM-MS unfolding measurements to specific residues using CID and ECD fragmentation techniques. The ECD results showed that acetylation occurred on K6, DEPC occurred on H68, and the succinylation modification reaction mostly affected the N-terminus of the protein while having no effect on K6.

In Chapter 4, The CCS of the WT-Ub and the acetylated Ub were compared and a considerable difference was found; the acetylated Ub was greatly destabilised. This difference indicates that the acetylated Ub has been unfolded, which can be explained by the unfolding of the compact conformation due to the addition of an acetyl group on the K6 residue. Charge distribution is considered the primary initiator of protein unfolding processes. Furthermore, the number of charges on the conformer has an impact on the structure's stability.

Histidine modification of the protein showed that it exhibits stronger destabilisation at intermediate impact energies but achieves the same unfolding population at higher energies. The first alteration is connected with H modification, whereas the second happens at K

residues. As a result, the resulted CIU data indicated that lysine modifications promote protein destabilisation.

From the succinylation results, it was concluded that modification of the N-terminus does not have an adverse effect on the stability of Ub's compact gas-phase structure. The second succinylation of the protein, on the other hand, was found to destabilise the structure.

A similar explanation was given for this as being the result of modifications made at K6. In summary, chemical modification coupled with CIU profiling could be used as a method for studying the stabilising role of intramolecular interactions within gas-phase protein structures.

In Chapter 5, CIU profiling method was applied to a variety of mutants that were created to investigate the effect of a K6 mutation on the stability of compact, folded Ub ions in the gas phase. First, the unfolding confirmation of the acetylated K6R and the K6R were shown to be very similar. It is possible that changing the lysine at K6 position to arginine causes acetylation to occur on another lysine region, having no effect on protein unfolding.

The acetylated mutant of the K6O protein is significantly less stable than the unmodified protein, presumably because the acetylation occurs at the protein's K6 location, disrupting the protein. It's important to note that the K6O Ub mutant contains a single lysine position that can be chemically modified by acetylation. since only the K6 position is modifiable, this highlights the critical role that K6 plays in the protein's stability. When acetylated and unacetylated K6O were compared, it was found that when acetylated, the drift trace looked like that of WT-Ub.

The results for the K6A mutant show excellent agreement with the AcUb in the first conformer, however the unfolded conformers indicate a greater value for the K6A mutant. Taken together, these findings demonstrate the critical role played by K6 in the stability of Ub's compact structure in the gas phase.

By combining IM-MS and molecular dynamic simulation approaches, the underlying mechanism of protein unfolding was investigated in Chapter 6. Studied effects on the K6 position in WT indicate that the interaction between protonated K6 and H68 is critical for stabilisation of the compact fold of Ub in the gas phase when changes in the K6 location occur.

The chemical mutation or mutagenesis of Ub destabilises the structure by altering its charge and the intermolecular interactions that act on it. This destabilisation results in the unfolding of Ub in the gas phase. The combination of CIU data from modified Ub and MD modelling has allowed structural insights into gas-phase proteins to be resolved down to the level of single amino acid interactions.

In Chapter 7, The result of UBA2 was discussed. UBA2, which is smaller compared to ubiquitin, showed a greater difference in charge state unfolding profile. The UBA2, when acetylated, lost its stable gas phase structure and unfolded rapidly at lower energy, whereas the unmodified UBA2 requires higher energy to unfold. In order to destabilise the overall structure, a K14A mutant was constructed by altering the K to A. This changes the neutral site rather than the protonated site. Like UBA2, the second mutant K14H changes the position of histidine but does not impact the structure or polarity of the residue.

In UBA2, The $[M+5H]^{+5}$ ion shows substantial variety in unfolding as energy increases, whereas the $[M+4H]^{+4}$ data shows no change between the variations. This implies that acetylation and mutation alter the gas-phase structure, and that the extra proton likely goes on lysine at position K14.

Finally, this work demonstrate that we can improve our understanding of protein structure in the gas phase by combining the CIU approach with chemical modification of proteins. This work has been published in *Chemistry – A European Journal*.

References

- Alberts, B., & Press, C. R. C. (2017). *Molecular biology of the cell*.
- Allison, T. M., Barran, P., Benesch, J. L. P., Cianferani, S., Degiacomi, M. T., Gabelica, V., . . . Thalassinou, K. (2020). Software Requirements for the Analysis and Interpretation of Native Ion Mobility Mass Spectrometry Data. *Analytical Chemistry*, *92*(16), 10881-10890. doi:10.1021/acs.analchem.9b05792
- Alonso, D. O., & Daggett, V. (1998). Molecular dynamics simulations of hydrophobic collapse of ubiquitin. *Protein science : a publication of the Protein Society*, *7*(4), 860-874. doi:10.1002/pro.5560070404
- Andersen, N. H. (2001). Protein Structure, Stability, and Folding. Methods in Molecular Biology. Volume 168 Edited by Kenneth P. Murphy (University of Iowa College of Medicine). Humana Press: Totowa, New Jersey. 2001. ix + 252 pp. \$89.50. ISBN 0-89603-682-0. *Journal of the American Chemical Society*, *123*(51), 12933-12934. doi:10.1021/ja0152815
- Ashcroft, A. E. (2003). Protein and peptide identification: the role of mass spectrometry in proteomics. *Nat Prod Rep*, *20*(2), 202-215.
- Badman, E. R., Myung, S., & Clemmer, D. E. (2002). Gas-phase separations of protein and peptide ion fragments generated by collision-induced dissociation in an ion trap. *Analytical Chemistry*, *74*(19), 4889-4894.
- Bafna, V., & Edwards, N. (2003). *On de novo interpretation of tandem mass spectra for peptide identification*. Paper presented at the Proceedings of the seventh annual international conference on Research in computational molecular biology.
- Bakhtiari, M., & Konermann, L. (2019). Protein Ions Generated by Native Electrospray Ionization: Comparison of Gas Phase, Solution, and Crystal Structures. *Journal of Physical Chemistry B*, *123*(8), 1784-1796. doi:10.1021/acs.jpcc.8b12173
- Banerjee, S., & Mazumdar, S. (2012). Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International journal of analytical chemistry*, 2012.

- Baslé, E., Joubert, N., & Pucheault, M. (2010). Protein Chemical Modification on Endogenous Amino Acids. *Chemistry & Biology*, 17(3), 213-227. doi:<https://doi.org/10.1016/j.chembiol.2010.02.008>
- Benesch, J. L. P., & Ruotolo, B. T. (2011). Mass spectrometry: come of age for structural and dynamical biology. *Current Opinion in Structural Biology*, 21(5), 641-649. doi:<https://doi.org/10.1016/j.sbi.2011.08.002>
- Biemann, K. (1990). Appendix 5. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol*, 193, 886-887. doi:10.1016/0076-6879(90)93460-3
- Bowers, M. T., Kemper, P. R., von Helden, G., & van Koppen, P. A. (1993). Gas-phase ion chromatography: transition metal state selection and carbon cluster formation. *Science*, 260(5113), 1446-1451. doi:10.1126/science.260.5113.1446
- Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., . . . Karplus, M. (2009). CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry*, 30(10), 1545-1614. doi:10.1002/jcc.21287
- Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V., & Ruotolo, B. T. (2010). Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Analytical Chemistry*, 82(22), 9557-9565. doi:10.1021/ac1022953
- Butterworth, P. J. (2005). Lehninger: principles of biochemistry (4th edn) D. L. Nelson and M. C. Cox, W. H. Freeman & Co., New York, 1119 pp (plus 17 pp glossary), ISBN 0-7167-4339-6 (2004). *Cell Biochemistry and Function*, 23(4), 293-294. doi:<https://doi.org/10.1002/cbf.1216>
- Chen, L., Shinde, U., Ortolan, T. G., & Madura, K. (2001). Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly. *EMBO Rep*, 2(10), 933-938. doi:10.1093/embo-reports/kve203
- Chowdhury, S., Katta, V., & Chait, B. (1990). Electrospray ionization mass spectrometric peptide mapping: a rapid, sensitive technique for protein structure analysis. *Biochemical and biophysical research communications*, 167(2), 686-692.

- Clark, C. R. (1997). A Stopped-Flow Kinetics Experiment for Advanced Undergraduate Laboratories: Formation of Iron(III) Thiocyanate. *Journal of Chemical Education*, 74(10), 1214. doi:10.1021/ed074p1214
- Clemmer, D. E., & Jarrold, M. F. (1997). Ion Mobility Measurements and their Applications to Clusters and Biomolecules. *Journal of Mass Spectrometry*, 32(6), 577-592. doi:10.1002/(SICI)1096-9888(199706)32:6<577::AID-JMS530>3.0.CO;2-4
- Cotter, R. J., Gardner, B. D., Ilchenko, S., & English, R. D. (2004). Tandem time-of-flight mass spectrometry with a curved field reflectron. *Analytical Chemistry*, 76(7), 1976-1981.
- Crabtree, M. D., & Shamma, S. L. (2018). Chapter Fourteen - Stopped-Flow Kinetic Techniques for Studying Binding Reactions of Intrinsically Disordered Proteins. In E. Rhoades (Ed.), *Methods Enzymol* (Vol. 611, pp. 423-457): Academic Press.
- Dage, J. L., Sun, H., & Halsall, H. B. (1998). Determination of diethylpyrocarbonate-modified amino acid residues in alpha 1-acid glycoprotein by high-performance liquid chromatography electrospray ionization-mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight-mass spectrometry. *Anal Biochem*, 257(2), 176-185. doi:10.1006/abio.1997.2552
- de Hoffmann, E. (1996). Tandem mass spectrometry: a primer. *Journal of Mass Spectrometry*, 31(2), 129-137.
- Díaz, I., de Castro, I. N., & Medina, M. A. (1993). An experiment on the chemical modification of essential histidine residues of lactate dehydrogenase. *Biochemical Education*, 21(4), 219-220. doi:10.1016/0307-4412(93)90106-A
- Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., & Alice, M. B. (1968). Molecular Beams of Macroions. *The Journal of Chemical Physics*, 49(5), 2240-2249. doi:10.1063/1.1670391
- Drahl, C., Cravatt, B. F., & Sorensen, E. J. (2005). Protein-reactive natural products. *Angewandte Chemie International Edition*, 44(36), 5788-5809.
- Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., . . . Obradovic, Z. (2001). Intrinsically disordered protein. *J Mol Graph Model*, 19(1), 26-59. doi:10.1016/s1093-3263(00)00138-8

- El-Faramawy, A., Siu, K. W. M., & Thomson, B. A. (2005). Efficiency of Nano-Electrospray Ionization. *Journal of the American Society for Mass Spectrometry*, 16(10), 1702-1707.
doi:<https://doi.org/10.1016/j.jasms.2005.06.011>
- Elviri, L. (2012). ETD and ECD Mass Spectrometry Fragmentation for the Characterization of Protein Post Translational Modifications. In.
- Eschweiler, J. D., Rabuck-Gibbons, J. N., Tian, Y., & Ruotolo, B. T. (2015). CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. *Analytical Chemistry*, 87(22), 11516-11522. doi:10.1021/acs.analchem.5b03292
- Ewing, M. A., Glover, M. S., & Clemmer, D. E. (2016). Hybrid ion mobility and mass spectrometry as a separation tool. *Journal of Chromatography A*, 1439, 3-25.
doi:<https://doi.org/10.1016/j.chroma.2015.10.080>
- Fersht, A. R. (1993). Protein folding and stability: the pathway of folding of barnase. *FEBS Letters*, 325(1), 5-16. doi:[https://doi.org/10.1016/0014-5793\(93\)81405-Q](https://doi.org/10.1016/0014-5793(93)81405-Q)
- Fox, T., & Kollman, P. A. (1996). The application of different solvation and electrostatic models in molecular dynamics simulations of ubiquitin: how well is the X-ray structure "maintained"? *Proteins*, 25(3), 315-334.
doi:10.1002/(sici)1097-0134(199607)25:3<315::Aid-prot4>3.0.Co;2-e
- Garcia, I. R., Giles, K., Bateman, R. H., & Gaskell, S. J. (2008). Studies of peptide a- and b-type fragment ions using stable isotope labeling and integrated ion mobility/tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry*, 19(12), 1781-1787.
- Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D., & Boyse, E. A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences of the United States of America*, 72(1), 11-15. doi:10.1073/pnas.72.1.11
- Gross, J. H. (2006). *Mass spectrometry: a textbook*. Springer Science & Business Media.
- Han, L., Hyung, S.-J., Mayers, J. J. S., & Ruotolo, B. T. (2011). Bound anions differentially stabilize multiprotein complexes in the absence of bulk solvent.

- Journal of the American Chemical Society*, 133(29), 11358-11367.
doi:10.1021/ja203527a
- Han, L., & Ruotolo, B. T. (2015). Ion Mobility-Mass Spectrometry Differentiates Protein Quaternary Structures Formed in Solution and in Electrospray Droplets. *Analytical Chemistry*, 87(13), 6808-6813.
doi:10.1021/acs.analchem.5b01010
- Heessen, S., Masucci, M. G., & Dantuma, N. P. (2005). The UBA2 Domain Functions as an Intrinsic Stabilization Signal that Protects Rad23 from Proteasomal Degradation. *Molecular Cell*, 18(2), 225-235.
doi:<https://doi.org/10.1016/j.molcel.2005.03.015>
- Hershko, A., & Ciechanover, A. (1998). The ubiquitin system. *Annual review of biochemistry*, 67(1), 425-479.
- Hicke, L. (1997). Ubiquitin-dependent internalization and down-regulation of plasma membrane proteins. *The FASEB Journal*, 11(14), 1215-1226.
- Hoaglund, C. S., Valentine, S. J., Sporleder, C. R., Reilly, J. P., & Clemmer, D. E. (1998). Three-Dimensional Ion Mobility/TOFMS Analysis of Electrosprayed Biomolecules. *Analytical Chemistry*, 70(11), 2236-2242.
doi:10.1021/ac980059c
- Hopper, J. T., & Oldham, N. J. (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(10), 1851-1858. doi:10.1016/j.jasms.2009.06.010
- Huang, J., & MacKerell Jr, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25), 2135-2145. doi:<https://doi.org/10.1002/jcc.23354>
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33-38. doi:[https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
- Hyung, S.-J., Robinson, C. V., & Ruotolo, B. T. (2009). Gas-Phase Unfolding and Disassembly Reveals Stability Differences in Ligand-Bound Multiprotein Complexes. *Chemistry & Biology*, 16(4), 382-390.
doi:<https://doi.org/10.1016/j.chembiol.2009.02.008>
- Im, H. (2011). The Inoue Method for Preparation and Transformation of Competent *E. coli*: "Ultra Competent" Cells. *Bio-protocol* e143(Bio101).

- Jenner, M., Ellis, J., Huang, W. C., Lloyd Raven, E., Roberts, G. C., & Oldham, N. J. (2011). Detection of a protein conformational equilibrium by electrospray ionisation-ion mobility-mass spectrometry. *Angewandte Chemie International Edition*, 50(36), 8291-8294. Retrieved from <https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201101077>
- Jensen, O. N. (2004). Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Current Opinion in Chemical Biology*, 8(1), 33-41. doi:10.1016/j.cbpa.2003.12.009
- Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., . . . Eaton, W. A. (1993). Fast events in protein folding initiated by nanosecond laser photolysis. *Proceedings of the National Academy of Sciences*, 90(24), 11860-11864. doi:10.1073/pnas.90.24.11860
- Kanu, A. B., Dwivedi, P., Tam, M., Matz, L., & Hill, H. H. (2008). Ion mobility–mass spectrometry. *Journal of Mass Spectrometry*, 43(1), 1-22. doi:10.1002/jms.1383
- Karplus, M., & Kuriyan, J. (2005). Molecular dynamics and protein function. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19), 6679-6685. doi:10.1073/pnas.0408930102
- Karplus, M., & Šali, A. (1995). Theoretical studies of protein folding and unfolding. *Current Opinion in Structural Biology*, 5(1), 58-73. doi:[https://doi.org/10.1016/0959-440X\(95\)80010-X](https://doi.org/10.1016/0959-440X(95)80010-X)
- Kebarle, P., & Peschke, M. (2000). On the Mechanisms by Which the Charged Droplets Produced by Electrospray Lead to Gas Phase Ions. *Analytica Chimica Acta*, 406, 11-35. doi:10.1016/S0003-2670(99)00598-X
- Kemper, P. R., Dupuis, N. F., & Bowers, M. T. (2009). A new, higher resolution, ion mobility mass spectrometer. *International Journal of Mass Spectrometry*, 287(1), 46-57. doi:<https://doi.org/10.1016/j.ijms.2009.01.012>
- Kimura, Y., & Tanaka, K. (2010). Regulatory mechanisms involved in the control of ubiquitin homeostasis. *J Biochem*, 147(6), 793-798. doi:10.1093/jb/mvq044
- Konermann, L. (2017). Molecular Dynamics Simulations on Gas-Phase Proteins with Mobile Protons: Inclusion of All-Atom Charge Solvation. *The Journal of Physical Chemistry B*, 121(34), 8102-8112. doi:10.1021/acs.jpccb.7b05703
- Konermann, L., & Douglas, D. (1998). Unfolding of proteins monitored by electrospray ionization mass spectrometry: a comparison of positive and

- negative ion modes. *Journal of the American Society for Mass Spectrometry*, 9(12), 1248-1254.
- Konermann, L., McAllister, R. G., & Metwally, H. (2014). Molecular Dynamics simulations of the electrospray process: formation of NaCl clusters via the charged residue mechanism. *Journal of Physical Chemistry B*, 118(41), 12025-12033. doi:10.1021/jp507635y
- Kony, D. B., Hünenberger, P. H., & van Gunsteren, W. F. (2007). Molecular dynamics simulations of the native and partially folded states of ubiquitin: influence of methanol cosolvent, pH, and temperature on the protein structure and dynamics. *Protein science : a publication of the Protein Society*, 16(6), 1101-1118. doi:10.1110/ps.062323407
- Lanucara, F., Holman, S. W., Gray, C. J., & Eyers, C. E. (2014). The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature Chemistry*, 6(4), 281-294. doi:10.1038/nchem.1889
- Lermyte, F., Lacki, M. K., Valkenborg, D., Gambin, A., & Sobott, F. (2017). Conformational Space and Stability of ETD Charge Reduction Products of Ubiquitin. *J Am Soc Mass Spectrom*, 28(1), 69-76. doi:10.1007/s13361-016-1444-7
- Lermyte, F., Łacki, M. K., Valkenborg, D., Gambin, A., & Sobott, F. (2017). Conformational Space and Stability of ETD Charge Reduction Products of Ubiquitin. *Journal of the American Society for Mass Spectrometry*, 28(1), 69-76. doi:10.1007/s13361-016-1444-7
- Lindorff-Larsen, K., Piana, S., Dror, R. O., & Shaw, D. E. (2011). How fast-folding proteins fold. *Science*, 334(6055), 517-520. doi:10.1126/science.1208351
- Mackerell, A. D., Jr. (2004). Empirical force fields for biological macromolecules: overview and issues. *J Comput Chem*, 25(13), 1584-1604. doi:10.1002/jcc.20082
- Marianayagam, N. J., & Jackson, S. E. (2005). Native-state dynamics of the ubiquitin family: implications for function and evolution. *Journal of the Royal Society, Interface*, 2(2), 47-54. doi:10.1098/rsif.2004.0025
- Marshall, A. G., Hendrickson, C. L., & Jackson, G. S. (1998). Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom Rev*,

- 17(1), 1-35. doi:[https://doi.org/10.1002/\(SICI\)1098-2787\(1998\)17:1<1::AID-MAS1>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K)
- Mason, E. A., & McDaniel, E. W. (1988). Transport properties of ions in gases. *NASA STI/Recon Technical Report A*, 89.
- McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, 267(5612), 585-590. doi:10.1038/267585a0
- McDaniel, E. W., & Mason, E. A. (1973). Mobility and diffusion of ions in gases.
- Mendoza, V. L., & Vachet, R. W. (2009). Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev*, 28(5), 785-815. doi:10.1002/mas.20203
- Merenbloom, S. I., Koeniger, S. L., Valentine, S. J., Plasencia, M. D., & Clemmer, D. E. (2006). IMS-IMS and IMS-IMS-IMS/MS for separating peptide and protein fragment ions. *Analytical Chemistry*, 78(8), 2802-2809. doi:10.1021/ac052208e
- Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., & Dobson, C. M. (1993). Detection of transient protein folding populations by mass spectrometry. *Science*, 262(5135), 896-900. doi:10.1126/science.8235611
- Mirgorodskaya, E., Roepstorff, P., & Zubarev, R. A. (1999). Localization of O-Glycosylation Sites in Peptides by Electron Capture Dissociation in a Fourier Transform Mass Spectrometer. *Analytical Chemistry*, 71(20), 4431-4436. doi:10.1021/ac990578v
- Niu, S., Rabuck, J. N., & Ruotolo, B. T. (2013). Ion mobility-mass spectrometry of intact protein–ligand complexes for pharmaceutical drug discovery and development. *Current Opinion in Chemical Biology*, 17(5), 809-817. doi:<https://doi.org/10.1016/j.cbpa.2013.06.019>
- Niu, S., & Ruotolo, B. T. (2015). Collisional unfolding of multiprotein complexes reveals cooperative stabilization upon ligand binding. *Protein Science*, 24(8), 1272-1281. doi:10.1002/pro.2699
- Piana, S., Lindorff-Larsen, K., & Shaw, D. E. (2013). Atomic-level description of ubiquitin folding. *Proceedings of the National Academy of Sciences*, 110(15), 5915-5920. doi:10.1073/pnas.1218321110
- Pickart, C. M., & Eddins, M. J. (2004). Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta*, 1695(1-3), 55-72. doi:10.1016/j.bbamcr.2004.09.019

- Polasky, D. A., Dixit, S. M., Fantin, S. M., & Ruotolo, B. T. (2019). CIUSuite 2: Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data. *Analytical Chemistry*, *91*(4), 3147-3155. doi:10.1021/acs.analchem.8b05762
- Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinos, K., . . . Scrivens, J. H. (2007). An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *International Journal of Mass Spectrometry*, *261*(1), 1-12. doi:<https://doi.org/10.1016/j.ijms.2006.07.021>
- Pucheault, M. (2008). Natural products: chemical instruments to apprehend biological symphony. *Organic & biomolecular chemistry*, *6*(3), 424-432.
- Revercomb, H. E., & Mason, E. A. (1975). Theory of plasma chromatography/gaseous electrophoresis. Review. *Analytical Chemistry*, *47*(7), 970-983. doi:10.1021/ac60357a043
- Roder, H., Maki, K., & Cheng, H. (2006). Early events in protein folding explored by rapid mixing methods. *Chemical Reviews*, *106*(5), 1836-1861. doi:10.1021/cr040430y
- Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J., & Robinson, C. V. (2008). Ion mobility–mass spectrometry analysis of large protein complexes. *Nature Protocols*, *3*, 1139. doi:10.1038/nprot.2008.78
- Ruotolo, B. T., Giles, K., Campuzano, I., Sandercock, A. M., Bateman, R. H., & Robinson, C. V. (2005). Evidence for macromolecular protein rings in the absence of bulk water. *Science*, *310*(5754), 1658-1661. doi:10.1126/science.1120177
- Sadoul, K., Boyault, C., Pabion, M., & Khochbin, S. (2008). Regulation of protein turnover by acetyltransferases and deacetylases. *Biochimie*, *90*(2), 306-312.
- Santos, L. S. (2009). Reactive intermediates : MS investigations in solution. Retrieved from http://www.123library.org/book_details/?id=13628
- Scarff, C. A., Thalassinos, K., Hilton, G. R., & Scrivens, J. H. (2008). Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Rapid communications in mass spectrometry*, *22*(20), 3297-3304. doi:<https://doi.org/10.1002/rcm.3737>
- Schaeffer-Reiss, C. (2008). A Brief Summary of the Different Types of Mass Spectrometers Used in Proteomics. In J. D. Thompson, M. Ueffing, & C.

- Schaeffer-Reiss (Eds.), *Functional Proteomics: Methods and Protocols* (pp. 3-16). Totowa, NJ: Humana Press.
- Schlick, T. (1996). Pursuing Laplace's Vision on Modern Computers. In J. P. Mesirov, K. Schulten, & D. W. Sumners (Eds.), *Mathematical Approaches to Biomolecular Structure and Dynamics* (pp. 219-247). New York, NY: Springer New York.
- Selkoe, D. J. (2003). Folding proteins in fatal ways. *Nature*, *426*(6968), 900-904. doi:10.1038/nature02264
- Shelimov, K. B., Clemmer, D. E., Hudgins, R. R., & Jarrold, M. F. (1997). Protein structure in vacuo: gas-phase conformations of BPTI and cytochrome c. *Journal of the American Chemical Society*, *119*(9), 2240-2248.
- Shelimov, K. B., & Jarrold, M. F. (1997). Conformations, Unfolding, and Refolding of Apomyoglobin in Vacuum: An Activation Barrier for Gas-Phase Protein Folding. *Journal of the American Chemical Society*, *119*(13), 2987-2994. doi:10.1021/ja962914k
- Shi, H., & Clemmer, D. E. (2014). Evidence for Two New Solution States of Ubiquitin by IMS-MS Analysis. *The Journal of Physical Chemistry. B*, *118*(13), 3498-3506. doi:10.1021/jp4097327
- Shi, H., Pierson, N. A., Valentine, S. J., & Clemmer, D. E. (2012). Conformation Types of Ubiquitin [M+8H]⁸⁺ Ions from Water:Methanol Solutions: Evidence for the N and A States in Aqueous Solution. *The Journal of Physical Chemistry B*, *116*(10), 3344-3352. doi:10.1021/jp210797x
- Sleno, L., & Volmer, D. A. (2004). Ion activation methods for tandem mass spectrometry. *Journal of Mass Spectrometry*, *39*(10), 1091-1112. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/jms.703>
- Sokratous, K., Roach, L. V., Channing, D., Strachan, J., Long, J., Searle, M. S., . . . Oldham, N. J. (2012). Probing Affinity and Ubiquitin Linkage Selectivity of Ubiquitin-Binding Domains Using Mass Spectrometry. *Journal of the American Chemical Society*, *134*(14), 6416-6424. doi:10.1021/ja300749d
- Suckau, D., Mak, M., & Przybylski, M. (1992). Protein surface topology-probing by selective chemical modification and mass spectrometric peptide mapping. *Proceedings of the National Academy of Sciences*, *89*(12), 5630-5634.
- Tian, Y., Han, L., Buckner, A. C., & Ruotolo, B. T. (2015). Collision Induced Unfolding of Intact Antibodies: Rapid Characterization of Disulfide Bonding

- Patterns, Glycosylation, and Structures. *Analytical Chemistry*, 87(22), 11509-11515. doi:10.1021/acs.analchem.5b03291
- Tian, Y., & T. Ruotolo, B. (2017). *Collision Induced Unfolding Detects Subtle Differences in Intact Antibody Glycoforms and Associated Fragments* (Vol. 425).
- Valentine, S. J., Anderson, J. G., Ellington, A. D., & Clemmer, D. E. (1997). Disulfide-intact and-reduced lysozyme in the gas phase: conformations and pathways of folding and unfolding. *The Journal of Physical Chemistry B*, 101(19), 3891-3900.
- Vijay-Kumar, S., Bugg, C. E., & Cook, W. J. (1987). Structure of ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology*, 194(3), 531-544. doi:10.1016/0022-2836(87)90679-6
- Wlodarski, T., & Zagrovic, B. (2009). Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin. *Proceedings of the National Academy of Sciences*, 106(46), 19346-19351. doi:10.1073/pnas.0906966106
- Wolff, M. M., & Stephens, W. E. (1953). A Pulsed Mass Spectrometer with Time Dispersion. *Review of Scientific Instruments*, 24(8), 616-617. doi:10.1063/1.1770801
- Woods, L. A., Radford, S. E., & Ashcroft, A. E. (2013). Advances in ion mobility spectrometry-mass spectrometry reveal key insights into amyloid assembly. *Biochim Biophys Acta*, 1834(6), 1257-1268. doi:10.1016/j.bbapap.2012.10.002
- Wytttenbach, T., Pierson, N. A., Clemmer, D. E., & Bowers, M. T. (2014). Ion mobility analysis of molecular dynamics. *Annu Rev Phys Chem*, 65, 175-196. doi:10.1146/annurev-physchem-040513-103644
- Wytttenbach, T., Pierson, N. A., Clemmer, D. E., & Bowers, M. T. (2014). Ion Mobility Analysis of Molecular Dynamics. *Annu Rev Phys Chem*, 65(1), 175-196. doi:10.1146/annurev-physchem-040513-103644
- Yang, R. (2020). Ubiquitin binding domains – from structure to application. *IOP Conference Series: Materials Science and Engineering*, 738, 012013. doi:10.1088/1757-899x/738/1/012013

- Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., & Zhao, Y. (2011). Identification of lysine succinylation as a new post-translational modification. *Nat Chem Biol*, 7(1), 58-63. doi:10.1038/nchembio.495
- Zhong, Y., Hyung, S.-J., & Ruotolo, B. T. (2012). Ion mobility–mass spectrometry for structural proteomics. *Expert Review of Proteomics*, 9(1), 47-58. doi:10.1586/epr.11.75
- Zhou, Y., & Vachet, R. W. (2012). Diethylpyrocarbonate labeling for the structural analysis of proteins: label scrambling in solution and how to avoid it. *J Am Soc Mass Spectrom*, 23(5), 899-907. doi:10.1007/s13361-012-0349-3