



# THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

# Resolving Intrinsically Disordered Proteins of the Cancer Genome with Ion Mobility Mass Spectrometry

Ewa Jurneczko



Ph.D.

The University of Edinburgh

2014

‘If you try and take a cat apart to see how it works, the first thing you have on your hands is a nonworking cat’

- Douglas Adams

‘If you can’t explain it simply, you don’t understand it well enough’

- Albert Einstein

## **Declaration**

This thesis is submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Edinburgh. Unless otherwise stated, this work is my own and has not been submitted for any other degree or professional qualification.

.....

Ewa Jurneczko

January 2014



## Acknowledgements

Firstly I would like to thank my supervisor Prof Perdi Barran for taking me on and letting me to explore and enjoy the arts of hands-on science from day one of my PhD. Some moments spent in the Barran lab will be always reminisced. Big TUNING times, drift cell, RF generators, taking ToF tube apart twice ('how clean is your ToF?'), changing MCPs, changing turbo pumps, pusher boards, cleaning the o-rings and fun times with the leak detector. It goes without saying but for me taking things apart and putting them back together = seeing better how things work!

A special thank you to all my fellow group members over the last four years for their invaluable assistance, support and discussions: Jason, Martin, Roland, Jude, Yana, Hattie, Luiz, Kama, Sophie, Bex, Jakub, Ashley, Ellie, Chris and Alex. My special thanks go to Dr Yana Berezovskaya for sharing her technical knowledge, to Dr Harriet Cole for advice on Origin issues. And special mention to Kamila Pacholarz for our frequent phone chats and being my travel buddy. I am also grateful to my talented project students Ellie, Ashley, Faye and Danting for all their hard work. Thank you also Dr Massimiliano Porrini for your priceless help with the calculations.

I am grateful to my collaborators: Dr Penka Nikolova, Prof Richard Kriwacki, people who can express proteins Dr Dave Clarke, Dr Jude Nicholson, Dr Erin Worrall and Ariele Follis and Dr Iain Campuzano for work at Waters. I am grateful for all the technical help provided by Stuart and Davey in the mechanical workshop, Donald in the electronic workshop, Hunter in the physics electronic workshop and Danny for being at the end of the phone when needed.

I am also appreciative of sending me to domestic and international conferences that provided me with an annual injection of inspiration and energy. Finally I am indebted to my family, and my friends for their continuous support and cheering me up. And capoeira, my love for martial arts, dance and music that together with my capoeira friends brought a balance into my scientific life and helped me to get through the tough times.

## Abstract

For proteins the link between their structure and their function is a central tenet of biology. A common approach to understanding protein function is to ‘solve’ its structure and subsequently probe interactions between the protein and its binding partners. The first part of this approach is non-trivial for proteins where localised regions or even their entire structure fail to fold into a three-dimensional structure and yet they possess function. These so called intrinsically or inherently disordered proteins (IDP’s) or intrinsically disordered regions (IDR’s) constitute up to 40% of all expressed proteins. IDPs which have crucial roles in molecular recognition, assembly, protein modification and entropic chain activities, are often dynamic with respect to both conformation and interaction, so in the course of a protein’s ‘lifespan’ it will sample many configurations and bind to several targets. For these proteins, there is a need to develop new methods for structure characterization which exploit their biophysical properties. The solvent free environment of a mass spectrometer is ideally suited to the study of intrinsic interactions and how they contribute to structure. Ion mobility mass spectrometry is uniquely able to observe the range of structures an IDP can occupy, and also the effect of selected binding partners on altering this conformational space.

This thesis details the technique of ion mobility mass spectrometry and illustrates its use in assessing the relative disorder of p53 protein. The tumour suppressor p53 is at the hub of a plethora of signalling pathways that maintain the integrity of the human genome and regulate the cell cycle. Deregulation of this protein has a great effect on carcinogenesis as mutated p53 can induce an amplified epigenetic instability of tumour cells, facilitating and accelerating the evolution of the tumour. Herein mass spectrometry provides a compelling, detailed insight into the conformational flexibility of the p53 DNA-binding domain. The plasticity of the p53 DNA-binding domain is reflected in the existence of more than one conformation, independent of any conformational changes prompted by binding. The *in vacuo* conformational phenotypes exhibited by common cancer-associated mutations are determined and

the second-site suppressor mutation from loop L1, H115N, is probed whether it could trigger conformational changes in p53 hotspot cancer mutations. The structural basis of the binding promiscuity of p53 protein is investigated; of particular interest is the molecular interaction of the p53 N-terminus with the oncoprotein murine double minute 2, as well as with the antiapoptotic factor B-cell lymphoma-extra-large.

## Lay Summary

Proteins are fundamental components of all living cells and are necessary for the proper functioning of an organism. The building blocks of proteins are called amino acids. Proteins differ from one another primarily in their sequence of amino acids, dictated by the nucleotide sequence of their genes, which usually results in folding of the protein into a specific three-dimensional shape that determines its activity. However proteins are not entirely rigid molecules. Proteins may transit between a number of related structures while they execute their functions. In the context of these functional rearrangements, these structures are usually referred to as conformations, and transitions between them are termed conformational changes.

This work is devoted to protein p53 that is a tumour suppressor, which in humans is encoded by the TP53 gene. More than fifty percent of human tumours contain a mutation or deletion of the TP53 gene. These mutations can cause errors in protein amino acid sequence, producing partially or completely non-functional proteins. Each cell, in order to function correctly, is dependent on a myriad of proteins to function in the right places at the right times. When a mutation affects a protein that plays a critical role in the body, a disease can develop. Mutant p53, however, does not function properly, allowing the cancer cells it routinely would target to slip through control mechanisms and proliferate. For this reason, the protein is a key target of research on cancer therapeutics. Here the various forms of p53 protein are captured and the conformational changes associated with binding to DNA and other interacting proteins are described. The diverse shapes of most common mutant proteins are also investigated.

## Abbreviations

ATD	arrival time distribution
AMBER	assisted model building with energy refinement
ATM	atomic force microscopy
Bad	B-cell lymphoma 2-associated death promoter
Bak	B-cell lymphoma 2 homologous antagonist/killer
Bax	B-cell lymphoma 2-associated X protein
BcL-2	B-cell lymphoma 2
BcL-xL	B-cell lymphoma-extra-large
BH3	B-cell lymphoma 2 homology 3
CCS	collision cross section
CD	circular dichroism
CDB3	cytoplasmic domain of band 3 protein
CE	collision energy
CSD	charge state distribution
CV	collision voltage
Da	Dalton
DBD	DNA-binding domain
DC	direct current
DNA	deoxyribonucleic acid
DT	drift tube
DT IM-MS	drift tube ion mobility mass spectrometry
EDC	enhanced duty cycle
EDTA	ethylene diamine tetraacetic acid
EHSS	exact hard sphere scattering
ESI	electrospray ionisation
EM	electron microscopy
FAIMS	field asymmetric ion mobility spectrometry
FRET	fluorescence resonance energy transfer
FTIR	fourier transform infrared spectroscopy

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HDMS	high definition mass spectrometry
HMGA	high mobility group A
IARC	international agency for research on cancer
IDP	intrinsically disordered protein
IDR	intrinsically disordered region
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
IUP	intrinsically unstructured protein
IM-MS	ion mobility mass spectrometry
IMS	ion mobility spectrometry
MALDI	matrix assisted laser desorption ionisation
MCP	microchannel plate detectors
Mdm2	mouse double minute 2
Mdm4	mouse double minute 4
MIRA-1	mutant p53 reactivation and induction of rapid apoptosis
MKRN1	makorin ring finger protein 1
MoQToF	mobility quadrupole time-of-flight
MoRF	molecular recognition feature
MS	mass spectrometry
$m/z$	mass to charge ratio
NaCl	sodium chloride
NaI	sodium iodide
NMR	nuclear magnetic resonance
n-ESI	nano-electrospray ionisation
OD	optical density
PA	projection approximation
PCB	printed circuit board
PDB	protein data bank
PONDR	predictor of naturally disordered regions
PRIMA-1	p53 reactivation and induction of massive apoptosis
PSA	projected superposition approximation
PUMA	p53 upregulated modulator of apoptosis

QToF	quadrupole time of flight
RING	really interesting new gene
RF	radio frequency
SANS	small angle neutron scattering
SAXS	small angle X-ray scattering
SRIG	stacked ring ion guide
SS	superstable
TAD	transcription activation domain
tATD	total arrival time distribution
TCEP	tris(2-carboxyethyl)phosphine
TDC	time to digital converter
TIC	total ion count
TM	trajectory method
ToF	time of flight
T-wave	travelling wave
Tris	tris(hydroxymethyl)aminomethane
TWIG	travelling wave ion guide
TW IM-MS	travelling wave ion mobility mass spectrometry
UV	ultraviolet
WT	wild type
VMD	visual molecular dynamics

## Physical Constants and Quantities

$e$	elementary charge
$E$	electric field strength
$K$	mobility
$K_0$	reduced mobility
$k_B$	Boltzmann constant
$L$	drift cell length
$m$	mass
$M_R$	relative molecular weight
$m/z$	mass to charge ratio
$N$	number density
$P$	pressure
$T$	temperature
$t_0$	dead time
$t_a$	arrival time
$t_d$	drift time
$v$	velocity
$V$	potential difference
$v_d$	drift velocity
$z$	nominal charge
$\mu$	reduced mass
$\nu$	RF
$\Omega$	momentum transfer integral



# Table of Contents

<b>Declaration.....</b>	<b>i</b>
<b>Acknowledgements.....</b>	<b>ii</b>
<b>Abstract.....</b>	<b>iii</b>
<b>Lay Summary .....</b>	<b>v</b>
<b>Abbreviations .....</b>	<b>vi</b>
<b>Physical Constants and Quantities .....</b>	<b>ix</b>

<b>1. Introduction and Motivation.....</b>	<b>- 1 -</b>
<b>1.1. The Rules of Disorder and Why Disorder Rules.....</b>	<b>- 1 -</b>
<b>1.2. Intrinsically Disordered Proteins: definition, detection and beyond - 2 -</b>	
1.2.1. Definition and Unique Properties of IDPs .....	- 2 -
1.2.2. Functions of IDPs: the D <sup>2</sup> Concept.....	- 5 -
1.2.3. Experimental Identification and Prediction Methods .....	- 6 -
<b>1.3. The Tumour Suppressor p53: Puzzle and Paradigm.....</b>	<b>- 8 -</b>
1.3.1. The Bigger Picture: Quaternary Structure of p53 .....	- 12 -
1.3.2. Post-translational Modifications of p53 .....	- 15 -
<b>1.4. p53:Mdm2 - the Affair that Never Ends .....</b>	<b>- 16 -</b>
<b>1.5. The Case for Mass Spectrometry as a Tool to Study IDP's .....</b>	<b>- 17 -</b>
<b>1.6. Ion Mobility Mass Spectrometry - to determine Conformation(s). - 19 -</b>	
1.6.1. Drift Tube Ion Mobility Mass Spectrometry .....	- 22 -
1.6.2. Travelling Wave Ion Mobility Mass Spectrometry .....	- 25 -
1.6.3. Field Asymmetric Waveform Ion Mobility Mass Spectrometry ..	- 27 -
<b>1.7. Summary .....</b>	<b>- 28 -</b>
<b>1.8. References .....</b>	<b>- 29 -</b>

## Chapter Two

<b>2. Experimental Methods .....</b>	<b>- 43 -</b>
<b>2.1. Reagents .....</b>	<b>- 43 -</b>
<b>2.2. Sample Introduction - Ionisation.....</b>	<b>- 43 -</b>
<b>2.3. Instruments .....</b>	<b>- 46 -</b>

<b>2.4. Drift Tube based IM-MS - the MoQToF .....</b>	<b>- 46 -</b>
2.4.1. Instrument Layout .....	- 46 -
2.4.2. Drift Tube Architecture.....	- 50 -
2.4.2.1. Drift Gas.....	- 52 -
2.4.2.2. Pressure and Temperature Regulation .....	- 52 -
2.4.3. Instrument Tuning Parameters .....	- 52 -
2.4.4. Data Acquisition and Analysis: an example .....	- 54 -
2.4.4.1. Multiple ATD Peak Fitting .....	- 60 -
<b>2.5. Travelling Wave based IM-MS - Synapt HDMS .....</b>	<b>- 62 -</b>
2.5.1. Instrument Layout and Operation .....	- 62 -
2.5.1.1. Travelling Wave Ion Guide Topology .....	- 63 -
2.5.1.2. Instrument Tuning Parameters and Typical Pressures .....	- 65 -
2.5.2. Calculation of CCS using TW IM-MS data.....	- 68 -
<b>2.6. Mass Calibration .....</b>	<b>- 69 -</b>
<b>2.7. Optimization of experimental conditions.....</b>	<b>- 70 -</b>
<b>2.8. Estimation of collision cross sections from PDB structures .....</b>	<b>- 71 -</b>
2.8.1. Projection Approximation Method .....	- 72 -
2.8.2. Exact Hard Sphere Scattering Method.....	- 73 -
2.8.3. Trajectory Method.....	- 73 -
2.8.4. Projected Superposition Approximation Method.....	- 73 -
<b>2.9. References .....</b>	<b>- 75 -</b>

### Chapter Three

<b>3. p53 DNA-binding domain <i>in vacuo</i> .....</b>	<b>- 80 -</b>
<b>3.1. Introduction .....</b>	<b>- 80 -</b>
<b>3.2. How p53 binds DNA.....</b>	<b>- 82 -</b>
<b>3.3. Methodology .....</b>	<b>- 84 -</b>
3.3.1. Protein Expression and Purification.....	- 84 -
3.3.2. Sample Preparation .....	- 84 -
3.3.2.1. DNA Oligonucleotides.....	- 85 -
3.3.3. Mass Spectrometry and Ion Mobility Mass Spectrometry.....	- 85 -
3.3.4. Theoretical Measurements of collision cross sections .....	- 86 -
3.3.5. Production of Waterfall Plots for data obtained on TW IM-MS ..	- 87 -

<b>3.4. Insights into Structure(s) and Dynamics of WT p53-DBD by MS and IM-MS .....</b>	<b>- 87 -</b>
3.4.1. MS of WT p53-DBD in buffered solution conditions .....	- 87 -
3.4.2. MS of WT p53-DBD in denaturing solution conditions.....	- 89 -
3.4.3. MS of WT p53-DBD in high buffer solution conditions .....	- 90 -
3.4.4. Conformational Landscapes occupied by WT p53-DBD .....	- 91 -
3.4.4.1. Collision Cross Sections of WT p53-DBD .....	- 97 -
<b>3.5. Interactions of WT p53-DBD with its consensus DNA binding site-</b>	<b>99 -</b>
3.5.1. WT p53-DBD dimer bound to ds 12 mer DNA.....	- 99 -
3.5.2. WT p53-DBD tetramer bound to ds 20 mer DNA.....	- 104 -
<b>3.6. Summary.....</b>	<b>- 106 -</b>
<b>3.7. References .....</b>	<b>- 107 -</b>

## Chapter Four

<b>4. Common p53 Cancer Mutants.....</b>	<b>- 112 -</b>
<b>4.1. Mapping Cancer Mutations on the DNA-binding Domain .....</b>	<b>- 112 -</b>
<b>4.2. Structural Classes of p53 Core Domain Cancer Mutants .....</b>	<b>- 113 -</b>
4.2.1. DNA-contact Mutations .....	- 113 -
4.2.2. Structural Mutations in the L3 loop .....	- 114 -
4.2.3. Disruption of the Zinc-binding Region.....	- 114 -
4.2.4. Structural Mutations at the Periphery of the DBD.....	- 115 -
4.2.5. $\beta$ -sandwich Mutations .....	- 115 -
<b>4.3. Rescue of p53 Mutants by ‘Conformational’ Drugs .....</b>	<b>- 116 -</b>
<b>4.4. Methodology .....</b>	<b>- 118 -</b>
4.4.1. Site-directed Mutagenesis, p53 Expression and Purification.....	- 118 -
4.4.2. Sample Preparation .....	- 119 -
4.4.3. Mass Spectrometry and Ion Mobility Mass Spectrometry.....	- 120 -
4.4.4. Production of Waterfall Plots for data obtained via TW IM-MS	- 121 -
4.4.5. Production of Waterfall Plots for data obtained via DT IM-MS	- 121 -
<b>4.5. Insights into the Diverse Conformational Nature of Common p53 Cancer Mutants .....</b>	<b>- 122 -</b>
4.5.1. Structural Mutation in the L3 loop - R249A.....	- 122 -
4.5.2. DNA Contact Mutation - R273H .....	- 124 -

4.5.3.	K292I Mutation .....	- 126 -
4.5.4.	A276Y Mutation .....	- 128 -
4.5.5.	Disruptions of the Zinc-binding Region - R175H .....	- 130 -
<b>4.6.</b>	<b>Mechanism of Rescue of Common p53 Cancer Mutants by Second-site Suppressor Mutations .....</b>	<b>- 132 -</b>
4.6.1.	Cold Spot Mutations used as Second-site Suppressors.....	- 133 -
4.6.2.	Engineered Mutation in the L1 loop - H115N .....	- 133 -
4.6.3.	DNA-binding properties of H115N Mutant.....	- 136 -
4.6.4.	Double Mutant R249A/H115N .....	- 142 -
4.6.5.	Double Mutant R273H/H115N .....	- 144 -
4.6.6.	Double Mutant R248Q/H115N .....	- 145 -
<b>4.7.</b>	<b>Summary .....</b>	<b>- 147 -</b>
<b>4.8.</b>	<b>References .....</b>	<b>- 149 -</b>

## Chapter Five

<b>5.</b>	<b>Promiscuous Interactions of p53 .....</b>	<b>- 155 -</b>
<b>5.1.</b>	<b>Methodology .....</b>	<b>- 156 -</b>
5.1.1.	Protein Expression and Purification.....	- 156 -
5.1.2.	Sample Preparation .....	- 157 -
5.1.3.	Mass Spectrometry and Ion Mobility Mass Spectrometry.....	- 158 -
<b>5.2.</b>	<b>Highly Disordered p53-N-Terminal Region .....</b>	<b>- 158 -</b>
5.2.1.	MS of p53-N-Terminus in buffered solution conditions.....	- 160 -
5.2.2.	Conformational Flexibility of p53-N-Terminal Region.....	- 161 -
<b>5.3.</b>	<b>Molecular Architecture of p53-N-DBD.....</b>	<b>- 163 -</b>
5.3.1.	MS of p53-N-DBD in buffered solution conditions .....	- 164 -
5.3.2.	Conformational Landscapes adopted by p53-N-DBD .....	- 165 -
<b>5.4.</b>	<b>The Quest for Stable p53 - Quadruple Mutant of p53-DBD: M113/V203/N239Y/N268D .....</b>	<b>- 168 -</b>
5.4.1.	MS of SS p53-DBD in buffered solution conditions .....	- 170 -
5.4.2.	Conformational Populations occupied by SS p53-DBD .....	- 172 -
<b>5.5.</b>	<b>Architecture of p53:Mdm2 Complex .....</b>	<b>- 175 -</b>
5.5.1.	Mdm2 N-terminal - MS and IM-MS perspective .....	- 176 -
5.5.2.	Interactions of p53-Mdm2 probed by IM-MS .....	- 180 -
<b>5.6.</b>	<b>BcL-xL-p53 Connection: Alternative Route to Cell Death .....</b>	<b>- 186 -</b>

5.6.1.	MS of BcL-xL $\Delta$ LAC in buffered solution conditions.....	- 188 -
5.6.2.	Collision cross sections of BcL-xL $\Delta$ LAC .....	- 189 -
5.6.3.	Molecular Interaction of N-terminal Domain of p53 with anti-apoptotic BcL-xL .....	- 192 -
<b>5.7.</b>	<b>Summary .....</b>	<b>- 193 -</b>
<b>5.8.</b>	<b>References .....</b>	<b>- 195 -</b>

## Chapter Six

<b>6.</b>	<b>Conclusions .....</b>	<b>- 201 -</b>
-----------	--------------------------	----------------

## Appendices

Appendix 1:	Amino Acid Abbreviations .....	- 204 -
Appendix 2:	Amino Acid R Groups.....	- 205 -
Appendix 3:	Properties of Drift Gases .....	- 206 -
Appendix 4:	Metric System Unit Prefixes .....	- 207 -
Appendix 5:	Physical Constants, Symbols and Units .....	- 208 -
Appendix 6:	Publications Arising from this Thesis .....	- 209 -

# 1

## Introduction and Motivation

*The very idea of structural disorder threatens the dogma of structural biology. The relevance of dynamically structured proteins in all kingdoms of life, in a myriad of cell functions and their significance in pathology means that conformational flexibility cannot be ignored. Mass spectrometry and ion mobility mass spectrometry can examine these heterogeneous flexible structures and their dynamics in the solvent free environment of a mass spectrometer. These techniques offer a capability to probe the conformational flexibility, binding and folding on binding and could also be adopted as a screening tool to assess the effects of interacting ligand molecules, or post transitional modifications, on intrinsically disordered protein conformations.*

## 1. Introduction and Motivation

### 1.1. The Rules of Disorder and Why Disorder Rules

The name protein is derived from the Greek '*proteos*' meaning 'of first importance' or 'primary'. According to Greek mythology, Proteus was a prophetic god of the sea who could change his form at will. '*Proteus modo se tenuabit in undas, Nunc leo, nunc arbor, nunc erit hirtus aper*' Like Proteus will melt into a sea wave, or will appear as a lion, or a tree, or a bristly wild boar (*Ars Amatoria*, 1, 755-766 di *Ovidio*). From this feature of Proteus comes the adjective 'protean' which is used to describe something extremely flexible, versatile or capable of assuming many forms. These connotations extremely well define the nature of the class of proteins that lack a unique three-dimensional structure, known as intrinsically disordered proteins (IDPs), which indeed are very 'protean' proteins. The intriguing protein family of natively unfolded proteins contradicts one of the cornerstones of molecular biology, that is, the structure-function paradigm: the concept that well-defined structure is the prerequisite of function. This notion, formulated by Fisher *et al.* (1) in 1894 as a lock-and-key model for explaining the remarkable specificity of the enzymatic hydrolysis of glucosides, has proved to be exceptionally fruitful. Uversky *et al.* (2) even considered the protein-structure paradigm as the big bang, forming the universe of modern protein science. The finding that a large fraction of proteins (over 30%) in eukaryotic cells lack a tertiary structure but are functional has forced the scientific community to review its understanding of the structure-function paradigm (3). The involvement of many of these intrinsically unstructured proteins (IUPs) in intracellular signalling and regulatory processes as well as their central positioning (as interaction hubs) in recently mapped protein interaction networks is particularly intriguing (4). The existence of IDPs has 'raised eyebrows' for a decade, but it is now becoming evident that a radical transition is taking place by 'unstructural' biology getting into the mainstream of molecular biology. Rapid growth of the field

has been marked by several excellent recent reviews (4-7) including a textbook of comprehensive coverage (8).

## **1.2. Intrinsically Disordered Proteins: definition, detection and beyond**

### **1.2.1. Definition and Unique Properties of IDPs**

Generally, IDPs can be defined as polypeptide chains that lack a unique fold, either entirely or in parts (9). Instead, they populate several heterogeneous conformations of similar energy that interconvert on a number of timescales and are more dynamic than folded globular proteins (7). IDPs and intrinsically disordered regions (IDRs) can be either collapsed (as a native molten globule), or extended (as a native coil or premolten globule) (10,11).

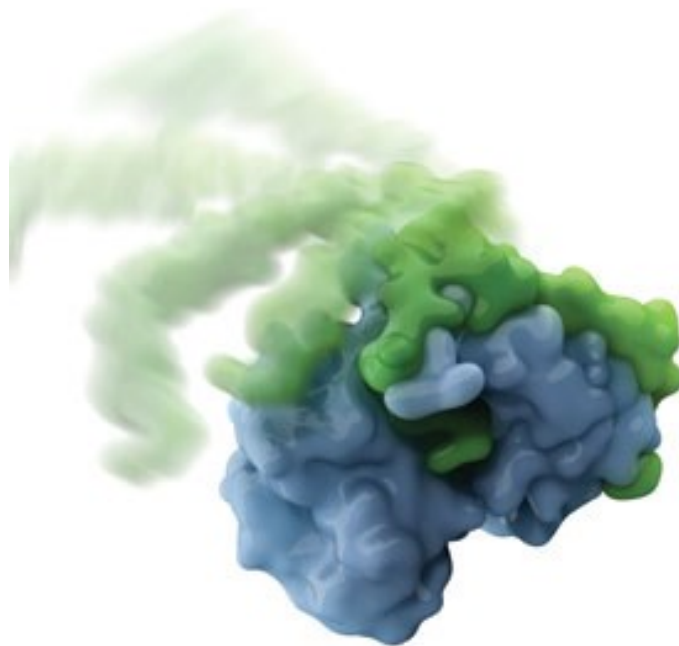
A number of terms have been used to indicate the disordered characteristics of these proteins, such as natively denatured (12), natively unfolded (13), intrinsically unstructured (14), intrinsically denatured (12), intrinsically unfolded (13), intrinsically disordered (11), floppy (15), flexible (16), mobile (17), partially folded (18), vulnerable (19), chameleon (20), malleable (21), dancing proteins (22), protein clouds (23). The most often used expression is intrinsically disordered proteins in recent literature and by a 'silent' agreement is considered as the most appropriate descriptor of the phenomenon.

A common trait of IDPs/IDRs is the low portion of bulky hydrophobic (Ile, Leu, and Val) and aromatic (Trp, Tyr, and Phe) groups, considered as 'ordering' amino acids (24). These order-promoting groups form and stabilise the hydrophobic cores of folded globular proteins that interact favourably with each other inside the fold, away from solvent. By contrast, IDPs are substantially enriched in polar (Arg, Gly, Gln, Ser, Pro, Glu, and Lys) and structure-breaking (Gly and Pro), amino acid residues considered as a source of disorder (9,25-27). These compositional differences



enabled a coarse prediction of disorder from primary sequence information alone via a numerous developed disorder predictors (28-30).

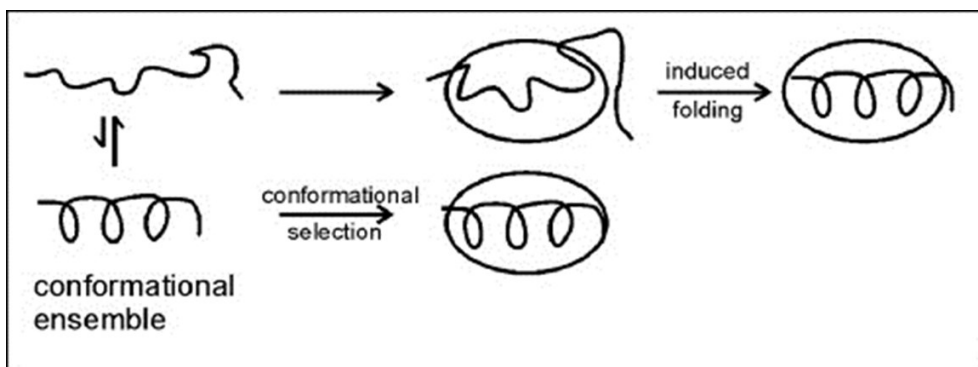
The binding of IDPs contrast with that of ordered proteins since they interact with their target via short molecular recognition features (MoRFs) which are motifs (10-70 amino acid residues) within unstructured segments, which undergo *disorder-to-order transition* upon binding to their partner. It has also been suggested by Gunasekaran *et al.* (31,32) that IDPs provide larger surface areas for intermolecular interactions than globular proteins of similar size. Such large and flexible interaction surfaces can be a major advantage and assist the assembly of intertwined multimeric complexes (33). Further, the high flexibility of intrinsically unstructured protein segments allows wrapping around their binding partners, thereby creating an extensive interaction surface (Figure 1.1).



**Figure 1.1: Folding only as needed, p27 (green) can wrap around many different partners. Illustration by AXS Biomedical Animation Studio (34).**

Many IDPs/IDRs *fold upon binding* to their physiological partner targets, a puzzling matter is whether folding occurs before binding or binding occurs before folding? Wright and Dyson *et al.* (35) proposed two extreme mechanisms: *induced folding*

and *conformational selection* (Figure 1.2). In the first process, the protein associates with its binding target in a disordered state and subsequently folds in association with the binding partner. In the conformational selection mechanism, the target protein ‘selects’ a conformational isomer which closely resembles that of the bound form from the repertoire of conformations populated by the IDP (36).

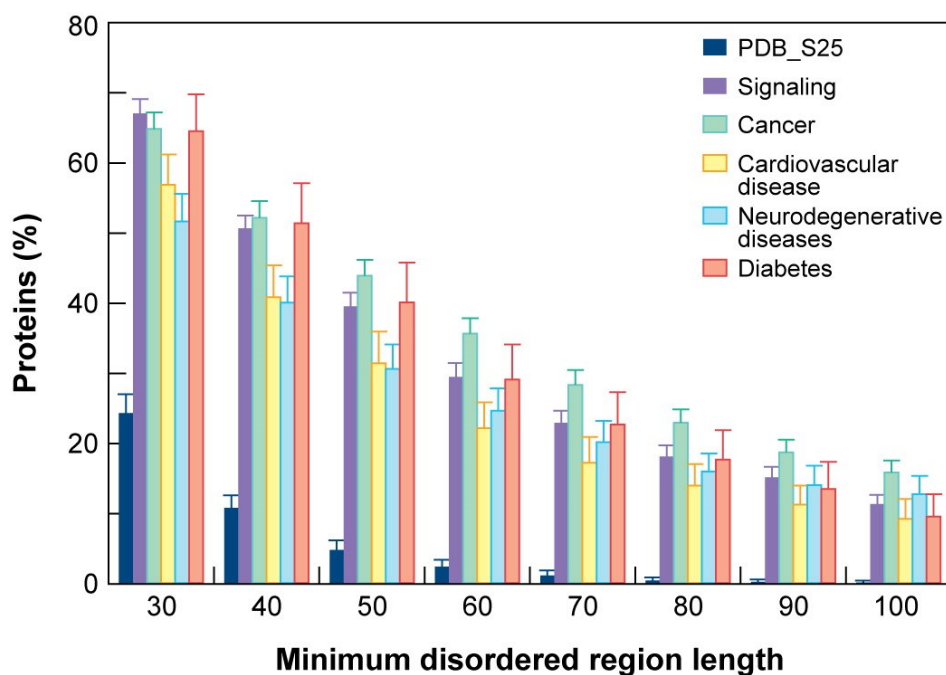


**Figure 1.2: Schematic showing two possible mechanisms of coupled folding and binding: folding upon binding and conformational selection. Image taken from Wright and Dyson *et al.* (35).**

This concept is equivalent to the induced fit-fluctuation fit (conformational selection) duality of molecular recognition. In real systems, one or other or both mechanism(s) are likely to be favoured. For many IUPs, coupled folding upon binding forms complexes with relatively low affinities, notably though, these complexes are in most cases highly specific, as their formation is facilitated by MoRFs and modulated by specific post translational modifications. The manifestation of these highly specific/low-affinity complexes could be invaluable for signalling proteins that need to be able to associate specifically to transmit a signal but also dissociate easily when conditions are altered (5,37). In contrast, according to Tompa *et al.* (38) many IDPs hardly ever become fully ordered even in the bound state, but show a significant amount of structural disorder or polymorphism in protein complexes. This phenomenon termed ‘fuzziness’ represents the extension of the paradigm of structural disorder to the functional-bound-state.

### 1.2.2. Functions of IDPs: the D<sup>2</sup> Concept

Intrinsic disorder is prevalently associated with human diseases, including cancer, cardiovascular disease, amyloidoses, neurodegenerative diseases, and diabetes, giving rise to the D<sup>2</sup> (disorder in disorders) concept (39). In 2002 Iakoucheva *et al.* (40) applied a neural network predictor of disorder to several data sets of proteins related to various diseases. 79% of cancer-related and 66% of cell-signalling proteins were found to contain predicted regions of disorder that were 30 residues or longer. The overall findings of another analysis performed by Uversky *et al.* (39) for proteins associated with cancer, cardiovascular disease and diabetes are shown below in Figure 1.3, which represents percentage of proteins with at least 30 consecutive amino acid residues predicted to be disordered.



**Figure 1.3: Abundance of intrinsic disorder in disease-associated proteins.** Percentages of disease-associated proteins with 30 - 100 consecutive residues predicted to be disordered. The error bars represent 95% confidence intervals and were calculated using 1000 bootstrap resampling. Corresponding data for signalling and ordered proteins are shown for comparison. Analysed sets of disease-related proteins included 1786, 487, 689, and 285 proteins for cancer, CVD, neurodegenerative disease, and diabetes, respectively. Figure and caption taken from Uversky *et al.* (39).

The high abundance of intrinsic disorder in diseases is due to the unique structural and functional peculiarities of IDPs and IDRs. IDPs perform crucial biological functions; they are among major cellular regulators, recognizers, and signal transducers. Their functions are tightly regulated via extensive post translational modifications and many IDPs can fold upon interaction with corresponding binding partners (partially or completely), ensuring low-affinity/high-specificity binding. They are promiscuous binders, they possess multiple binding specificity and participate in *one-to-many* and *many-to-one* signalling interactions (binding plasticity) (41).

### **1.2.3. Experimental Identification and Prediction Methods**

Intrinsic disorder in proteins can be identified and characterised by using a wide arsenal of experimental methods (42,43). Naturally, the lack of a well-defined structure in disordered proteins complicates their investigation, since the determination of the unique high resolution structure(s) of IDPs is extremely difficult. Only a strategic combination of complementary structural and biophysical techniques would allow one to decipher their mode of action at the structural level. The main techniques are briefly specified below:

Disordered states of proteins, being highly heterogeneous and dynamic, are not accessible to X-ray crystallography, however this method repeatedly defines missing electron density in many structures, which may correspond to disordered and dynamic region(s) of the protein. The increased flexibility of atoms in such region(s) leads to the non-coherent X-ray scattering, making them unobserved (9).

Heteronuclear multidimensional Nuclear Magnetic Resonance (NMR) spectroscopy is by far the most powerful technique for determination of high resolution three-dimensional structural information in solution and for the portrayal of protein dynamics. The backbone motions (in the NMR timescale) throughout the polypeptide chain can be revealed by this technique. Recent advances in NMR spectroscopy have

allowed the complete assignment of several unfolded and partially folded proteins, as well as the disordered fragments of folded proteins (2,44).

Small angle X-ray scattering (SAXS) and the small angle neutron scattering (SANS), on the other hand typically provides information on the hydrodynamic behaviour and topology of the polypeptide chain (45).

Low content of secondary structure in partially intrinsically disordered proteins can be provided by a number of spectroscopic techniques including far-ultraviolet (UV) circular dichroism (CD) (46), Fourier transform infrared spectroscopy (FTIR), Raman optical activity and deep UV Raman spectroscopy (47).

Additional data of on the intramolecular mobility and compactness of the protein can be detected via the analysis of different fluorescence methods, including fluorescence resonance energy transfer (FRET) (48).

Mass spectrometry based techniques (reviewed by Beveridge *et al.* (49)) can also be applied to the study of IDPs, including hydrogen-deuterium exchange (HDX) (50), limited proteolysis (51-53), proteomics (54) and ion mobility mass spectrometry (IM-MS), which is described in detail later in this chapter.

Bioinformatic predictions still play a decisive role in studies of structural disorder. Based on the compositional bias of IDPs, many of different predictor programs have been developed in recent years (28-30,55), including PONDR (26) (predictor of naturally disordered regions), charge-hydropathy plots (CH-plots) (24), and IUPred (56), to name a few. The available predictors can be classified into two categories (8,57): (i) predictors that are based on the amino acid sequence, (ii) machine learning algorithms that are trained on the existing experimental data sets which indicate the disorder (e.g. missing electron density in crystallographic data). One of the key findings of these computational approaches to the completely sequenced genomes is that eukaryotes exhibit a significantly larger portion (up to 30%) of their proteome that is intrinsically disordered than prokaryotes (3).

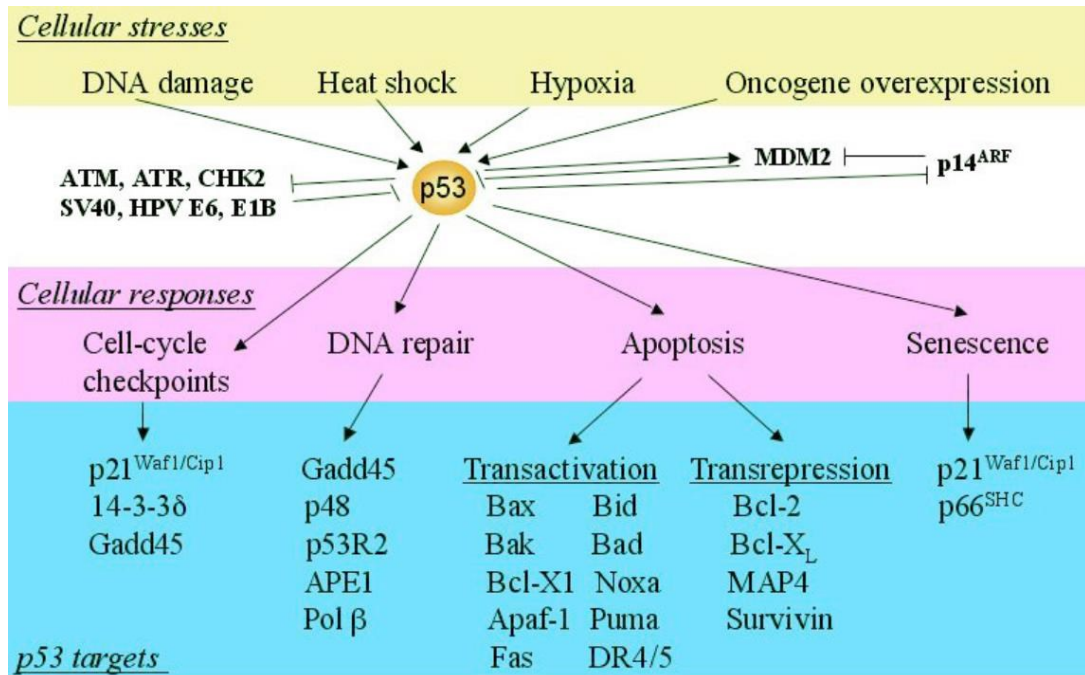
Even if feasible, determination of a single set of three-dimensional atomic coordinates, may not be instructive for a highly disordered protein. Rather, the objective is to identify the rules that define the dynamic conformational behaviour, particularly in terms of an ensemble description of interconverting structures and determining timescales that are involved in the conformational transitions in order to develop a complete picture. IDPs inhabit a vast conformational landscape, and mapping of this intricate conformational energy landscape requires an exploitation of complementary experimental techniques.

### **1.3. The Tumour Suppressor p53: Puzzle and Paradigm**

p53 belongs to the growing family of loosely folded or partially unstructured native proteins (58). Research on the p53 protein is at fever pitch. In the past couple of years it has become evident that inactivation of its tumour-suppressor activity is an almost universal step in the formation of human cancers (59). But even thirty years after its discovery by Levine *et al.* (60), the multi-functional tumour suppressor p53 is still somewhat of an enigma. p53 nicknamed the ‘guardian of the genome’ is at the hub of a plethora of signalling networks by a multitude of protein-protein interactions that regulate the cell cycle and maintain the integrity of the human genome, and modulated by extensive post translational modifications. In response to oncogenic or other cellular stress stimuli, p53 is activated and prompts up- or down regulation of genes involved in cell cycle arrest, DNA repair, senescence, or apoptosis (61-63). It is therefore not surprising that this functional complexity is mirrored in its structure.

p53 protein has been extensively studied leading to a significant number of discovered protein-protein interactors; so far 485 partners according to BioGRID database, 266 direct protein interactors according to the Human Protein Reference Database, 303 according Interologous Interaction Database (SwissProt ID: P04637, 315 from POINeT database, have been brought forward by different experimental methods (Figure 1.4). Both transactivation and C-terminal domain are particularly





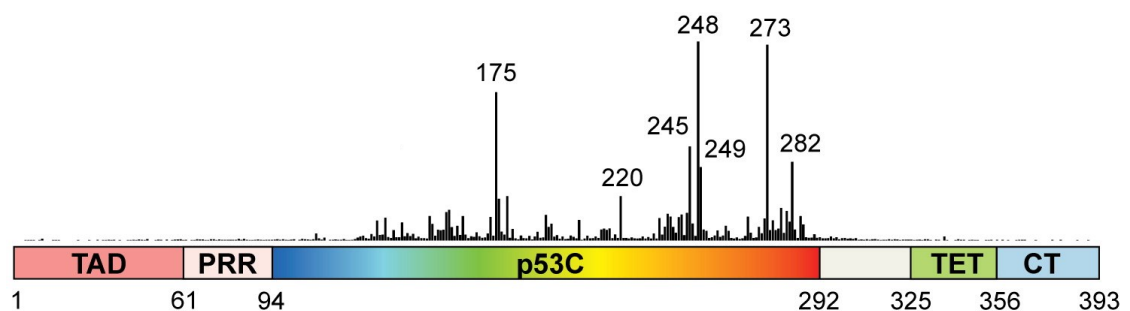
**Figure 1.5: p53 locating at the crossroads of complex networks of stress response pathways. Various intercellular or extracellular stresses elicit cellular responses directly or indirectly through p53 activation. p53 activates its downstream targets to perform various functions including cell cycle arrest, DNA repair, apoptosis, and senescence. Figure and caption taken from Bai *et al.* (68).**

As a result of efforts to identify targets of p53 in various experimental systems, hundreds of physiologically p53 responsive genes have been reported, such as genes for p21<sup>Waf1/Cip1</sup>, Gadd45 (growth arrest and DNA-damage-inducible protein) and genes of the Bcl-2 family (63). Elevated levels of p53 can also cause repression of genes which include cyclin B1, MAP4, bcl-2, bcl-x, and survivin, some of them are negative regulators of apoptosis (63). The functions of p53 target genes are diverse, corresponding to p53's activity as a multifunctional protein.

p53 functions as a homotetramer comprising 4 x 393 amino acid residues. It has a modular domain structure, consisting of independently folded DNA-binding and tetramerization domains, flanked by natively unfolded regions at both the amino- and carboxy-termini, which poses a challenge to the structural biologists (Figure 1.6). The N-terminal region consists of an intrinsically disordered transactivation domain



(TAD, residues 1-61) that induces transcriptional activation of genes regulated by p53, followed by a proline-rich region (PRR, residues 62-94), and a central folded DNA-binding domain (DBD, residues 94-292) that is responsible for recognition of the consensus sequences associated with p53 target gene promoters. This domain is linked to a tetramerization domain (TET, residues 325-356) that regulates the oligomerization state of p53. At its C-terminus (CT, residues 356-393) p53 contains a strongly basic regulatory domain (mainly lysines) which binds DNA non-specifically. Although the core DBD forms a so called 'ordered' region of p53, and possesses many disordered loops and it is in these that most p53 inactivation mutations associated with nearly 50% of human cancer occur (69). Overall, approximately 70% of p53 interactions are mediated by intrinsically disordered regions (IDRs) (70). A bias toward intrinsic disorder is particularly expressed in the sites of posttranslational modifications, with 90%, 86%, and 100% of detected phosphorylation, acetylation, and protein conjugation sites, respectively found in IDRs (70).



**Figure 1.6: Domain structure of full-length p53 consisting of an N-terminal transactivation domain (TAD), followed by a proline-rich region (PRR), the central DNA-binding domain (p53C), the tetramerization domain (TET), and the extreme C-terminus (CT). p53C, the domain where most cancer-associated p53 mutations are located, is highlighted with a rainbow color gradient from blue at its N-terminus to red at its C-terminus. The bars above the diagram indicate the relative frequency of oncogenic missense mutations at each residue according to version R11 (October 2006, N = 17,015) of the International Agency for Research on Cancer TP53 mutation database. Schematic and caption taken from Joerger *et al.* (71).**

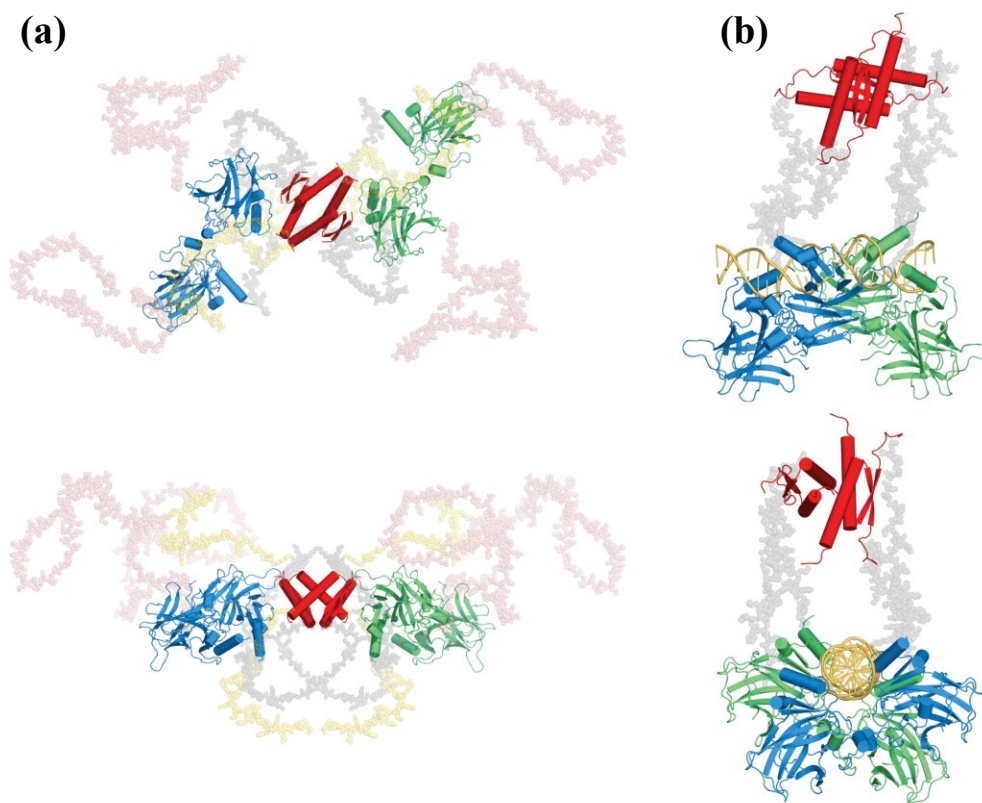
In the mid-1990s, several pioneering studies elucidated structural details of individual modules of p53 structure, the first structures to be solved were the core

DBD in a complex with DNA (72), and the tetramerization domain (73-76). Shortly after, atomic structure of a transactivation domain in a complex was solved (77), and recently fragments of the C-terminal regulatory domain also in a complex with binding partners were determined (78-80). But in the decade that followed, relatively small progress was achieved in revealing the structural foundation of p53 functionality or its inactivation in tumour, considering that the scientific literature has been inundated with p53-associated publications.

### **1.3.1. The Bigger Picture: Quaternary Structure of p53**

Only in recent years has a full grasp of the quaternary structure of p53 and its dynamic behaviour as a whole has become possible, owing to combination of sophisticated computational methods, innovative protein engineering and classical structural biology. Many structural studies on the full-length p53 have been hindered both by its intrinsic instability and the presence of large regions of intrinsic disorder. A rapidly growing number of identified proteins have globular domains connected by intrinsically disordered regions, and so solving such structures is a recurring problem. There are particular obstacles in tackling the full-length structure of tumour suppressor p53; the protein is too flexible for crystallographic measurements, and too big for conventional NMR spectroscopy. A multi-technique strategy was applied by Tidow (81) and Wells *et al.* (82). The quaternary structure of the stabilized full-length p53 and different domain constructs in their free form and in complex with response element DNA was determined by using state-of-the-art NMR spectroscopy to identify domain-domain interfaces, SAXS experiments and electron microscopy (EM). The X-Ray crystal structures of individual p53 domains were modelled as rigid bodies to fit the SAXS and EM data, missing residues were modeled *ab initio*. Pivotal to this work was Nikolova's design of an engineered stabilized variant of p53 core domain M133L/V203A/N239Y/N268D that considerably improved the overall stability of the protein (83,84).

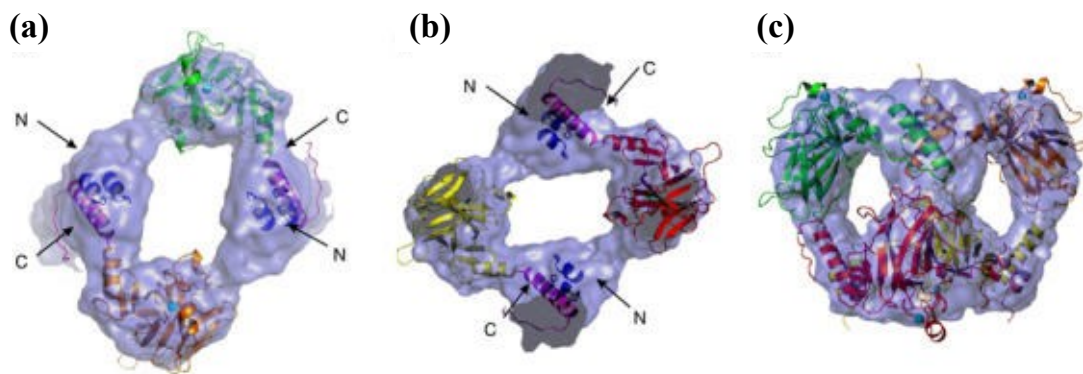
According to Tidow *et al.* (81), the free p53 protein in solution forms an elongated cross-shaped structure with the tetramerization domain at its center, loosely assembled core domain dimers and extended N and C termini (Figure 1.7). In this open conformation, the DBDs can interact freely with cognate DNA and partner proteins in the cell cycle.



**Figure 1.7: Rigid body models of free (a) and DNA-bound (b) p53 in solution from small-angle X-ray scattering data (81). Both models are shown in two different orientations. Core (blue and green) and tetramerization (red) domains are shown as cartoon representations, with core domains binding to the half-site in the same colour. Flexible connecting linkers (gray), N termini (pink), and C termini (yellow) are shown as semi-transparent space-filled models. Models of N and C termini in the free protein are approximations to illustrate the space occupied by these flexible regions, rather than representing defined conformations. Image and caption taken from Joerger *et al.* (71).**

EM images for full-length p53 are heterogeneous in shape and size, indicating an ensemble of different, predominantly open conformations (81). A radically different structural model of the unbound full-length murine p53 protein was proposed by

Okorokov *et al.* (85). This cryo-electron microscopy study reported a closed, compact structure resembling a hollow skewed cube, in which the tetramer is formed through interactions via its N and C termini (Figure 1.8). Upon DNA binding (Figure 1.7b), p53 wraps around the DNA double helix and becomes more compact. Domain rearrangements to allow DNA binding are facilitated via the flexible linker between the DNA-binding and tetramerization domain, and the overall structure becomes more rigid.



**Figure 1.8: Structural organization of p53.** (a) The upper density layer of the p53 three-dimensional map, with fitted structures of two core domains shown in green and orange. N and C represent positions of the N- (blue) and C-termini (magenta), respectively. (a) Cut away view of the lower density layer with the second pair of core domains (coloured red and yellow) fitted in. (c) Domain structures are fitted into the corresponding nodes of the tetramer. Image and caption adapted from Okorokov *et al.* (85).

This process is in accord with the X-ray crystal structures of the tetrameric p53-DNA complexes by Kitayner *et al.* (86). The N termini are extended in both the free and DNA-bound p53 modeled structures by Tidow *et al.* (81), consonant with the fact that they interact with myriad of signaling proteins and are subjected to extensive post translational modifications (87,88).

Not only is this complex (un)structural biology vital for full understanding the role of p53 in the cell cycle and in other activities, but the structures may also help in the design of drugs that affect the cell cycle.

### 1.3.2. Post-translational Modifications of p53

In response to stress levels, p53 participates in complex networks of covalent post-translational modification, including phosphorylation, acetylation, ADP-ribosylation, ubiquitylation, sumoylation, neddylation, and cytoplasmic sequestration (63,89). Post-translational modifications affect protein activity, stabilization, localisation, turnover, and interaction properties, governing fully the protein function. Recent studies have shown that most of these modifications occur in unstructured parts of proteins, mainly in the N- and C-terminal regions. Phosphorylation and acetylation are the key modifications improving the transcription activating ability of p53 because these modifications result in p53 stabilization and accumulation in the nucleus, where p53 interacts with sequence-specific sites of its target genes (90). Most of the phosphorylation sites have been identified in the N-terminal domain at the Ser6, Ser9, Ser15, Thr18, Ser20, Ser33, Ser37, Ser46, Thr55, and Thr81 residues, but there are also other sites in the C-terminal domain at Ser315, Ser366, Ser371, Ser376, Ser378, Thr387, and Ser392, and also in the central core DNA-binding domain at Ser149, Ser150, Ser155 (89,91). Many protein kinases have been implicated in phosphorylating events of p53, including ATM, ATR, Chk1, Chk2, JNK and p38 (92-94). p53 is also targeted for degradation via Mdm2 via ubiquitination of specific lysine residues localised within its C-terminus. In addition p53 can be modified by ubiquitin like proteins, NEDD8, and SUMO which regulate the p53-Mdm2 interaction and protein stability. Acetylation takes place on residues used for ubiquitination thus resulting in p53 stabilization as a consequence blocking Mdm2 catalyzed ubiquitination (95). Not only p53 is activated by blocking Mdm2 ubiquitination, but acetylation induces conformational changes within p53 which stimulates its sequence-specific DNA binding activity (95,96). Two histone acetyltransferases are known to acetylate p53: p300 or/and CBP acetylates the C-terminus of p53 at Lys305, Lys372, Lys373, Lys381, and Lys382, whereas PCAF (p300/CBP-associated factor) acetylates Lys320 (95,97). These various modifications form a 'binding code', as they alter the local charge density and hydrophobicity, affecting the structure and folding properties of the local chain and its surface (4). Generally, post-translational modifications provide dynamicity in the

protein-protein interactions and can adapt the function precisely in in space and time scales to achieve diverse and specific biological cell outcomes.

## **1.4. p53:Mdm2 - the Affair that Never Ends**

The interaction between p53 and Mdm2 (murine double minute 2) represents the best-studied relationship between a tumour suppressor protein which acts as a transcription factor and an oncoprotein, which functions primarily as an E3 protein ligase (also called a ubiquitin ligase, that combines with a ubiquitin containing E2 ubiquitin-conjugating enzyme, recognizes the target protein and causes the attachment of ubiquitin to a lysine on the target protein) (98). This intimate relationship is tightly regulated by a complex arsenal of post-translational modifications, which in turn dictates the stability and activity of p53 and Mdm2. Mdm2 down regulates the activity of p53 via a negative feedback loop through binding to the N-terminal transactivation domain of p53 (99). In addition to blocking the transcriptional activity of p53, Mdm2 also exports p53 from the nucleus into the cytoplasm and promotes it for proteasomal degradation via ubiquitination (100-102). The E3 activity of Mdm2 is centered within the RING finger domain (103). The ability of Mdm2 to associate with, and induce p53 degradation is highly dependent on the phosphorylation status of p53, as well as on the interaction of p53 with other cellular proteins. In response to cellular stress, when p53 phosphorylation occurs on multiple serine residues in its N-terminal transactivation domain, including those spanning Mdm2 binding, the association of p53 with Mdm2 is no longer possible (104,105). Amplification of the Mdm2 gene is observed in approximately 7% of human cancers, with a much higher incidence in soft tissue tumours (20%), osteosarcomas (16%) and oesophageal carcinomas (13%) (106). Overexpression of Mdm2 effectively abolishes p53 function, and in turn tumours become less susceptible to chemotherapeutic agents and do not undergo programmed cell death, or apoptosis. The myriad of roles which p53 and Mdm2 play within the cell led to the realization that the p53-Mdm2 interaction is a 'druggable' target (77). A series of p53-Mdm2 antagonists have been developed including 'stapled' peptides (107), cis-

imidizalones (108), meta-chloro piperidinones (109), benzodiazepinediones (110) and chromenotriazol- pyrimidines (111).

Mdm2 itself is a multi-domain 55 kDa protein consisting of 491 amino acids. Overall it is composed of an N-terminal domain with a hydrophobic pocket, a disordered central acidic domain containing a nuclear localization signal, a zinc finger, and a C-terminal really interesting new gene (RING) which is responsible for its ligase activity (112).

## **1.5. The Case for Mass Spectrometry as a Tool to Study IDP's**

Knowledge of the active conformation and interaction partners of proteins and their complexes is paramount to understanding their intercellular function. Many protein systems are dynamic with respect to both conformation and interaction, so in the course of a protein's 'lifespan' it will sample many configurations and bind to several targets. This is particularly true for IDP's which have crucial roles in molecular recognition, assembly, protein modification and entropic chain activities, but often intrinsic structural disorder until required to function. It is desirable therefore to observe the range of structures an IDP can occupy, and also the effect of selected binding partners on altering this conformational space. In the past 10 years, a substantial body of work on IDP's has been performed *in silico*, but due to their flexibility and 'fuzziness' they have been less examined as a class of proteins by structural characterisation (39). Since the solvent free environment of a mass spectrometer is ideally suited to the study of intrinsic interactions and how they contribute to structure, it is a very exciting venue for the study of IDP's.

Mass Spectrometry has established itself as a *bone fide* method for structural biology (113-115). Pioneering work has shown that bioactive complexes of significant size can be transported intact into the controlled environment of a mass spectrometer for subsequent analysis (116). There is of course controversy about whether a solution

phase structure is retained in its entirety in the solvent free environment of a mass spectrometer, but for large macromolecular systems bound by many non-covalent interactions, there is evidence to suggest that macroscopic features of solution and even *in vivo* structures are retained (117). When applied to IDP's, the fact that a range of conformations can be observed for a given protein system depending on the method by which it is ionised is a rather exciting prospect. This control and selectivity over the way in which a given protein that is intrinsically disordered can be studied means that its inherent stability and how that is affected by binding partners can be probed. The growth and success of studies of macromolecular complexes by mass spectrometry increasingly places biological mass spectrometry as the first step in structural analysis of an unknown protein and of its targets; in short, it now has a role as a rapid information rich predictive tool which has a particular suitability for IDP's.

Careful optimisation of the nano electrospray ionisation (n-ESI), source optic voltages and pressures allows for the non-covalent interactions to be preserved and a gentle transfer of intact macromolecules from solution into the solvent-free environment of a mass spectrometer. Following n-ESI, all proteins are present in a number of different charge states, commonly as (multiply) protonated species (for positive ionisation) or (multiply) deprotonated (negative ionisation) - this is termed a charge-state distribution (CSD). Structured, folded proteins with little conformational flexibility in solutions where the protein is buffered appropriate to its pI, give rise to gaseous ions carrying a relatively low number of charges, presenting a narrow charge state distribution centered on high values of  $m/z$  (118). By contrast, less structured or unfolded proteins can accommodate a larger variation in the number of protons related to the varied availability of ionisable sites on the protein surface, in turn a feature of the conformational flexibility in solution (119). Analysis of the relative intensities of ions in CSDs enables an indirect assessment of conformational families present in solution. Frimpong *et al.* (120) used CSD analysis following ESI-MS to determine the conformational heterogeneity of monomeric  $\alpha$ -synuclein, a neuronal protein which has a special relevance for understanding Parkinsons disease (121). Although, this protein is highly disordered, it was found to populate four distinct



conformational states, ranging from a compact, highly structured one, to a random coil (120). IM-MS provides another dimension to gas-phase studies of biomolecules garnering more direct information on the molecular species ‘shape’ (collision cross-section, in  $\text{\AA}^2$ ) along with the CSD, but does not provide atomic resolution compared to NMR and X-ray crystallography. IM-MS has been successfully applied to examine protein dynamics (118,120), protein-ligand and protein-protein interactions (122) and shows great promise for the structural characterisation of IDPs. Bernstein *et al.* (123) performed experiments on  $\alpha$ -synuclein which combined n-ESI-MS with IM measurements. Their findings indicate the presence of compact structures when spraying from low pH solutions, while higher charge states occurring at physiological pH (6.8) have larger collision cross sections and are relatively unfolded. The transition from compact to extended conformations occurs from charge state  $z = -8$  to  $z = -9$  where a 50% increase in cross section is observed. IM-MS has also been utilized to study the structure of the DNA binding domain of the tumour suppressor protein p53 with and without the functional zinc ion (124). Other research using IM-MS to study IDPs has examined changes in conformational equilibrium of the intrinsically disordered High Mobility Group A (HMGA) chromatin factors, due to post translational modifications or sequence deletion (125), and the ability of IDPs to fold on binding to a partner using the salivary tannin binding proteins (126).

A variety of configurations of IM instrumentation have been developed: drift tube (DT) IM-MS, travelling wave (TW) IM-MS and field asymmetric waveform ion mobility-spectrometry (FAIMS). These techniques are briefly described in the following section.

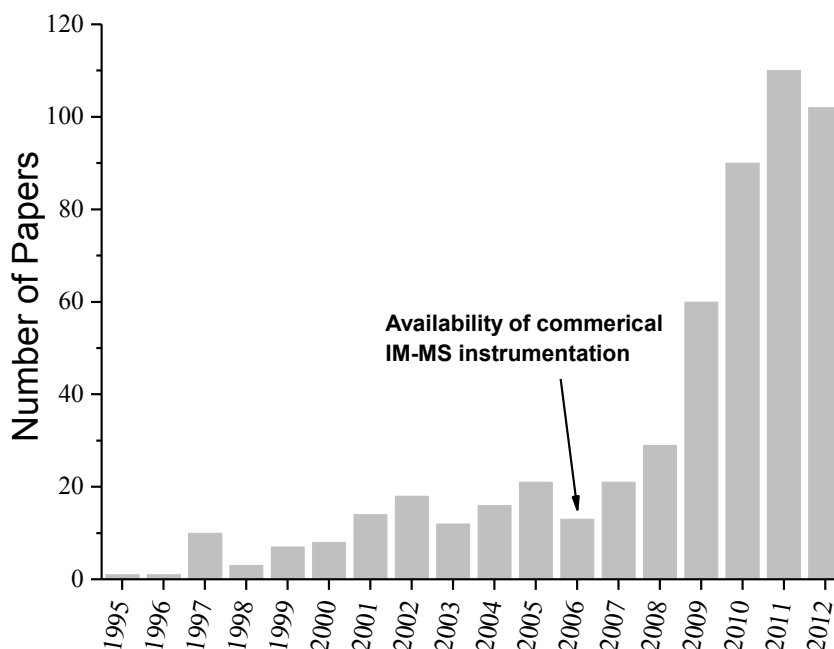
## **1.6. Ion Mobility Mass Spectrometry - to determine Conformation(s)**

The start of the analytical technique termed ‘ion mobility spectrometry’ (IMS) can be traced to the beginnings of the 20<sup>th</sup> century, when experimental (127,128) and theoretical (129) physicists became interested in the movement of ions in gases. In

these early studies, as in all modern ion mobility methods, what was measured is the velocity of ions in a mass spectrometer where they ‘drift’ due to the presence of a weak electric field while experiencing collisions with a buffer gas. Following the experiments of Erikson (127) and Bradbury (128), in the 1950’s McDaniel (130,131) constructed low field drift tubes which bear a great resemblance to the Drift Time IMS mobility instrumentation still used today. In 1965 Kebarle coupled such a drift tube to a mass spectrometer creating the hybrid technique of ion mobility mass spectrometry (IM-MS) (132,133). This work prompted the development of different configurations of ion mobility and mass spectrometry instrumentation over the past 40 years, and their application to an ever wider range of analyte species (134). Details of the early parallel development of ion mobility and of mass spectrometry and its applications to large macromolecules can be found in an excellent review by Uetrecht *et al.* (135) which focusses on applications, and one on instrumentation in a comprehensive summary by Kanu *et al.* (134). In the past 5 years commercial instrumentation for IM-MS (136) has been developed and the technique has become more widely applied as a research tool. Despite differences in the configuration of mobility devices there are some features common to all IM-MS experiments. A typical mobility experiment requires the injection of a pulse of ions into a chamber filled with a known gas at a known pressure across which is applied an electric field. The time taken for the ions to pass through the chamber is measured. Upon injection into the chamber or drift cell, the ions experience an electrostatic force pulling them through the cell; this force is countered by collisions between the ions and the buffer gas. The behaviour of an ion moving through a gas under the influence of an electric field ( $E$ ) is dependent on its energy which is determined by the ratio of electric field strength to buffer gas number density,  $E/N$  (137). At low  $E/N$  the ions are said to be in the low field limit, under this regime the ions have low velocities which are independent of the field strength. At higher  $E/N$  the ions may align in the field and their motion becomes dependent on the field strength. In the low field limit, the motion of the ions can be described in more simple terms and most ion mobility measurements are therefore performed in the low field limit. Under these conditions the low field mobility is defined, ( $K$ ) as the constant of proportionality between drift velocity, ( $v_d$ ) and electric field,  $E$  as shown by Equation 1.1.

$$v_d = KE \quad \text{Equation 1.1}$$

For experiments performed at sufficiently high pressure, and low drift voltages, ions will quickly be thermally equilibrated and reach a constant drift velocity  $v_d$  which is determined by the magnitude of the electric field ( $E$ ) and its influence on the charge of the ion as well as by the number of collisions that the ion makes with the buffer gas. Hence, the mobility,  $K$ , must contain terms relating to molecular shape, charge and the buffer gas pressure. These latter properties combine to provide the rotationally averaged collision cross section ( $\Omega$ ) for each ion which is critically dependent on the chosen buffer gas as well as on its temperature (see Equation 1.2). Hence determining  $v_d$  elucidates  $\Omega$  which can be interpreted as the shape adopted by a given molecular ion under particular gas phase conditions. To determine  $v_d$ , and hence  $K$  and  $\Omega$ , accurately in ion mobility instrumentation, the length of the drift region must be known precisely and then the average arrival times of ions can be recorded at a detector post the drift region. In IM-MS instrumentation, detection will also provide mass resolution and so the arrival time for a given mass to charge ratio is found. Most laboratories will provide  $K$  in terms of the reduced mobility  $K_0$ , normalised for pressure and temperature. In recent years IM-MS coupled with the soft ionisation technique electrospray ionisation (ESI) has gained importance as a tool for the analysis of macromolecules and particularly for its application in determining the conformations adopted by biological molecules (123,138-141). Efforts in commercial development have concentrated on improving the low duty cycle that can be significant in linear ion mobility experiments, and also on increasing the resolution obtained. Initial studies focused as much on structural measurements as on mixture separation, reflecting the growing potential of mass spectrometry as a tool to provide detailed conformational information on biological moieties (113). The use of IM-MS to study proteins and protein complexes has been particularly rapid. Figure 1.9 illustrates this, the number of publications as searched for in the ISI Web of Knowledge database containing the words ‘protein + ion mobility mass spectrometry’ was plotted versus the year of their publication.



**Figure 1.9: Results from a search on ISI Web of Knowledge for the words protein + ion mobility mass spectrometry appearing in title or abstract fields. The number of published articles per year is plotted. The graph starts with the first hit (142).**

Of note on this figure are the inflections when Clemmer developed instrumentation in 1997 and then more recently when the Synapt instruments were launched in 2006. There are three principal types of ion mobility instrumentation that have been successfully combined with mass spectrometry.

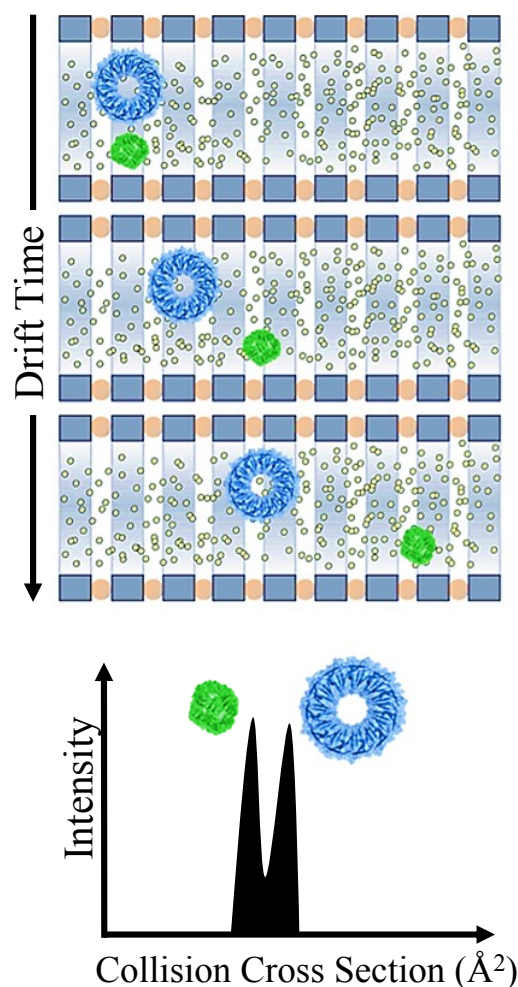
### **1.6.1. Drift Tube Ion Mobility Mass Spectrometry**

Following from the apparatus used by Kebarle, the simplest configuration of IM-MS instrumentation consists of a tube located within a vacuum system filled with a buffer gas of interest across which is applied a weak electric field ( $5 - 100 \text{ Vcm}^{-1}$ ). Ions are steered and injected into a small orifice at one end of the drift tube. They drift due to the field and collisions with the buffer gas until they exit through another small orifice. By measuring the drift velocity of an ion through such a linear drift field and solving Equation 1.1, it is possible to determine its collision cross section

(CCS) with some degree of accuracy (137) according to the relationship in Equation 1.2:

$$K_0 = \frac{3ze}{16N} \left( \frac{2\pi}{\mu k_B T} \right)^{1/2} \frac{1}{\Omega} \quad \text{Equation 1.2}$$

where  $K_0$  is the reduced mobility (the measured mobility  $K$  standardised for pressure and temperature to 273.15 K and 760 Torr),  $z$  is the ion charge state,  $e$  is the elementary charge,  $N$  is the gas number density,  $\mu$  is the reduced mass of the ion-neutral pair ( $\mu = \frac{m_1 m_2}{m_1 + m_2}$ , where  $m_1$  is the mass of the ion, and  $m_2$  is mass of the drift gas)  $k_B$  is the Boltzmann constant and  $T$  is the gas temperature. As shown here, the mobility of an ion is inversely related to its rationally-averaged collision cross section. In other words, a more compact ion will undergo fewer collisions and hence travel faster through the drift cell than a more extended ion (Figure 1.10). This is used to separate ions of the same mass but different conformations.



**Figure 1.10:** In the schematic, two ions of the coincident mass to charge ratios are resolved by DT IM-MS. The more extended conformer (blue) takes longer to traverse the drift cell because it collides more often with the buffer gas than the compact (green) conformer. Image adapted from the website of Ruotolo Research Group at the University of Michigan (143).

This experimental CCS ( $\Omega$ ) which is buffer gas dependent (144) can be compared to cross sections derived from co-ordinates obtained from other structural investigations (for example crystal structure co-ordinates), or from computational predictions, to obtain atomically detailed conformational information.

Several IM-MS instruments which incorporate drift cells over which a linear field gradient is applied have been reported. In particular there have been notable efforts made in the ‘modern’ era of IM-MS in designing and building DT IM-MS instrumentation by Bowers with Kemper (145,146), and by Jarrold (147,148),

Clemmer (149-154), and Hill (155,156). In a seminal early study, Clemmer and Jarrold used a modified ion mobility-time of flight (TOF) mass spectrometer equipped with an ESI source for analysis of small organic and biological molecules, in particular bradykinin as well as the larger proteins cytochrome c and ubiquitin. Another ESI capable IM-MS instrument IM-Quadrupole reported by Wyttenbach, Kemper and Bowers (157), has been applied to the study of an impressive range of biologically relevant molecules including the amyloidogenic proteins alpha-synuclein (123) and the amyloid  $\beta$  protein (158), PNA/dsDNA complexes (159) and a systematic study of variants of the hormone GnRH (160). The studies from this group are frequently accompanied by extensive MD simulations of the analytes of interest, allowing comparison between calculated and measured CCSs. The Clemmer group at Indiana University have been responsible for a number of notable advances in instrument and technique development for the study of large proteomic datasets, (161,162) including multidimensional ion mobility (149,163-165). In our group an ion mobility mass spectrometer, capable of performing temperature variable measurements of CCSs has been built and applied to the examination of the conformations of biological molecules and complexes (166). McLean *et al.* have coupled a DT IM- MS with imaging mass spectrometry (167) - the results show the spatial localization of lipids in biological tissues.

### **1.6.2. Travelling Wave Ion Mobility Mass Spectrometry**

Waters MS Technologies introduced the first commercially available integrated IM-MS instrument the Synapt HDMS in 2006 (136). With its origins in the pioneering work of Gerlich (168) this mobility separator consists of three stacked ring ion guides (called the Triwave) with RF applied to consecutive electrodes; to propel ions through the device, a travelling wave (TW) comprising a series of transient DC voltages is superimposed on top of the RF voltage. In 2009, Waters launched the second generation Synapt G2 System which improves some of the design. In the short time these instruments have been available, they have made a significant impact on the use of IM-MS for the study of macromolecular systems and in

particular on the study of intact proteins and protein complexes. The mobility of ions through these travelling wave devices is not as directly related to collision cross section as in DT IM-MS instruments. Shvartsburg *et al.* (169) have made notable efforts to formally understand the relationship between collision cross section and mobility in a travelling wave device, but as yet there is no adoption of such an approach by the IM-MS community. Despite this, with a careful use of standards previously measured in helium via DT IM-MS, it is possible to convert mobilities measured with nitrogen as a drift gas in Synapt instrumentation into mass selected 'helium based' collision cross sections (170). Using Synapt HDMS system, Ruotolo *et al.* (171) have assessed conformations of multimeric proteins, and also the disassembly of complexes viewing the partial unfolding of monomer units whilst still retaining some of the integrity of the complex (172). Smith *et al.* (173) has used TW IM-MS methodology to examine prefibrillar aggregates of amyloidogenic proteins. Scrivens has applied Synapt technology to several applications for biomolecular science including a study of the protein haemoglobin, (174) and with Bowers a study of the calcium binding protein calmodulin. Scrivens *et al.* (175) has also employed this technique to extract the spectra of N-glycans released with PNGase F from a serial titration of recombinantly expressed envelope glycoprotein, gp120, from the human immunodeficiency virus. Heck and co-workers have examined very large species with TW IM-MS including chaperonin complexes (176) and intact virus particles (177). Oldham has performed promising experiments which apply IM-MS to the stabilisation of proteins by ligands (178). As can be seen in Figure 1.7 the number of publications which employ IM-MS to examine proteins has grown significantly since the introduction of the Synapt instruments by Waters. Only a flavour of the range of biological systems amenable to this analysis has been provided here, rather than an extensive list.



### **1.6.3. Field Asymmetric Waveform Ion Mobility Mass Spectrometry**

The final type of ion mobility separation, commonly used in conjunction with mass spectrometry is field asymmetric waveform ion mobility spectrometry (FAIMS) or differential mobility spectrometry. First developed in the USSR in the early 1980's, (179) it was successfully coupled with ESI-MS by Purves, Gruevremont and co-workers in the late 1990's (180,181). The FAIMS/MS platform was made commercially available in 2003 as a front-end device that could be coupled to a number of mass spectrometers. In this technique ions pass between two electrodes in the presence of a tangential gas flow. FAIMS devices are usually placed at the source end of mass spectrometers where the gas flow may well be due to the drying gas requirements of an electrospray source. The voltages applied to the electrodes and the speed of the gas flow can be tuned such that ions of a specific mobility  $K$  are transmitted successfully through the device. This said, separation in FAIMS is independent of the absolute value of  $K$ , ions are separated by the difference in their mobilities at high and low values of  $E$ . Ashcroft and co-workers have employed a front end mobility separator available from Thermo-Electron coupled to a quadrupole TOF instrument to examine conformers of  $\beta$ -microglobulin (182). The group of Smith have been extremely successful in both the implementation with respect to separation of complex mixtures (183) and the interpretation of FAIMS (184). From this group, of particular note is a recent report which discusses the pendular alignment of proteins in gases under the influence of electric fields which has great promise as a new analytical approach for studying protein unfolding and for conformer selection of large systems (185). Xuan *et al.* (186) has also used FAIMS to help separate proteomic mixtures, with an emphasis on elucidation of phosphopeptides. This latter study and the work of Smith *et al.* (183) probably best represent the strengths and future applications of the FAIMS approach to ion mobility as an analytical technique for separation of mixtures, essentially gas phase chromatography, rather than for precise conformational detail.

## **1.7. Summary**

Mass spectrometry is a technique that can sample the broad repertoire of structural and functional states occupied by IDP's. MS informs on all levels of protein organization, and enables quantitative assessments of their intrinsic dynamics in the gas phase. The key advantages of MS are that it is a sensitive separation technique with wide applicability, and thereby allows the individual interrogation of transient conformers and oligomers comprising a heterogeneous ensemble. In this thesis MS and IM-MS is used to provide insights into the disordered nature of tumour suppressor, multi-domain transcription factor p53 protein.

## 1.8. References

1. Fischer, E. (1894) Einfluss der configuration auf die wirkung der enzyme. *Berichte der deutschen chemischen Gesellschaft*, **27**, 2985-2993.
2. Uversky, V.N. (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Science*, **11**, 739-756.
3. Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F. and Jones, D.T. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *Journal of Molecular Biology*, **337**, 635-645.
4. Gsponer, J. and Madan Babu, M. (2009) The rules of disorder or why disorder rules. *Progress in Biophysics and Molecular Biology*, **99**, 94-103.
5. Dyson, H.J. and Wright, P.E. (2005) Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology*, **6**, 197-208.
6. Dunker, A.K., Silman, I., Uversky, V.N. and Sussman, J.L. (2008) Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology*, **18**, 756-764.
7. Uversky, V.N. and Dunker, A.K. (2010) Understanding protein non-folding. *Biochimica et Biophysica Acta*, **1804**, 1231-1264.
8. Tompa, P. (2009) *Structure and function of intrinsically disordered proteins*. CRC Press (Taylor and Francis Group), Boca Raton, FL.
9. Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.R., Hipps, K.W. *et al.* (2001) Intrinsically disordered protein. *Journal of Molecular Graphics & Modelling*, **19**, 26-59.
10. Uversky, V.N. (2003) Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? *Cellular and Molecular Life Science*, **60**, 1852-1871.
11. Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W. *et al.* (2001) Intrinsically disordered protein. *Journal of Molecular Graphics and Modeling*, **19**, 26-59.
12. Schweers, O., Schonbrunn-Hanebeck, E., Marx, A. and Mandelkow, E. (1994) Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *Journal of Biological Chemistry*, **269**, 24290-24297.
13. Weinreb, P.H., Zhen, W., Poon, A.W., Conway, K.A. and Lansbury, P.T., Jr. (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*, **35**, 13709-13715.
14. Tompa, P. (2002) Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, **27**, 527-533.
15. Lee, J., O'Kane, D.J. and Visser, A.J. (1985) Spectral properties and function of two lumazine proteins from photobacterium. *Biochemistry*, **24**, 1476-1483.
16. Pullen, R.A., Jenkins, J.A., Tickle, I.J., Wood, S.P. and Blundell, T.L. (1975) The relation of polypeptide hormone structure and flexibility to receptor binding: the relevance of X-ray studies on insulins, glucagon and human placental lactogen. *Molecular and Cellular Biochemistry*, **8**, 5-20.

17. Cary, P.D., Moss, T. and Bradbury, E.M. (1978) High-resolution proton-magnetic-resonance studies of chromatin core particles. *European Journal of Biochemistry*, **89**, 475-482.
18. Neurath, H. (1959) Protein structure and enzyme action. *Reviews of Modern Physics*, **31**, 185-190.
19. Chen, J., Liang, H. and Fernandez, A. (2008) Protein structure protection commits gene expression patterns. *Genome Biology*, **9**, R107.
20. Uversky, V.N. (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *Journal of Biomolecular Structure & Dynamics*, **21**, 211-234.
21. Fuxreiter, M., Tompa, P., Simon, I., Uversky, V.N., Hansen, J.C. and Asturias, F.J. (2008) Malleable machines take shape in eukaryotic transcriptional regulation. *Nature Chemical Biology*, **4**, 728-737.
22. Livesay, D.R. (2010) Protein dynamics: dancing on an ever-changing free energy stage. *Current Opinion in Pharmacology*, **10**, 706-708.
23. Uversky, V.N. (2011) Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. *Chemical Society Reviews*, **40**, 1623-1634.
24. Uversky, V.N., Gillespie, J.R. and Fink, A.L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins-Structure Function and Genetics*, **41**, 415-427.
25. Radivojac, P., Iakoucheva, L.M., Oldfield, C.J., Obradovic, Z., Uversky, V.N. and Dunker, A.K. (2007) Intrinsic disorder and functional proteomics. *Biophysical Journal*, **92**, 1439-1456.
26. Romero, P., Obradovic, Z., Li, X.H., Garner, E.C., Brown, C.J. and Dunker, A.K. (2001) Sequence complexity of disordered protein. *Proteins-Structure Function and Genetics*, **42**, 38-48.
27. Muller-Spath, S., Soranno, A., Hirschfeld, V., Hofmann, H., Ruegger, S., Reymond, L., Nettels, D. and Schuler, B. (2010) From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proceedings of National Academy of Sciences of the United States of America*, **107**, 14609-14614.
28. Ferron, F., Longhi, S., Canard, B. and Karlin, D. (2006) A practical overview of protein disorder prediction methods. *Proteins-Structure Function and Bioinformatics*, **65**, 1-14.
29. Oldfield, C.J., Cheng, Y., Cortese, M.S., Brown, C.J., Uversky, V.N. and Dunker, A.K. (2005) Comparing and combining predictors of mostly disordered proteins. *Biochemistry*, **44**, 1989-2000.
30. Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V.N. *et al.* (2007) DisProt: the database of disordered proteins. *Nucleic Acids Residues*, **35**, D786-793.
31. Gunasekaran, K., Tsai, C.J., Kumar, S., Zanuy, D. and Nussinov, R. (2003) Extended disordered proteins: targeting function with less scaffold. *Trends in Biochemecial Sciences*, **28**, 81-85.
32. Gunasekaran, K., Tsai, C.J. and Nussinov, R. (2004) Analysis of ordered and disordered protein complexes reveals structural features discriminating

- between stable and unstable monomers. *Journal of Molecular Biology*, **341**, 1327-1341.
33. Meszaros, B., Tompa, P., Simon, I. and Dosztanyi, Z. (2007) Molecular principles of the interactions of disordered proteins. *Journal of Molecular Biology*, **372**, 549-561.
  34. <http://www.scientificamerican.com/article.cfm?id=the-orderly-chaos-of-proteins>.
  35. Wright, P.E. and Dyson, H.J. (2009) Linking folding and binding. *Current Opinion in Structural Biology*, **19**, 31-38.
  36. Tsai, C.J., Ma, B., Sham, Y.Y., Kumar, S. and Nussinov, R. (2001) Structured disorder and conformational selection. *Proteins*, **44**, 418-427.
  37. Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, **293**, 321-331.
  38. Tompa, P. and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in Biochemical Sciences*, **33**, 2-8.
  39. Uversky, V.N., Oldfield, C.J. and Dunker, A.K. (2008) Intrinsically disordered proteins in human diseases: Introducing the D<sup>2</sup> concept. *Annual Review of Biophysics*, **37**, 215-246.
  40. Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradovic, Z. and Dunker, A.K. (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of Molecular Biology*, **323**, 573-584.
  41. Tompa, P., Szasz, C. and Buday, L. (2005) Structural disorder throws new light on moonlighting. *Trends in Biochemical Sciences*, **30**, 484-489.
  42. Eliezer, D. (2009) Biophysical characterization of intrinsically disordered proteins. *Current Opinions in Structural Biology*, **19**, 23-30.
  43. Mittag, T. and Forman-Kay, J.D. (2007) Atomic-level characterization of disordered protein ensembles. *Current Opinions in Structural Biology*, **17**, 3-14.
  44. Dyson, H.J. and Wright, P.E. (2004) Unfolded proteins and protein folding studied by NMR. *Chemistry Reviews*, **104**, 3607-3622.
  45. Receveur-Brechot, V. and Durand, D. (2012) How random are intrinsically disordered proteins? A small angle scattering perspective. *Current Protein and Peptide Science*, **13**, 55-75.
  46. Keiderling, T.A. and Xu, Q. (2002) Unfolded peptides and proteins studied with infrared absorption and vibrational circular dichroism spectra. *Advances in Protein Chemistry*, **62**, 111-161.
  47. Barron, L.D., Blanch, E.W. and Hecht, L. (2002) Unfolded proteins studied by Raman optical activity. *Advances in Protein Chemistry*, **62**, 51-90.
  48. Sakon, J.J. and Weninger, K.R. (2010) Detecting the conformation of individual proteins in live cells. *Nature Methods*, **7**, 203-205.
  49. Beveridge, R., Chappuis, Q., Macphree, C. and Barran, P. (2013) Mass spectrometry methods for intrinsically disordered proteins. *Analyst*, **138**, 32-42.

50. Konermann, L., Tong, X. and Pan, Y. (2008) Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. *Journal of Mass Spectrometry*, **43**, 1021-1036.
51. Hubbard, S.J., Eisenmenger, F. and Thornton, J.M. (1994) Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Science*, **3**, 757-768.
52. Cohen, S.L., Ferre-D'Amare, A.R., Burley, S.K. and Chait, B.T. (1995) Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. *Protein Science*, **4**, 1088-1099.
53. Massotte, D., Yamamoto, M., Scianimanico, S., Sorokine, O., van Dorsselaer, A., Nakatani, Y., Ourisson, G. and Pattus, F. (1993) Structure of the membrane-bound form of the pore-forming domain of colicin A: a partial proteolysis and mass spectrometry study. *Biochemistry*, **32**, 13787-13794.
54. Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature*, **422**, 198-207.
55. Bordoli, L., Kiefer, F. and Schwede, T. (2007) Assessment of disorder predictions in CASP7. *Proteins*, **69 Suppl 8**, 129-136.
56. Dosztanyi, Z., Csizmok, V., Tompa, P. and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*, **21**, 3433-3434.
57. Csizmók, V. and Tompa, P. (2009) In Ovádi, J. and Orosz, F. (eds.), *Protein Folding and Misfolding: Neurodegenerative Diseases*. Springer Netherlands, Vol. 7, pp. 1-19.
58. Bell, S., Klein, C., Muller, L., Hansen, S. and Buchner, J. (2002) p53 contains large unstructured regions in its native state. *Journal of Molecular Biology*, **322**, 917-927.
59. Hollstein, M., Sidransky, D., Vogelstein, B. and Harris, C.C. (1991) p53 mutations in human cancers. *Science*, **253**, 49-53.
60. Reich, N.C. and Levine, A.J. (1982) Specific interaction of the SV40 T antigen-cellular p53 protein complex with SV40 DNA. *Virology*, **117**, 286-290.
61. Levine, A.J., Hu, W. and Feng, Z. (2006) The P53 pathway: what questions remain to be explored? *Cell Death and Differentiation*, **13**, 1027-1036.
62. Vogelstein, B., Lane, D. and Levine, A.J. (2000) Surfing the p53 network. *Nature*, **408**, 307-310.
63. Vousden, K.H. and Lu, X. (2002) Live or let die: the cell's response to p53. *Nature Reviews Cancer*, **2**, 594-604.
64. Anderson, C.W. and Appella, E. (2004) Signaling to the p53 tumor suppressor through pathways activated by genotoxic and nongenotoxic stress. *Bradshaw RA; Dennis EA ed, Academic Press*.
65. Huart, A.S. and Hupp, T.R. (2013) Evolution of conformational disorder & diversity of the p53 interactome *BioDiscovery*, **8**, 5.
66. Matsuoka, S., Huang, M. and Elledge, S.J. (1998) Linkage of ATM to cell cycle regulation by the chk2 protein kinase. *Science*, **282**, 1893-1897.
67. Banin, S., Moyal, L., Shieh, S.-Y., Taya, Y., Anderson, C.W., Chessa, L., Smorodinsky, N.I., Prives, C., Reiss, Y., Shiloh, Y. *et al.* (1998) Enhanced

- phosphorylation of p53 by ATM in response to DNA damage. *Science*, **281**, 1674-1677.
68. Bai, L. and Zhu, W.G. (2006) p53: structure, function and therapeutic applications. *Journal of Cancer Molecules*, **2**, 141-153.
  69. Olivier, M., Eeles, R., Hollstein, M., Khan, M.A., Harris, C.C. and Hainaut, P. (2002) The IARC TP53 database: new online mutation analysis and recommendations to users. *Human Mutations*, **19**, 607-614.
  70. Oldfield, C.J., Cheng, Y., Cortese, M.S., Romero, P., Uversky, V.N. and Dunker, A.K. (2007) Intrinsic disorder in protein-protein interaction networks: case studies of complexes involving p53 and 14-3-3. *BMC Genomics*, **BIOCOMP**, 553-566.
  71. Joerger, A.C. and Fersht, A.R. (2008) Structural biology of the tumor suppressor p53. *Annual Review of Biochemistry*, **77**, 557-582.
  72. Cho, Y., Gorina, S., Jeffrey, P. and Pavletich, N. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346-355.
  73. Clore, G.M., Omichinski, J.G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E. and Gronenborn, A.M. (1994) High-resolution structure of the oligomerization domain of p53 by multidimensional NMR. *Science*, **265**, 386-391.
  74. Clore, G.M., Ernst, J., Clubb, R., Omichinski, J.G., Kennedy, W.M.P., Sakaguchi, K., Appella, E. and Gronenborn, A.M. (1995) Refined solution structure of the oligomerization domain of the tumour suppressor p53. *Nature Structural & Molecular Biology*, **2**, 321-333.
  75. Lee, W., Harvey, T.S., Yin, Y., Yau, P., Litchfield, D. and Arrowsmith, C.H. (1994) Solution structure of the tetrameric minimum transforming domain of p53. *Nature Structural Biology*, **1**, 877-890.
  76. Jeffrey, P.D., Gorina, S. and Pavletich, N.P. (1995) Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. *Science*, **267**, 1498-1502.
  77. Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A.J. and Pavletich, N.P. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science*, **274**, 948-953.
  78. Rustandi, R.R., Baldisseri, D.M. and Weber, D.J. (2000) Structure of the negative regulatory domain of p53 bound to S100B(beta-beta). *Nature Structural Biology*, **7**, 570-574.
  79. Avalos, J.L., Bever, K.M. and Wolberger, C. (2005) Mechanism of sirtuin inhibition by nicotinamide: Altering the NAD<sup>+</sup> cosubstrate specificity of a Sir2 enzyme. *Molecular Cell*, **17**, 855-868.
  80. Lowe, E.D., Tews, I., Cheng, K.Y., Brown, N.R., Gul, S., Noble, M.E., Gamblin, S.J. and Johnson, L.N. (2002) Specificity determinants of recruitment peptides bound to phospho-CDK2/cyclin A. *Biochemistry*, **41**, 15625-15634.
  81. Tidow, H., Melero, R., Mylonas, E., Freund, S.M.V., Grossmann, J.G., Carazo, J.M., Svergun, D.I., Valle, M. and Fersht, A.R. (2007) Quaternary structures of tumor suppressor p53 and a specific p53-DNA complex.

- Proceedings of the National Academy of Sciences of the United States of America*, **104**, 12324-12329.
82. Wells, M., Tidow, H., Rutherford, T.J., Markwick, P., Jensen, M.R., Mylonas, E., Svergun, D.I., Blackledge, M. and Fersht, A.R. (2008) Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proceedings of National Academy of Sciences of the United States of America*, **105**, 5762-5767.
  83. Joerger, A.C., Allen, M.D. and Fersht, A.R. (2004) Crystal structure of a superstable mutant of human p53 core domain - Insights into the mechanism of rescuing oncogenic mutations. *Journal of Biological Chemistry*, **279**, 1291-1296.
  84. Nikolova, P.V., Henckel, J., Lane, D.P. and Fersht, A.R. (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proceedings of the National Academy of Sciences*, **95**, 14675-14680.
  85. Okorokov, A.L., Sherman, M.B., Plisson, C., Grinkevich, V., Sigmundsson, K., Selivanova, G., Milner, J. and Orlova, E.V. (2006) The structure of p53 tumour suppressor protein reveals the basis for its functional plasticity. *Embo Journal*, **25**, 5191-5200.
  86. Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E. and Shakked, Z. (2006) Structural basis of DNA recognition by p53 tetramers. *Molecular Cell*, **22**, 741-753.
  87. Toledo, F. and Wahl, G.M. (2006) Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. *Nature Reviews Cancer*, **6**, 909-923.
  88. Lavin, M.F. and Gueven, N. (2006) The complexity of p53 stabilization and activation. *Cell Death & Differentiation*, **13**, 941-950.
  89. Bode, A.M. and Dong, Z. (2004) Post-translational modification of p53 in tumorigenesis. *Nature Reviewa Cancer*, **4**, 793-805.
  90. Xu, Y. (0000) Regulation of p53 responses by post-translational modifications. *Cell Death & Differentiation*, **10**, 400-403.
  91. Siliciano, J.D., Canman, C.E., Taya, Y., Sakaguchi, K., Appella, E. and Kastan, M.B. (1997 ) DNA damage induces phosphorylation of the amino terminus of p53. *Genes & Development*, **15**, 3471-3481.
  92. Shieh, S.Y., Ahn, J., Tamai, K., Taya, Y. and Prives, C. (2000) The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. *Genes & Development*, **14**, 289-300.
  93. Tibbetts, R.S., Brumbaugh, K.M., Williams, J.M., Sarkaria, J.N., Cliby, W.A., Shieh, S.Y., Y., T., Prives, C. and Abraham, R.T. (1999) A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes & Development*, **15**, 152-157.
  94. Canman, C.E., Lim, D.S., Cimprich, K.A., Taya, Y., Tamai, K., Sakaguchi, K., Appella, E., Kastan, M.B. and Siliciano, J.D. (1998) Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science*, **281**, 1677-1679.
  95. Gu, W. and Roeder, R.G. (1997) Activation of p53 Sequence-Specific DNA Binding by Acetylation of the p53 C-Terminal Domain. *Cell*, **90**, 595-606.



96. Friedler, A., Veprintsev, D.B., Freund, S.M.V., von Glos, K.I. and Fersht, A.R. (2005) Modulation of Binding of DNA to the C-Terminal Domain of p53 by Acetylation. *Structure*, **13**, 629-636.
97. Wang, Y.H., Tsay, Y.G., Tan, B.C., Lo, W.Y. and Lee, S.C. (2003) Identification and characterization of a novel p300-mediated p53 acetylation site, lysine 305. *Journal of Biological Chemistry*, **278**, 25568-25576.
98. Alarcon-Vargas, D. and Ronai, Z.e. (2002) p53-Mdm2 - the affair that never ends. *Carcinogenesis*, **23**, 541-547.
99. Wu, X., Bayle, J.H., Olson, D. and Levine, A.J. (1993) The p53-mdm-2 autoregulatory feedback loop. *Genes & Development*, **7**, 1126-1132.
100. Freedman, D.A. and Levine, A.J. (1998) Nuclear export Is required for degradation of endogenous p53 by Mdm2 and human papillomavirus E6. *Molecular and Cellular Biology*, **18**, 7288-7293.
101. Tao, W. and Levine, A.J. (1999) Nucleocytoplasmic shuttling of oncoprotein Hdm2 is required for Hdm2-mediated degradation of p53. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 3077-3080.
102. Pickart, C.M. (2001) Mechanisms underlying ubiquitination. *Annual Review of Biochemistry*, **70**, 503-533.
103. Fang, S., Jensen, J.P., Ludwig, R.L., Vousden, K.H. and Weissman, A.M. (2000) Mdm2 is a RING finger-dependent ubiquitin protein ligase for Itself and p53. *Journal of Biological Chemistry*, **275**, 8945-8951.
104. Shieh, S.Y., Ikeda, M., Taya, Y. and Prives, C. (1997) DNA damage-induced phosphorylation of p53 alleviates inhibition by MDM2. *Cell*, **91**, 325-334.
105. Fuchs, S.Y., Adler, V., Buschmann, T., Wu, X.W. and Ronai, Z. (1998) Mdm2 association with p53 targets its ubiquitination. *Oncogene*, **17**, 2543-2547.
106. Momand, J., Jung, D., Wilczynski, S. and Niland, J. (1998) The MDM2 gene amplification database. *Nucleic Acids Residues*, **26**, 3453-3459.
107. Chang, Y.S., Graves, B., Guerlavais, V., Tovar, C., Packman, K., To, K.-H., Olson, K.A., Kesavan, K., Gangurde, P., Mukherjee, A. *et al.* (2013) Stapled  $\alpha$ -helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proceedings of the National Academy of Sciences*.
108. Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C. *et al.* (2004) In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science*, **303**, 844-848.
109. Rew, Y., Sun, D., Gonzalez-Lopez De Turiso, F., Bartberger, M.D., Beck, H.P., Canon, J., Chen, A., Chow, D., Deignan, J., Fox, B.M. *et al.* (2012) Structure-Based Design of Novel Inhibitors of the MDM2-p53 Interaction. *Journal of Medicinal Chemistry*, **55**, 4936-4954.
110. Grasberger, B.L., Lu, T., Schubert, C., Parks, D.J., Carver, T.E., Koblisch, H.K., Cummings, M.D., LaFrance, L.V., Milkiewicz, K.L., Calvo, R.R. *et al.* (2005) Discovery and Cocrystal Structure of Benzodiazepinedione HDM2 Antagonists That Activate p53 in Cells. *Journal of Medicinal Chemistry*, **48**, 909-912.

111. Allen, J.G., Bourbeau, M.P., Wohlieter, G.E., Bartberger, M.D., Michelsen, K., Hungate, R., Gadwood, R.C., Gaston, R.D., Evans, B., Mann, L.W. *et al.* (2009) Discovery and Optimization of Chromenotriazolopyrimidines as Potent Inhibitors of the Mouse Double Minute 2–Tumor Protein 53 Protein–Protein Interaction. *Journal of Medicinal Chemistry*, **52**, 7044-7053.
112. Shloush, J., Vlassov, J.E., Engson, I., Duan, S., Saridakis, V., Dhe-Paganon, S., Raught, B., Sheng, Y. and Arrowsmith, C.H. (2011) Structural and functional comparison of the RING domains of two p53 E3 ligases, Mdm2 and Pirh2. *Journal of Biological Chemistry*, **286**, 4796-4808.
113. Benesch, J.L.P., Ruotolo, B.T., Simmons, D.A. and Robinson, C.V. (2007) Protein complexes in the gas phase: Technology for structural genomics and proteomics. *Chemical Reviews*, **107**, 3544-3567.
114. Jarrold, M.F. (2000) Peptides and proteins in the vapor phase. *Annual Review of Physical Chemistry*, **51**, 179-207.
115. van den Heuvel, R.H. and Heck, A.J. (2004) Native protein mass spectrometry: from intact oligomers to functional machineries. *Current Opinion in Chemical Biology*, **8**, 519-526.
116. Loo, J.A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews*, **16**, 1-23.
117. Ruotolo, B.T. and Robinson, C.V. (2006) Aspects of native proteins are retained in vacuum. *Current Opinion in Chemical Biology*, **10**, 402-408.
118. Konermann, L. and Douglas, D.J. (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**, 1248-1254.
119. Uversky, V.N. and Dunker, A.K. (2011) Multiparametric analysis of intrinsically disordered proteins: Looking at intrinsic disorder through compound eyes. *Analytical Chemistry*, **84**, 2096-2104.
120. Frimpong, A.K., Abzalimov, R.R., Uversky, V.N. and Kaltashov, I.A. (2010) Characterization of intrinsically disordered proteins with electrospray ionization mass spectrometry: conformational heterogeneity of alpha-synuclein. *Proteins*, **78**, 714-722.
121. Bennett, M.C. (2005) The role of alpha-synuclein in neurodegenerative diseases. *Pharmacology & Therapeutics*, **105**, 311-331.
122. Griffith, W.P. and Kaltashov, I.A. (2003) Highly asymmetric interactions between globin chains during hemoglobin assembly revealed by electrospray ionization mass spectrometry. *Biochemistry*, **42**, 10024-10033.
123. Bernstein, S., Liu, D., Wytenbach, T., Bowers, M., Lee, J., Gray, H. and Winkler, J. (2004)  $\alpha$ -Synuclein: Stable compact and extended monomeric structures and pH dependence of dimer formation. *Journal of the American Society for Mass Spectrometry*, **15**, 1435-1443.
124. Faull, P.A., Florance, H.V., Schmidt, C.Q., Tomczyk, N., Barlow, P.N., Hupp, T.R., Nikolova, P.V. and Barran, P.E. (2010) Utilising ion mobility-mass spectrometry to interrogate macromolecules: Factor H complement control protein modules 10–15 and 19–20 and the DNA-binding core domain of tumour suppressor p53. *International Journal of Mass Spectrometry*, **298**, 99-110.

125. Maurizio, E., Cravello, L., Brady, L., Spolaore, B., Arnoldo, L., Giancotti, V., Manfioletti, G. and Sgarra, R. (2011) Conformational role for the C-terminal tail of the intrinsically disordered high mobility group A (HMGA) chromatin factors. *Journal of Proteome Research*, **10**, 3283-3291.
126. Canon, F., Ballivian, R., Chirot, F., Antoine, R., Sarni-Manchado, P., Lemoine, J. and Dugourd, P. (2011) Folding of a salivary intrinsically disordered protein upon binding to tannins. *Journal of the American Chemical Society*, **133**, 7847-7852.
127. Erikson, H.A. (1927) The Mobility of Argon and Hydrogen Ions in Air. *Physical Review*, **30**, 339-348.
128. Bradbury, N.E. (1932) The absolute values of the mobility of gaseous ions in pure gases. *Physical Review*, **40**, 508-523.
129. Langevin, P. (1903), *Annales de Chimie et de Physique*, Vol. 28, pp. 433-530.
130. McDaniel, E.W. (1962) Drift tube-mass spectrometer for studies of low-energy ion-molecule reactions *Review of Scientific Instruments*, **33**, 2.
131. Barnes, W.S., Martin, D.W. and McDaniel, E.W. (1961) Mass Spectrographic Identification of the Ion Observed in Hydrogen Mobility Experiments. *Physical Review Letters*, **6**, 110-111.
132. Hogg, A.M. and Kebarle, P. (1965) Mass-spectrometric study of ions at near-atmospheric pressure. II. Ammonium ions produced by the  $\alpha$ -radiolysis of ammonia and their solvation in the gas phase by ammonia and water molecules. *Journal of Chemical Physics* **43**, 449-456.
133. Kebarle, P. and Hogg, A.M. (1965) Mass-spectrometric study of ions at near atmospheric pressures. I. The ionic polymerization of ethylene. *Journal of Chemical Physics* **42**, 668-674.
134. Kanu, A.B., Dwivedi, P., Tam, M., Matz, L. and Hill, H.H., Jr. (2008) Ion mobility-mass spectrometry. *Journal of Mass Spectrometry*, **43**, 1-22.
135. Uetrecht, C., Rose, R.J., van Duijn, E., Lorenzen, K. and Heck, A.J.R. (2010) Ion mobility mass spectrometry of proteins and protein assemblies. *Chemical Society Reviews*, **39**, 1633-1655.
136. Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinou, K., Bateman, R.H., Bowers, M.T. and 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**, 1-12.
137. Mason, E.A. and McDaniel, E.W. (1988) *Transport properties of ions in gases*. Wiley, New York.
138. Shelimov, K.B., Clemmer, D.E., Hudgins, R.R. and Jarrold, M.F. (1997) Protein structure in vacuo: gas-phase conformations of BPTI and cytochrome c. *Journal of the American Chemical Society*, **119**, 2240-2248.
139. Badman, E.R., Hoaglund-Hyzer, C.S. and Clemmer, D.E. (2001) Monitoring structural changes of proteins in an ion trap over ~10–200 ms: unfolding transitions in cytochrome c ions. *Analytical Chemistry*, **73**, 6000-6007.
140. Valentine, S. and Clemmer, D. (2002) Temperature-dependent H/D exchange of compact and elongated cytochrome c ions in the gas phase. *Journal of the American Society for Mass Spectrometry*, **13**, 506-517.

141. Gidden, J., Ferzoco, A., Baker, E.S. and Bowers, M.T. (2004) Duplex formation and the onset of helicity in poly d(CG)<sub>n</sub> oligonucleotides in a solvent-free environment. *Journal of the American Chemical Society*, **126**, 15132-15140.
142. Clemmer, D.E., Hudgins, R.R. and Jarrold, M.F. (1995) Naked protein conformations: Cytochrome c in the gas phase. *Journal of the American Chemical Society*, **117**, 10141-10142.
143. <http://www.umich.edu/~ruotolo/instruments.html>.
144. Matz, L.M., Hill, H.H., Jr., Beegle, L.W. and Kanik, I. (2002) Investigation of drift gas selectivity in high resolution ion mobility spectrometry with mass spectrometry detection. *Journal of the American Society for Mass Spectrometry*, **13**, 300-307.
145. Kemper, P. and Bowers, M. (1990) A hybrid double-focusing mass spectrometer-High-pressure drift reaction cell to study thermal energy reactions of mass-selected ions. *Journal of the American Society for Mass Spectrometry*, **1**, 197-207.
146. Kemper, P.R., Dupuis, N.F. and Bowers, M.T. (2009) A new, higher resolution, ion mobility mass spectrometer. *International Journal of Mass Spectrometry*, **287**, 46-57.
147. Dugourd, P., Hudgins, R.R., Clemmer, D.E. and Jarrold, M.F. (1997) High-resolution ion mobility measurements. *Review of Scientific Instruments*, **68**, 1122-1129.
148. Jarrold, M.F. (1995) Drift tube studies of atomic clusters. *The Journal of Physical Chemistry*, **99**, 11-21.
149. Merenbloom, S.I., Koeniger, S.L., Valentine, S.J., Plasencia, M.D. and Clemmer, D.E. (2006) IMS-IMS and IMS-IMS-IMS/MS for separating peptide and protein fragment ions. *Analytical Chemistry*, **78**, 2802-2809.
150. Valentine, S.J., Koeniger, S.L. and Clemmer, D.E. (2003) A split-field drift tube for separation and efficient fragmentation of biomolecular ions. *Analytical Chemistry*, **75**, 6202-6208.
151. Myung, S., Lee, Y.J., Moon, M.H., Taraszka, J., Sowell, R., Koeniger, S., Hilderbrand, A.E., Valentine, S.J., Cherbas, L., Cherbas, P. *et al.* (2003) Development of high-sensitivity ion trap ion mobility spectrometry time-of-flight techniques: A high-throughput nano-LC-IMS-TOF separation of peptides rising from a drosophila protein extract. *Analytical Chemistry*, **75**, 5137-5145.
152. Hoaglund-Hyzer, C.S. and Clemmer, D.E. (2000) Ion trap/ion mobility/quadrupole/time-of-flight mass spectrometry for peptide mixture analysis. *Analytical Chemistry*, **73**, 177-184.
153. Hoaglund, C.S., Valentine, S.J., Sporleder, C.R., Reilly, J.P. and Clemmer, D.E. (1998) Three-dimensional ion mobility/TOFMS analysis of electrosprayed biomolecules. *Analytical Chemistry*, **70**, 2236-2242.
154. Hoaglund, C.S., Valentine, S.J. and Clemmer, D.E. (1997) An ion trap interface for ESI-ion mobility experiments. *Analytical Chemistry*, **69**, 4156-4161.
155. Tang, X., Bruce, J.E. and Hill, H.H., Jr. (2007) Design and performance of an atmospheric pressure ion mobility Fourier transform ion cyclotron resonance

- mass spectrometer. *Rapid Communication in Mass Spectrometry*, **21**, 1115-1122.
156. Wittmer, D., Chen, Y.H., Luckenbill, B.K. and Hill, H.H. (1994) Electrospray ionization ion mobility spectrometry. *Analytical Chemistry*, **66**, 2348-2355.
  157. Wytenbach, T., Kemper, P.R. and Bowers, M.T. (2001) Design of a new electrospray ion mobility mass spectrometer. *International Journal of Mass Spectrometry*, **212**, 13-23.
  158. Bernstein, S.L., Wytenbach, T., Baumketner, A., Shea, J.-E., Bitan, G., Teplow, D.B. and Bowers, M.T. (2005) Amyloid  $\beta$ -protein: Monomer structure and early aggregation states of A $\beta$ 42 and Its Pro19 alloform. *Journal of the American Chemical Society*, **127**, 2075-2084.
  159. Baker, E.S., Hong, J.W., Gaylord, B.S., Bazan, G.C. and Bowers, M.T. (2006) PNA/dsDNA complexes: Site specific binding and dsDNA biosensor applications. *Journal of the American Chemical Society*, **128**, 8484-8492.
  160. Barran, P.E., Roeske, R.W., Pawson, A.J., Sellar, R., Bowers, M.T., Morgan, K., Lu, Z.-L., Tsuda, M., Kusakabe, T. and Millar, R.P. (2005) Evolution of constrained gonadotropin-releasing hormone ligand conformation and receptor selectivity. *Journal of Biological Chemistry*, **280**, 38569-38575.
  161. Taraszka, J.A., Counterman, A.E. and Clemmer, D.E. (2001) Gas-phase separations of complex tryptic peptide mixtures. *Fresenius Journal of Analytical Chemistry*, **369**, 234-245.
  162. Valentine, S.J., Kulchania, M., Barnes, C.A.S. and Clemmer, D.E. (2001) Multidimensional separations of complex peptide mixtures: a combined high-performance liquid chromatography/ion mobility/time-of-flight mass spectrometry approach. *International Journal of Mass Spectrometry*, **212**, 97-109.
  163. Hilderbrand, A.E., Myung, S., Srebalus Barnes, C.A. and Clemmer, D.E. (2003) Development of LC-IMS-CID-TOFMS techniques: Analysis of a 256 component tetrapeptide combinatorial library. *Journal of the American Society for Mass Spectrometry*, **14**, 1424-1436.
  164. Moon, M.H., Myung, S., Plasencia, M., Hilderbrand, A.E. and Clemmer, D.E. (2003) Nanoflow LC/ion mobility/CID/TOF for proteomics: Analysis of a human urinary proteome. *Journal of Proteome Research*, **2**, 589-597.
  165. Xun, Z., Sowell, R.A., Kaufman, T.C. and Clemmer, D.E. (2006) Protein expression in a drosophila model of parkinson's disease. *Journal of Proteome Research*, **6**, 348-357.
  166. McCullough, B.J., Kalapothakis, J., Eastwood, H., Kemper, P., MacMillan, D., Taylor, K., Dorin, J. and Barran, P.E. (2008) Development of an ion mobility quadrupole time of flight mass spectrometer. *Analytical Chemistry*, **80**, 6336-6344.
  167. McLean, J.A., Ridenour, W.B. and Caprioli, R.M. (2007) Profiling and imaging of tissues by imaging ion mobility-mass spectrometry. *Journal of Mass Spectrometry*, **42**, 1099-1105.
  168. Gerlich, D. (2007), *Advances in Chemical Physics*. John Wiley & Sons, Inc., pp. 1-176.
  169. Shvartsburg, A.A. and Smith, R.D. (2008) Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry*, **80**, 9689-9699.

170. Ruotolo, B.T., Benesch, J.L., Sandercock, A.M., Hyung, S.J. and Robinson, C.V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nature Protocols*, **3**, 1139-1152.
171. Ruotolo, B.T., Giles, K., Campuzano, I., Sandercock, A.M., Bateman, R.H. and Robinson, C.V. (2005) Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science*, **310**, 1658-1661.
172. Ruotolo, B.T., Hyung, S.-J., Robinson, P.M., Giles, K., Bateman, R.H. and Robinson, C.V. (2007) Ion Mobility–Mass Spectrometry Reveals Long-Lived, Unfolded Intermediates in the Dissociation of Protein Complexes. *Angewandte Chemie International Edition*, **46**, 8001-8004.
173. Smith, D.P., Radford, S.E. and Ashcroft, A.E. (2010) Elongated oligomers in beta2-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry. *Proceedings of National Academy of Sciences of the United States of America*, **107**, 6794-6798.
174. Scarff, C.A., Patel, V.J., Thalassinou, K. and Scrivens, J.H. (2009) Probing hemoglobin structure by means of traveling-wave ion mobility mass spectrometry. *Journal of the American Society for Mass Spectrometry*, **20**, 625-631.
175. Harvey, D.J., Sobott, F., Crispin, M., Wrobel, A., Bonomelli, C., Vasiljevic, S., Scanlan, C.N., Scarff, C.A., Thalassinou, K. and Scrivens, J.H. (2011) Ion mobility mass spectrometry for extracting spectra of N-glycans directly from incubation mixtures following glycan release: application to glycans from engineered glycoforms of intact, folded HIV gp120. *Journal of the American Society for Mass Spectrometry*, **22**, 568-581.
176. van Duijn, E., Barendregt, A., Synowsky, S., Versluis, C. and Heck, A.J. (2009) Chaperonin complexes monitored by ion mobility mass spectrometry. *Journal of the American Chemical Society*, **131**, 1452-1459.
177. Uetrecht, C., Versluis, C., Watts, N.R., Wingfield, P.T., Steven, A.C. and Heck, A.J. (2008) Stability and shape of hepatitis B virus capsids in vacuo. *Angewandte Chemie International Edition*, **47**, 6247-6251.
178. Hopper, J.T. and 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. *Journal of the American Society for Mass Spectrometry*, **20**, 1851-1858.
179. Buryakov, I.A., Krylov, E.V., Nazarov, E.G. and Rasulev, U.K. (1993) A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field. *International Journal of Mass Spectrometry and Ion Processes*, **128**, 143-148.
180. Purves, R.W., Guevremont, R., Day, S., Pipich, C.W. and Matyjaszczyk, M.S. (1998) Mass spectrometric characterization of a high-field asymmetric waveform ion mobility spectrometer. *Review of Scientific Instruments*, **69**, 4094-4105.
181. Barnett, D.A., Ells, B., Guevremont, R. and Purves, R.W. (1999) Separation of leucine and isoleucine by electrospray ionization–high field asymmetric waveform ion mobility spectrometry–mass spectrometry. *Journal of the American Society for Mass Spectrometry*, **10**, 1279-1284.

182. Borysik, A.J., Read, P., Little, D.R., Bateman, R.H., Radford, S.E. and Ashcroft, A.E. (2004) Separation of beta2-microglobulin conformers by high-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled to electrospray ionisation mass spectrometry. *Rapid Communication in Mass Spectrometry*, **18**, 2229-2234.
183. Tang, K., Li, F., Shvartsburg, A.A., Strittmatter, E.F. and Smith, R.D. (2005) Two-dimensional gas-phase separations coupled to mass spectrometry for analysis of complex mixtures. *Analytical Chemistry*, **77**, 6381-6388.
184. Shvartsburg, A.A., Tang, K. and Smith, R.D. (2005) Optimization of the design and operation of FAIMS analyzers. *Journal of the American Society for Mass Spectrometry*, **16**, 2-12.
185. Shvartsburg, A.A., Noskov, S.Y., Purves, R.W. and Smith, R.D. (2009) Pendular proteins in gases and new avenues for characterization of macromolecules by ion mobility spectrometry. *Proceedings of National Academy of Sciences of the United States of America*, **106**, 6495-6500.
186. Xuan, Y., Creese, A.J., Horner, J.A. and Cooper, H.J. (2009) High-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled with high-resolution electron transfer dissociation mass spectrometry for the analysis of isobaric phosphopeptides. *Rapid Communication in Mass Spectrometry*, **23**, 1963-1969.

# 2

## Experimental Methods

*Instrumentation used in the experiments is described with schematic diagrams and typical operating and tuning parameters. The drift tube and travelling wave based ion mobility mass spectrometers are set out. Illustration of a typical workflow for ion mobility mass spectrometry data acquisition and analysis is given. Information on nano-electrospray ionisation conditions, reagents used and mass calibration procedure is provided. Theoretical approaches for calculating rotationally averaged collision cross sections are also explained.*



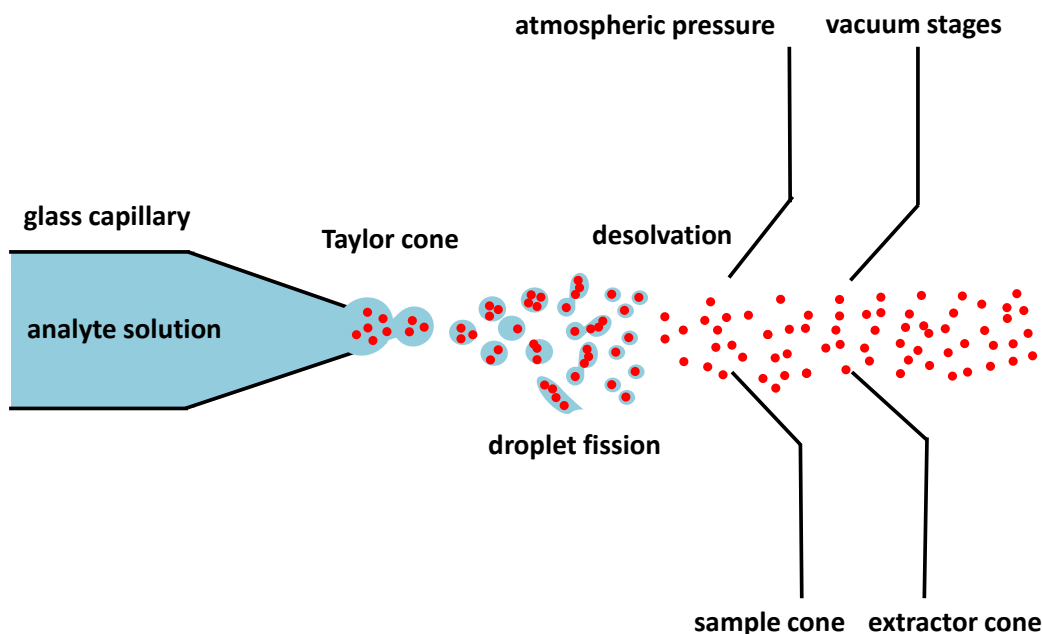
## 2. Experimental Methods

### 2.1. Reagents

For all experiments high purity water was used, obtained from an Arium 611 water purification system (Sartorius, Göttingen, Germany) or supplied by Fisher Scientific Ltd (Loughborough, UK). All organic solvents were either LC-MS or HPLC grade and supplied by Fisher Scientific Ltd (Loughborough, UK) or by Sigma Aldrich (Dorset, UK). Acids were provided by VMW International Ltd (UK). Ammonium acetate (77.08 g/mol,  $\geq 99\%$  purity) was purchased from Fisher Scientific Ltd (Loughborough, UK). Information on protein samples can be found in the relevant chapters.

### 2.2. Sample Introduction - Ionisation

Ionisation is a critical step in the transfer of intact biological molecules from solution into the vacuum environment of the mass spectrometer. Electrospray ionisation (ESI) was pioneered by Dole in 1968 (1) and advanced by a Nobel laureate Prof Fenn in 2002 (2,3). The principle of the electrospray process is shown in Figure 2.1. ESI is an atmospheric pressure ionisation technique that can produce either positive or negative ions. The sample is introduced into the ionisation source via a metal capillary to which a strong electric potential is applied (2 - 5 kV). Charging leads to droplets with an elongated meniscus, a Taylor cone, and releases droplets containing charged sample ions (4). The charged droplets are directed away from the capillary by a co-axial flow of neublizing gas (usually nitrogen). Evaporation of the solvent from droplets is achieved via the heating of the source region and another flow of nitrogen gas, termed drying gas. The size of droplets gradually decreases and due to charge conservation, Coulombic forces overcome the liquid surface tension at a critical diameter, leading to fission of smaller charged droplets. Recurring evaporation and fission events lead to charged droplets which are the gas-phase ion precursors. The desolvated analyte ions pass into the mass spectrometer through the cone, propelled by decreasing potential and pressure gradient.



**Figure 2.1:** Schematic of a n-ESI source, illustrating Taylor cone formation, production of multiply charged droplets, droplet fission, and entry of sample ions into the source of a mass spectrometer via the sample cone.

The mechanism by which a charged droplet forms a desolvated ion is believed to occur by one or two models (5). The charged residue model, proposed by Dole (1), whereby droplets undergo a sequential fission until a single analyte ion remains within a droplet. Total desolvation of this droplet leaves the gaseous ‘charged residue’. This model is widely accepted for large biological molecules such as proteins which are the topic of this thesis (6). Conversely, a competing process is the evaporation model, introduced by Iribarne and Thomson (7). This model is favoured by small ions, where analyte ion is ejected from droplets when coulombic repulsive forces increase as the droplet decreases in size.

In large molecules, such as proteins, the charged observed is interrelated to the availability of ionisable sites, which links with the key mechanism of ESI; it produces multiply charged ions from large biomolecules due to charge accumulation in the droplets. However protons might not be necessarily localised, according to the prevailing mobile proton model (8). The model assumes that ionizing protons are initially located at basic sites such as the N-terminus or side chains of lysine, arginine

and histidine. Upon ion activation the ionising proton(s) can move from less basic sites on the polypeptide chain triggering various charge-site-initiated mechanisms.

Miniaturisation of the ESI interface - nano-ESI by Wilm and Mann (9) has vastly expanded the area of application for this technique. n-ESI uses much lower flow rates, analyte volumes and concentrations. n-ESI employs glass capillaries with a micron-wide internal diameter, instead of the metal capillaries (typical internal diameter of 0.1 mm) utilised in ESI. Electric charge is provided to n-ESI capillaries via a conductive coating on the inside of the glass capillary, or to a metal wire inserted into the analyte solution. Droplets formed by n-ESI are an order of magnitude smaller than those produced by conventional ESI, following from that, the final droplets will contain less concentrated amounts of unwanted impurities or salts, resulting in a cleaner mass spectrum (10). n-ESI is a gentler desolvation process which preserves the non-covalent interactions (11,12).

All MS and IM-MS instruments utilised in this research were fitted with n-ESI sources. n-ESI glass capillaries were prepared in-house from thin-walled borosilicate capillaries (Precision Instruments, Stevenage, UK) using a Flaming/Brown micropipette puller Model P-97 (Sutter Instrument Company, Novato, CA, USA) and then filled with sample solution by means of gel micro-loading tips (Eppendorf, Hamburg, Germany). Alternatively pre-pulled platinum coated glass capillaries were used (Waters nanoflow probe tips, long thin wall, M956232AD1-S). The filled capillary was placed onto a holder and a thin platinum wire was inserted in order to provide an electrical contact between sample solution and voltage supply. This was secured to an adjustable xyz-axes stage of a Z-spray source allowing the position of the capillary to be altered with respect to the sample cone. The Micromass Z-spray ion source derives its name from the approximate shape of trajectory taken by ions from sample capillary to orthogonally positioned sample and extraction cones (Figure 2.2). The off-axis geometry ensures an excellent separation of ions from neutral molecules or excess solvent so that they diffuse away toward the vacuum pump. This circumvents the problem of clogging the skimmers by neutral species thus preventing a frequent maintenance downtime for cleaning of the source

components. Spray potential was adjusted to provide optimum signal. The precise value of voltage used depended mainly on the solution composition and spray geometry but voltages were typically in the range of 1.0-2.0 kV. In case of samples prone to aggregation a nitrogen backing gas was used to maintain a stable spray. Unless stated otherwise the source block was heated 80°C to aid desolvation process.

## **2.3. Instruments**

All of the experimental work reported here was performed on either Micromass quadrupole time-of-flight mass spectrometer that has been significantly modified in-house (13) or commercially available Synapt HDMS instrument (Waters, Manchester, UK) (14).

## **2.4. Drift Tube based IM-MS - the MoQToF**

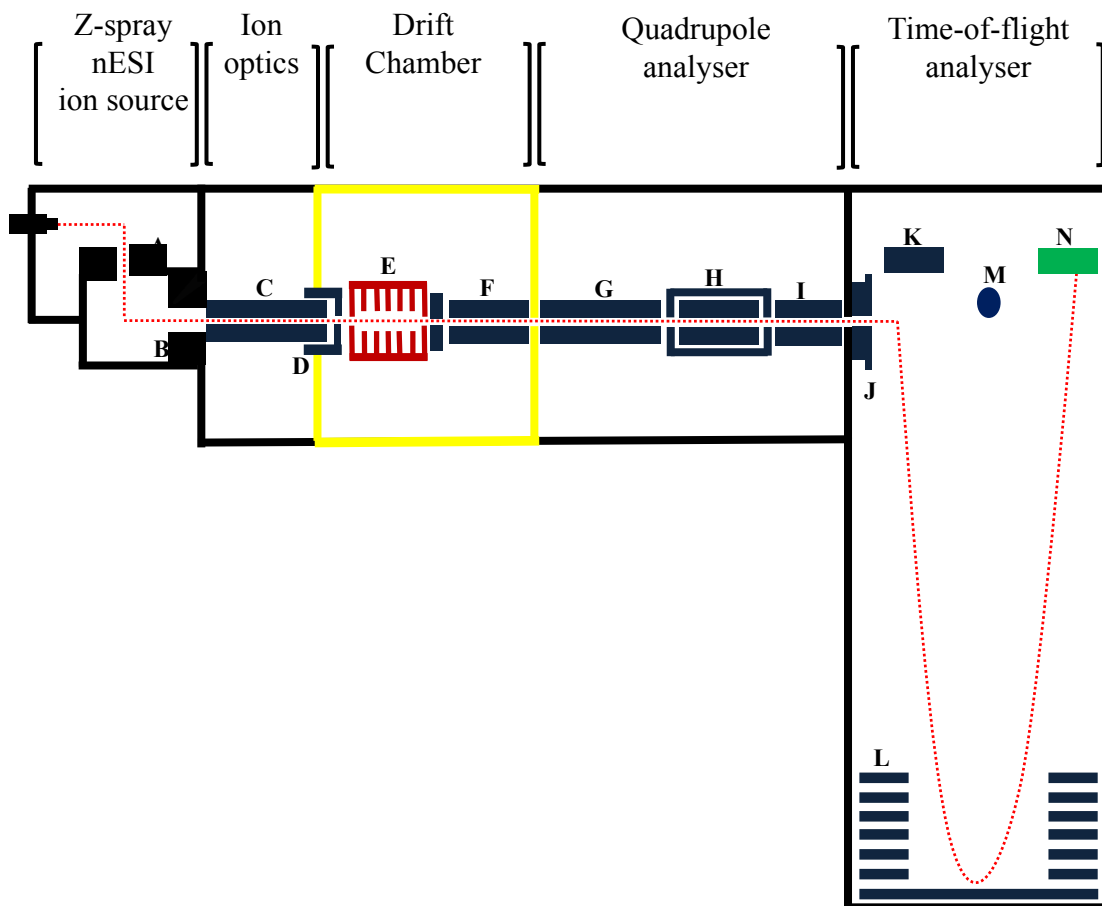
The MoQToF (Mobility Quadrupole Time-of-Flight) instrument was built in-house in 2006 in collaboration with Waters Micromass and Prof. Bowers (University of California, Santa Barbara, USA). It is a low resolution drift tube based ion mobility mass spectrometer capable of providing temperature dependant absolute collision cross section measurements.

### **2.4.1. Instrument Layout**

The MoQToF is a Q-ToF mass spectrometer that has been modified by an inclusion of a vacuum chamber containing a copper drift cell and ancillary ion optics between the first transfer hexapole and the quadrupole analyser, as shown in Figure 2.2 (13). Typically the drift cell is filed with a buffer gas helium at a pressure of 3.5 - 4.0 Torr (4.7- 5.3 mBar) and a weak electric field is applied to propel ions to drift through. Although the temperature of the drift cell can be controlled by heating or cooling the buffer gas during operation, all measurements presented here were made at 300K (300 K ~ 27°C).

Ions are produced in the n-ESI source and transferred through the source to the first hexapole region (20 cm). From here they are focused into the drift cell using a three-element Einzel lens stack. Design of the instrument allows for operation in both mass spectrometry (MS) and ion mobility mass spectrometry (IM-MS) modes. In MS mode, a continuous beam of ions drifts through the cell, whereas in the latter ion packets are released into the drift cell in discrete pulses. Ions are accumulated in the pre-cell hexapole by raising a potential on the top hat lens (TH1) located at the end of the pre-cell hexapole, immediately prior to the focusing lenses of the drift cell. This stopping voltage is lowered periodically for 40  $\mu$ s (performed via a bipolar pulsar unit) to allow the stored ion packets to enter the drift cell. The frequency of operation of this pulse is regulated by a Stanford DG535 digital delay generator (Stanford Research Systems, Sunnyvale, CA, USA) which is to be one two-hundredth of the frequency of the pusher pulse into the ToF region of the mass spectrometer. Typically, a pusher period of 115  $\mu$ s is used to cover a mass to charge ( $m/z$ ) range of 4000. Hence any two mobility events take place every  $200 \times 115 \mu\text{s} = 23 \text{ ms}$ , or at a frequency of 43.5 Hz ( $\nu = 1/t = 1/23 \text{ ms} = 43.5 \text{ Hz}$ ).

Upon exiting the drift cell, ions are focused into 12.5 cm post-cell hexapole via a single lens (L4). Subsequently ions are transmitted into the next vacuum chamber which contains a quadrupole mass analyser, collision cell and a final transfer hexapole prior to reaching the ToF analyser. In a quadrupole mode ions are detected on a point detector and on the microchannel plate detector (MCP) when operating in ToF mode. In the ToF-MS, the velocity of an ion is inversely proportional to its  $m/z$  ratio. Here the ions are pulsed down the flight tube by the pusher until they reach the reflectron, where focused and reflected towards the MCP detector. The signal from arriving ions at the detector is converted via 4 GHz time-to-digital card (TDC) into a total arrival time distribution (tATD) and processed by MassLynx software (version 4.1, Waters, Manchester, UK).



**Figure 2.2:** Schematic of MoQToF instrument showing n-ESI source, ion optics, drift chamber, quadrupole and time-of-flight analysers. Ion flight path is shown in red dotted line. (A) sample cone, B - extractor cone, C - pre-cell hexapole, D - top hat lens, E - drift cell, F - post-cell hexapole, G - quadrupole, H - collision cell, I - post-collision cell hexapole, J - acceleration and focusing lenses, K - pusher lens stack, L - reflectron, M - point detector (for quadrupole analyser), N - MCP detector (for ToF analyser). The non-standard part of a QToF is highlighted in yellow (drift chamber including drift cell, E and post-cell hexapole, F)

All voltages applied to the original parts of MoQToF (i.e., the source, quadrupole, ToF, and ion optics) are generated by the original instrument power supplies and controlled *via* Masslynx software (version 4.1, Waters, Manchester, UK). Voltages required for the drift cell and associated ion optics are generated by an external in-house built power supply.

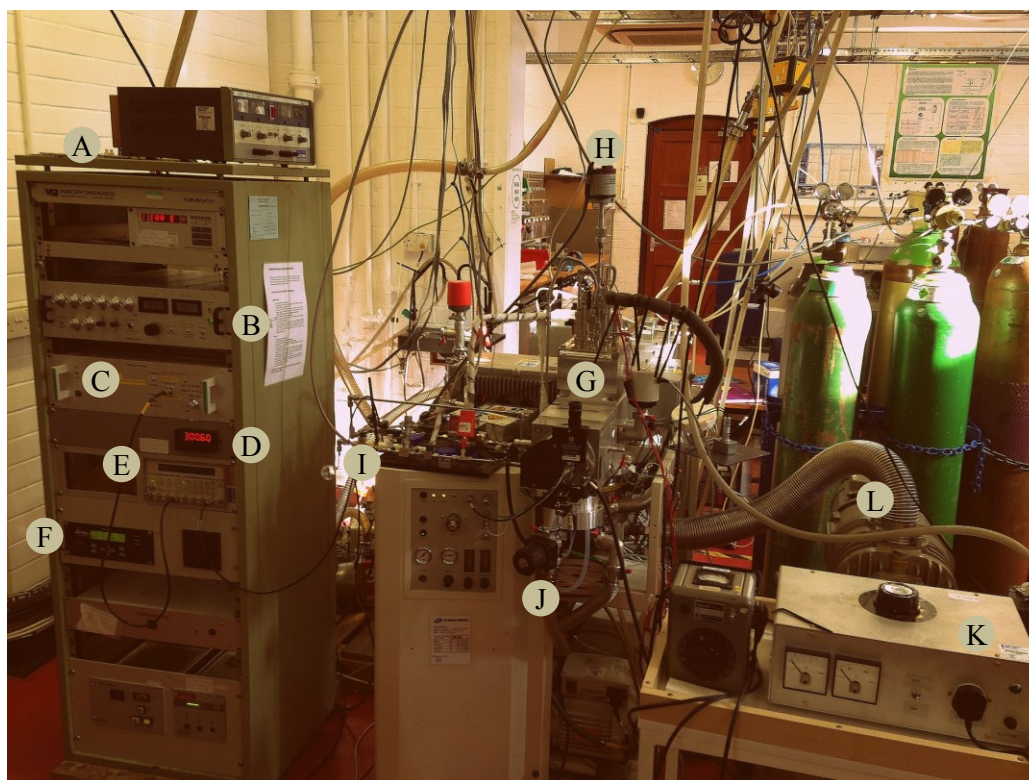
Typical operating vacuum readings attained in the MoQToF with the buffer gas in the cell are tabulated in a Table 2.1. Optimisation of the source pressure was carried

out for each system investigated. The elevated source pressure was achieved by throttling of the source rotary pump (Edwards E28) through insertion of a Speedivalve in line isolation valve to the vacuum line or *via* presence of background argon gas in the source region.

Region	Pressure / mBar
Source	3.3
Analyser	$1.7 \times 10^{-3}$
ToF	$3.0 \times 10^{-7}$

**Table 2.1:** Typical MoQToF instrument pressures.

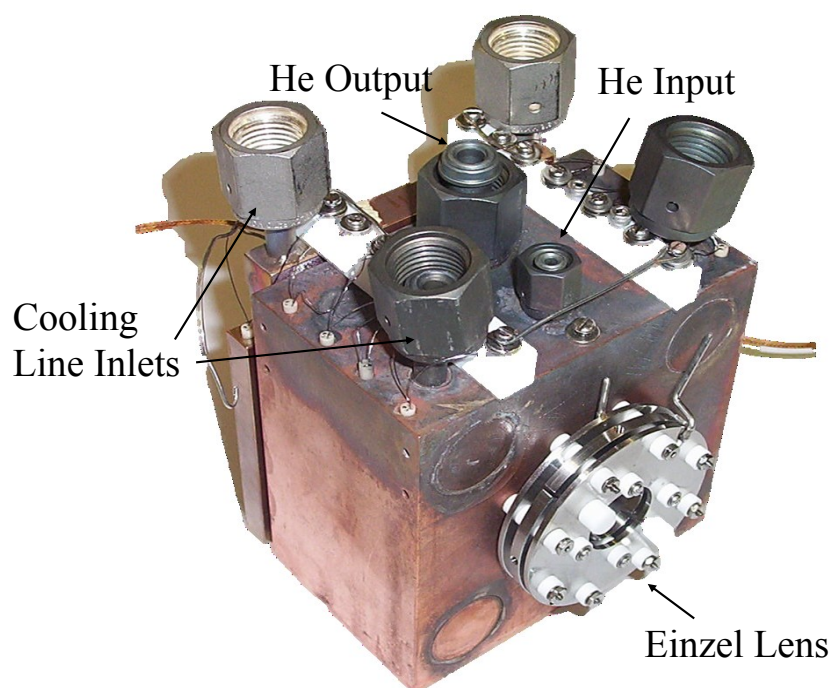
The photograph of the MoQToF instrument (Figure 2.3) highlights the modified elements from standard Q-ToF and additional components required for the mobility experiments.



**Figure 2.3:** Photograph of the in-house linear drift tube IM-MS instrument – the ‘MoQToF’. A - power supply, B - Einzel lens stack controls, C - bipolar pulsar unit, D - drift cell temperature read out, E - Stanford generator, F - drift cell pressure read out, G - drift cell, H - baratron, I - argon gas line, J - source speed valve, K - variable transformers, L - booster pump.

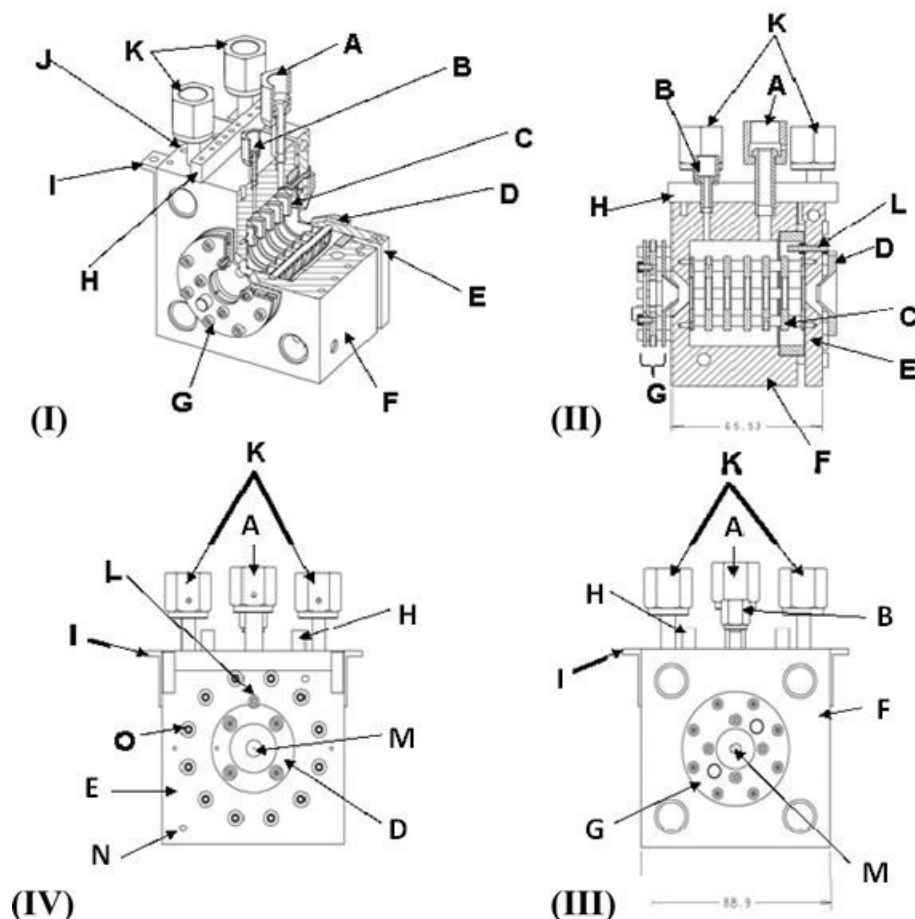
### 2.4.2. Drift Tube Architecture

The low resolution drift tube is based on the design used by the Bowers group at UCSB first reported in 1990 (15). The cell is made from a copper block and a copper end cap separated by a ceramic space ring (photograph of the cell is shown in Figure 2.4). The cell contains five copper drift rings spaced 5.5 mm apart, separated by ceramic spacers, and connected in series by 1 M $\Omega$  resistors. Figure 2.5 shows perspective cross sectional and sectional views of the drift cell used in the MoQToF instrument. Here the linear drift field is achieved by applying a potential difference between the cell body and the end cap with a total drift length being equal to 5.1 cm. Entrance and exit orifices are made from molybdenum discs (30.5 mm o.d. x 15.2 mm i.d x 1.6 mm width) and held in place by copper retainer rings.



**Figure 2.4: Photograph of MoQToF drift cell showing the Einzel lens stack (L1, L2 and L3), gas inlets and outlets.**





**Figure 2.5: Cross sectional drawings for MoQToF drift cell. (I) 3D section through cell (II) section through cell viewed from side (III) front elevation (IV) rear elevation. Parts are labelled as follows: A - Baratron connection, B - gas in, C- drift rings, D - exit lens (L4), E - end cap (C2); F - cell body (C1), G -Einzel lens (L1, L2 and L3); H - heater terminal block, I - mounting brackets, J - heaters; K - cooling line inlets, L - feedthrough to drift rings, M - molybdenum orifice, N - thermocouple mounting, O - cell screws.**

As ions leave the top hat lens (TH1), they are focused and steered into the entrance orifice of the drift cell via an Einzel lens stack. The Einzel lens is used to focus ions without increasing their energy (16). The Einzel lens consists of three stainless steel cylindrical lenses (L1, L2 and L3). The first lens L1 is an initial guiding lens for ions travelling from the pre-cell hexapole, L2 is subdivided into four elements at the  $xy$  plane to allow for directional steering of the ion beam and L3 acts as a final focusing lens. Exiting ions are focused through the final post-cell lens L4 and transferred to the quadrupole chamber.

### **2.4.2.1. Drift Gas**

CP Grade helium gas (99.999% purity, BOC Specialty Gases Ltd, Guildford, UK) permeates through a moisture filter (Inert Gas Purifier, Mykrolis cooperation) and enters the drift cell through a flexible stainless steel bellow. To maintain a correct pressure with additional gas load in this vacuum region a 500 L/s Pfeiffer TMH520 turbomolecular pump (Pfeiffer Vacuum Ltd, Newport Pagnell, UK) backed by the Edwards two-stage E2M40 rotary vane pump (Edwards Vacuum, Crawley, UK) and also a mechanical booster pump is utilised.

### **2.4.2.2. Pressure and Temperature Regulation**

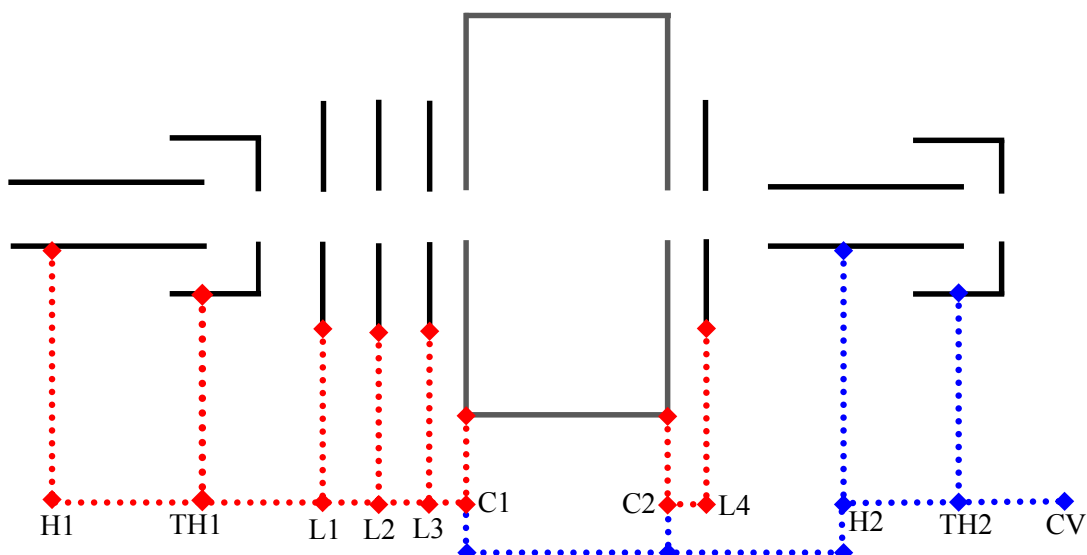
The temperature of the drift cell is carefully monitored using three k-type thermocouples, one located on the cell body and two on the top and bottom of the end cap. The design of the cell enables the temperature to be varied between 80 and 700 K. Heating can be implemented via ceramic rods threaded with tantalum wire (0.25 mm diameter, Goodfellow Cambridge Ltd., Huntingdon, UK) which are located in the copper block of the cell. A current (supplied by two external variable transformers) is passed through the chain of heaters causing the wire to resistively heat the cell. Cooling of the cell can be achieved with a stream of a chilled nitrogen gas that passes through the channels inside the copper block.

The pressure of the drift cell is measured using a MKS Baratron (MKS Instruments, Andover, Massachusetts, USA) attached to the top of the cell *via* a Swagelok VCR fitting.

### **2.4.3. Instrument Tuning Parameters**

The voltages applied to the drift cell and associated ion optics are supplied *via* an external in-house build power supply.

The required voltages and their grounds are presented in the schematic below (Figure 2.6).



**Figure 2.6:** Schematic of MoQToF drift cell voltages with surrounding lenses. Horizontal lines indicate the ground for each voltage. Voltages shown in red dotted line are set manually and supplied by external power supply whereas voltages indicated by a blue line are supplied from the instrument and controlled via the software. H1 is the pre-cell hexapole, H2 the post-cell hexapole, TH1 and TH2 pre and post cell top hat lenses. L1, L2 and L3 pre-cell lenses, L4 post-cell lens. C1 cell body one, C2 end cap voltage. CV is a collision voltage.

All voltages float above the collision cell voltage (CV) i.e. increasing its value by 4 Volts will in effect increase all other voltages by 4 Volts. Additionally the pre-cell lenses (H1, TH1, L1, L2 and L3) are referenced to cell body one (C1) and the post-cell lens (L4) is referenced to the end cap voltage (C2). The voltage ranges of each tuning supply and typical settings used are tabulated in *Table 2.2*.

Tuning	Voltage Range /	Output Voltage /	Typical Setting
H1	0 to 200	CV + C1 + H1	35 (on C1 – 105)
TH1	0 to 200	CV + C1 + TH1	15 (on C1 – 85)
L1	0 to -350	CV + C1 + L1	-40 (on C1 – 30)
L2	0 to -350	CV + C1 + L2	-80 (on C1 – 30)
L3	+50 to -50	CV + C1 + L3	20 (on C1 – 90)
C1	0 to 200	CV + C1	70 (ref CV)
C2	0 to 75	CV + C2	10 (ref CV)
L4	+50 to -50	CV + C2 + L4	-9 (on C2 – 1)

**Table 2.2:** Voltage requirements and typical drift cell lens stack settings for MoQToF for IM-MS experiment.

Typical instrument settings used in the presented work are shown in Table 2.3.

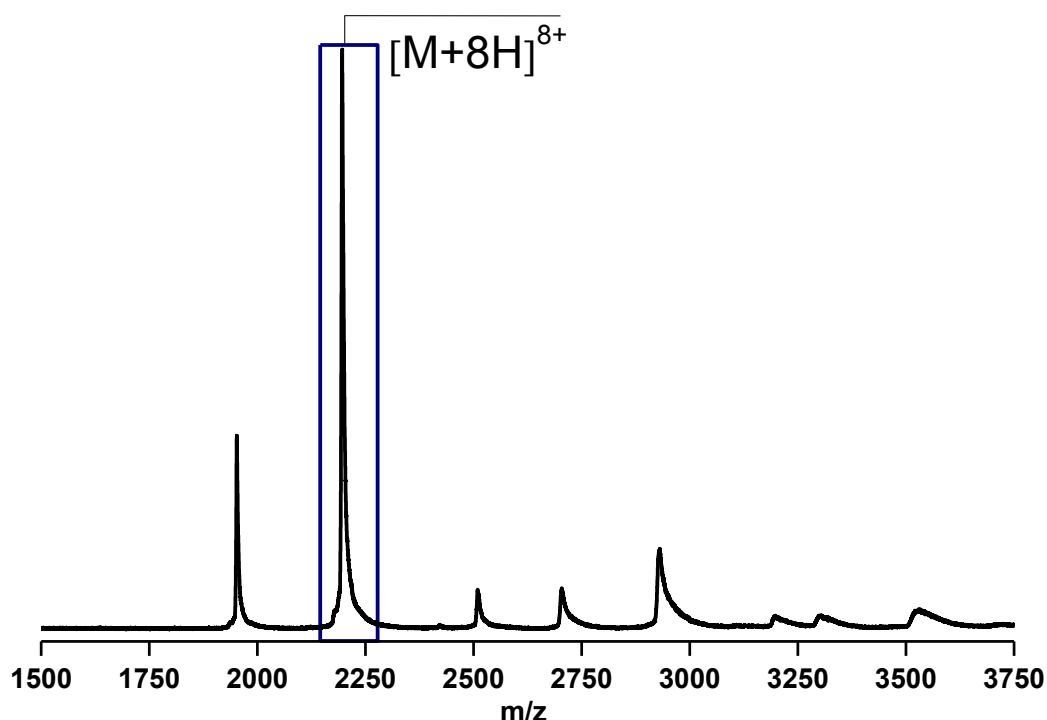
Instrument Parameters	Settings
Capillary / kV	1.0-2.0
Cone	60
Extractor	120
RF Lens	1.35
Source Temperature / °C	80
Collision Energy	2.3
Ion Energy	2.0
Steering	0.00
Entrance	39.9
Pre-filter	10.8
Transport	4.1
Aperture 2	12.6
Acceleration	200
Focus	1
Tube Lens	80
Guard	64.3
TOF / kV	7.20
Reflectron	35.00
Pusher Cycle Time / $\mu$ s*	120
Pusher Frequency / Hz	8333.3333
Multiplier	651
MCP	3000
TDC Start / mV	300.0
TDC Stop / mV	40.0
TDC Threshold	0.0

**Table 2.3:** Typical settings for the MoQToF in positive polarity. Values are shown in volts, unless stated otherwise. Extractor and sample cone voltage settings are the read outs given at cell drift potential 60V. (\*) The pusher cycle time value is adjusted according to  $m/z$  range.

#### 2.4.4. Data Acquisition and Analysis: an example

The following experiment demonstrates the procedure for acquisition and analysis of IM-MS data on the in-house MoQToF instrument. Proof of concept is illustrated by study of holo-myoglobin protein in the presence of three different drift gases (helium, nitrogen, and argon) in the drift cell.

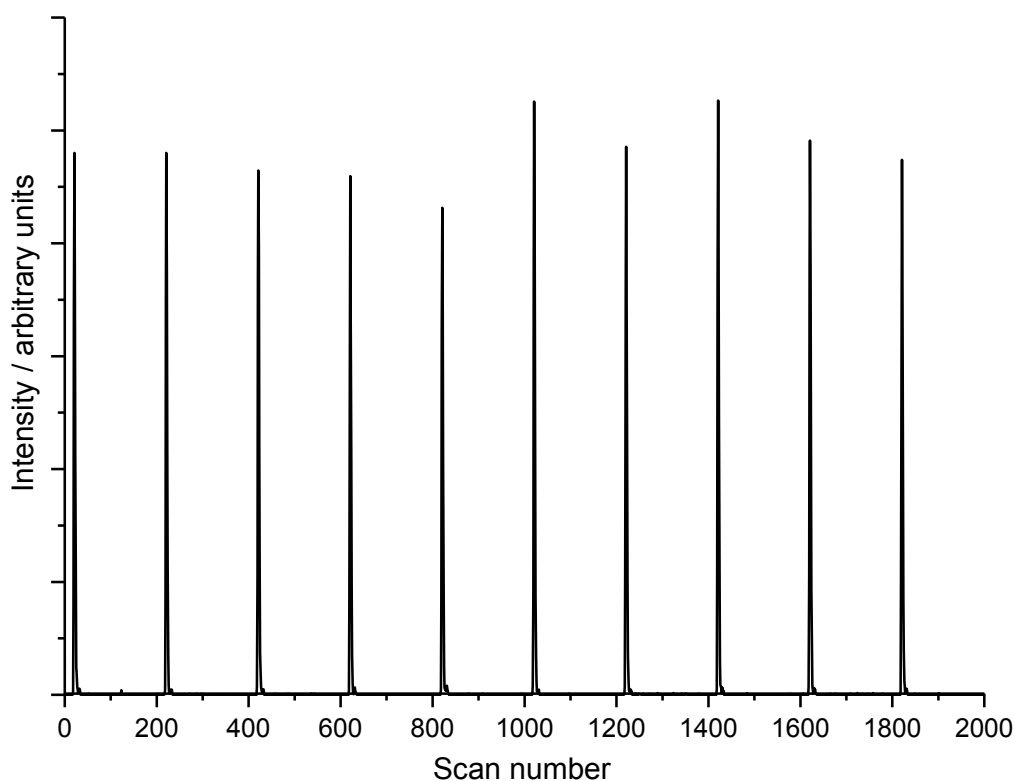
Mass spectra were obtained for myoglobin sprayed from buffered solution conditions (10 mM ammonium acetate, pH 6.8) in the presence of different drift gases in the drift cell (Figure 2.7 shows a mass spectrum of myoglobin obtained in the presence of helium). Irrespective of the drift gas used, the resulting mass spectra were very similar. In the buffered solution, charge state distributions of the heme bound holo-myoglobin monomer were recorded with  $7 \leq z \leq 10$  ( $M_R = 17566.8 \text{ Da} \pm 0.7$ ) and additional multimeric species were also observed. In all cases, the most intense peak corresponds to the monomeric  $[M + 8H]^{8+}$  holo-myoglobin ( $m/z = 2197.0$ )



**Figure 2.7:** n-ESI spectrum of myoglobin sprayed from buffered solution conditions (10 mM ammonium acetate, pH 6.8) obtained in helium.

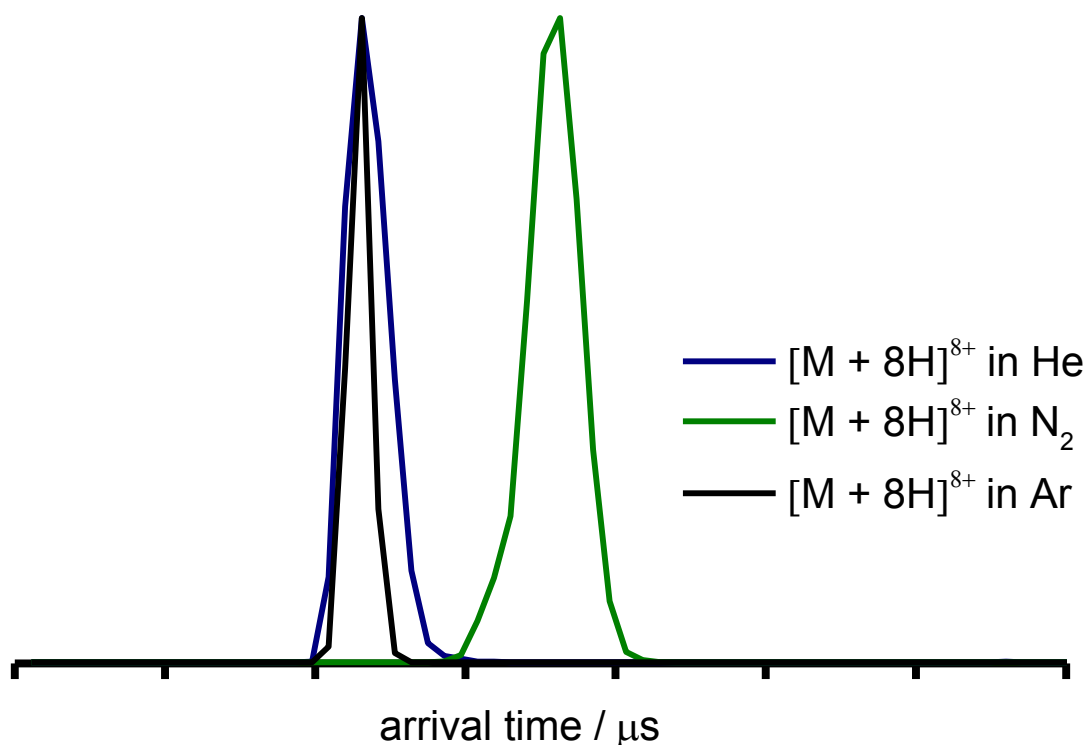
In mobility mode ion arrival time distributions (ATD) were recorded by synchronization of the release of ions into the drift cell with mass spectral acquisition. A total arrival time distribution (tATD) was generated once every 200 scans, where each 200 scans sums to 200 ‘pushes’ of ToF pusher plate (Figure 2.8). A minimum of ten tATD (2000 scans) was acquired for myoglobin at eight different

drift voltages (the electric potential difference across the cell): 60 V, 50 V, 40 V, 35 V, 30 V, 25 V, 20 V and 15 V.



**Figure 2.8: Myoglobin total ion ATD at a drift voltage of 30 V recorded in helium. Each peak is a collation of 200 scans.**

The ion mobility experiment consisting of eight drift voltages requires approximately one hour to be completed due to software waiting times and the time needed to alter the drift voltage. From the tATD a mass spectrum can be produced (Figure 2.7) and from this a deconvoluted ATD of an individual  $m/z$  peak generated. Data for each drift voltage is copied from Mass Lynx software (version 4.1, Waters, Manchester, UK) into an Excel spreadsheet where, the average arrival time ( $t_a$ ) is obtained by multiplying the average scan number of selected species by the pusher period value for each drift voltage.



**Figure 2.9:** Arrival time distributions acquired for the  $[M + 8H]^{8+}$  holo-myoglobin peak ( $m/z = 2197.0$ ) taken at a drift voltage of 30 V at 300K in helium, nitrogen and argon.

ATDs acquired for the  $8^+$  holo monomer ( $m/z = 2197.0$ ) at a drift voltage of 30 V are shown in Figure 2.9. A single resolvable species is observed in each of the three buffer gases, although the actual time spent in the drift cell varies, due to the different pressures used, as well as the varying mobility in each gas. There are differences in the width of the ATD for each gas. This is best seen by comparing the ATD in helium with that in argon where the average drift times are similar but the width of the peaks differs; This observation may be attributed to different gas pressures in the drift cell and the surrounding chamber and/or the nature of the buffer gas dependent collisions on the conformational spread of the ions.

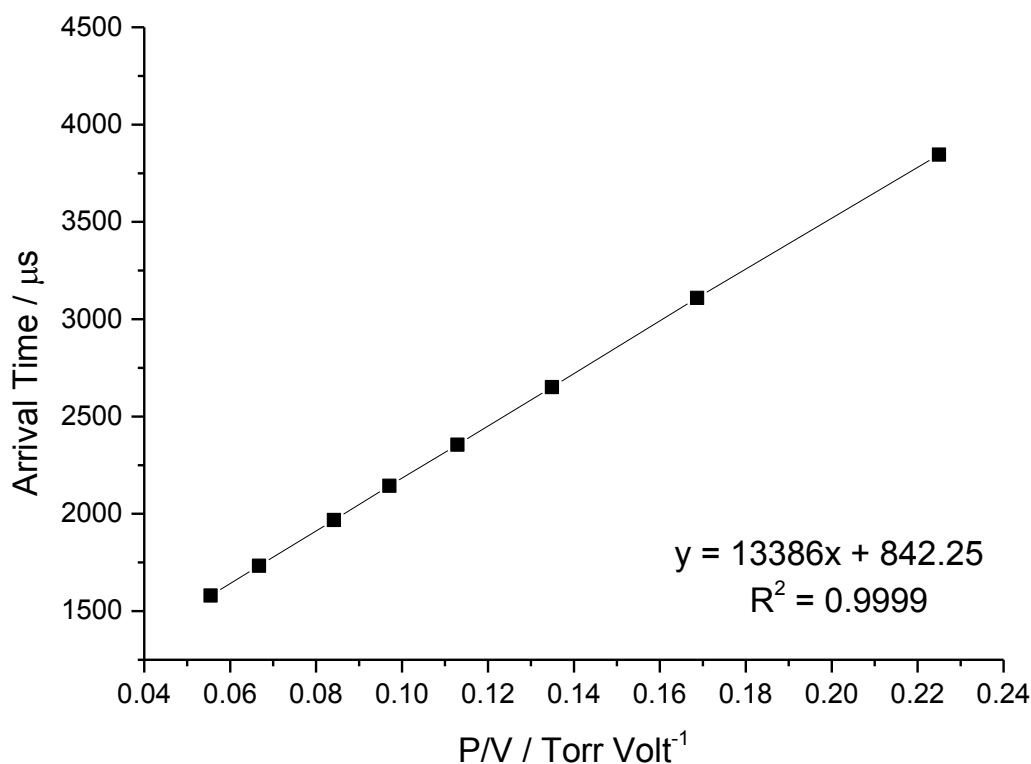
The average arrival time,  $t_a$ , combines the time taken to traverse the drift cell ( $t_d$ ) and the time spent outside the drift cell until the ions reach the detector (known as the dead time  $t_0$ ).

$$t_a = t_d + t_0$$

**Equation 2.1**

Typically  $t_0$  are less than 1 ms. Average arrival times are plotted versus pressure/drift voltage ( $P/V$ ) where an intercept is equal to  $t_0$  and the slope of gradient is inversely proportional to the mobility ( $1/K$ ) (Figure 2.10). The pressure and temperature of the gas present in the drift cell is measured and recorded regularly, at the start and end of each mobility event each drift voltage and the average of each, is taken. The  $t_d$  can be calculated from the equation 2.2 where  $K_0$  is the reduced mobility,  $L$  is the length of the drift cell, and  $P$  and  $T$  are averaged experimental buffer gas pressure and temperature corrected to standard conditions of temperature ( $T_0$  273.15 K) and pressure ( $P_0$  760 Torr).

$$t_d = t_a - t_0 = \frac{L^2 T_0 P}{K_0 T P_0 V} \quad \text{Equation 2.2}$$



**Figure 2.10:** Graph of  $P/V$  versus arrival time for the  $[M+8H]^{8+}$  holo-myoglobin ion obtained in helium.

The goodness-of-fit value  $R^2$  value should be above  $> 0.9995$ . If the fit is good then the gradient can be used to determine the rotationally averaged CCS of the ion using



Equation 2.3.  $\Omega$  is the rotationally averaged CCS,  $z$  is the integer ion charge,  $e$  is electron charge ( $1.602 \times 10^{-19}$  C),  $N$  is the number density of the drift gas,  $\mu$  is the reduced mass of an ion- neutral gas pair,  $k_B$  is the Boltzmann constant ( $1.381 \times 10^{-23}$  J K<sup>-1</sup>),  $T$  is the effective temperature of the drift gas and  $K_0$  is reduced mobility.

$$\Omega = \frac{3ze}{16N} \left( \frac{2\pi}{\mu k_B T} \right)^{1/2} \frac{1}{K_0} \quad \text{Equation 2.3}$$

Under buffered conditions where the charge state distribution of holo-myoglobin is centered on a lower  $m/z$  range (Figure 2.4), the corresponding values for CCS for  $6 \leq z \leq 9$  are shown in Figure 2.11.

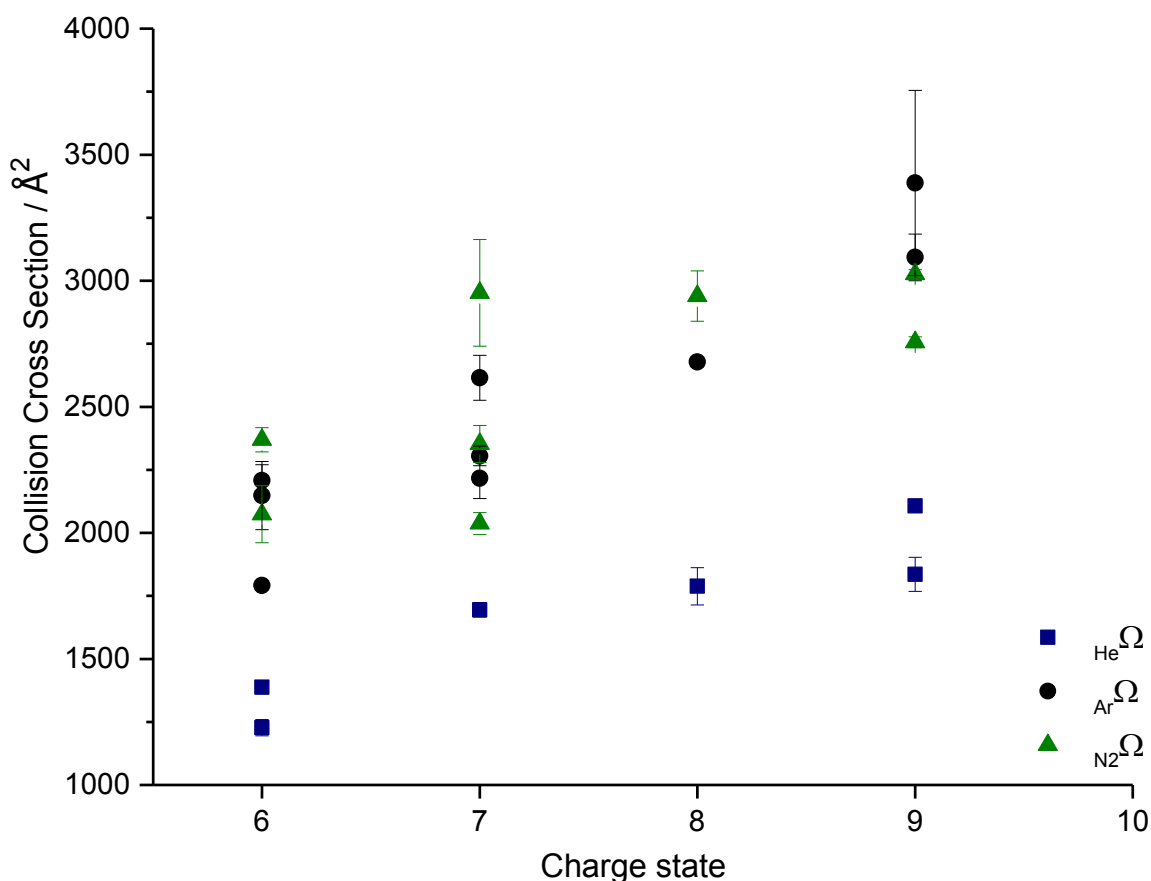


Figure 2.11: CCS of all of the observed charge states of holo-myoglobin obtained in three different drift gases (helium, nitrogen and argon) from pH 6.8 solution at 300 K. The error bars are the standard deviation values from an average of three replicates, referring to the maxima of the fitted peaks.

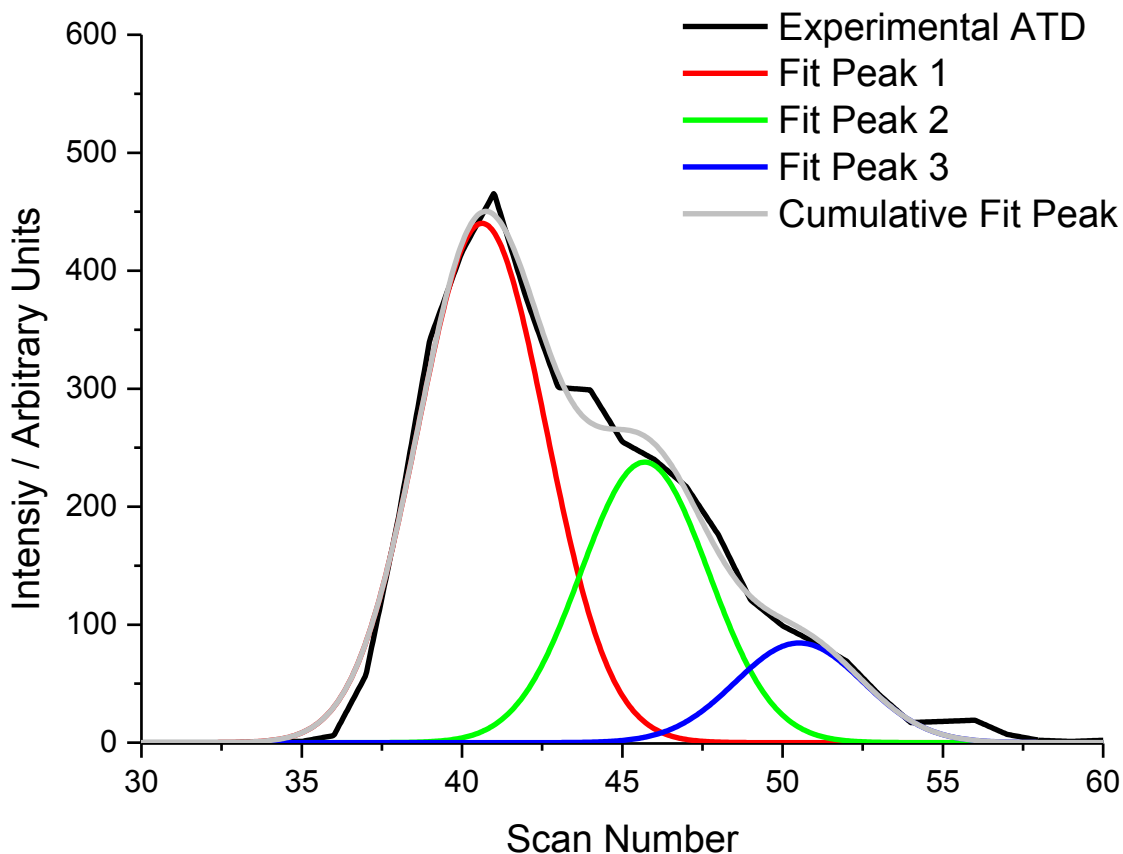
For conformations of myoglobin with charge states that are considered to be ‘native-like’, the dominant factors in determining CCS are the mass and the polarizability of the buffer gas. The larger, diatomic more polarizable molecule effectively lengthens the drift time, resulting in correspondingly larger values of CCS. Diatomic nitrogen gas will interact more strongly at short distances with the molecular ion than argon or helium: each nitrogen atom is actually negatively charged, and the center of the nitrogen molecule is positively charged, molecular nitrogen is also larger than argon. In addition, it may be that the greater effects of momentum transfer into rotational and vibrational energy for the diatomic gas compared to the monatomic gas causes the increase in mobility for nitrogen. When we compare the CCS for each charge state the relationship of  $_{\text{He}}\Omega < _{\text{Ar}}\Omega < _{\text{N}_2}\Omega$  is found. Where  $z = 6$ , the  $_{\text{He}}\Omega$  shows two conformers, both of which indicate a collapsed state, whereas for  $z = 7, 8$ , and  $9$  there is little change in the values found for  $_{\text{He}}\Omega$ , although a more extended form appears where  $z = 9$ . For the heavier gases, at all other charge states around the dominant  $[\text{M} + 8\text{H}]^{8+}$  species, there is more than one conformer present. This is an interesting finding, suggesting that the stability of the holo- protein in a single resolvable fold is favoured for  $[\text{M} + 8\text{H}]^{8+}$  and starts to be more disrupted for other charge states.

In summary, different drift gases provide different CCS for each observed charge state of the same protein. In addition different gases can reveal different conformations and populations of molecular ions present in the gas-phase, particularly for native forms.

#### 2.4.4.1. Multiple ATD Peak Fitting

For many species, the presence of peak shoulders or individual peaks that are not base line resolved in ATD obtained is an indication of potential multiple conformers. In this case maxima for the resulting peaks (and hence the arrival times) can be achieved by fitting a Gaussian distribution using Origin 8.5.1 (OriginLab, Northampton, MA, USA). The Gaussian function used is specified in Equation 2.4, where  $y_0$  is the offset,  $A$  is the area,  $x_c$  the centre of the fitted peak and  $w$  the peak width:

$$y = y_0 + \frac{A}{w\sqrt{\frac{\pi}{2}}} e^{-2\frac{(x-x_c)^2}{w^2}} \quad \text{Equation 2.4}$$



**Figure 2.12:** Experimental ATD of holo-myoglobin peak ( $m/z = 2510.7$ ) taken at a drift voltage of 20 V at 300 K in nitrogen. Three species can be resolved. Gaussian fits (red, green and blue lines) allow the centre of each species to be calculated shown in a dashed line.

The centre of the Gaussian fitted peak  $x_c$  is then converted into average arrival time for each species at each drift voltage. For each peak, goodness-of-fit  $R^2$  value above  $R^2 > 0.90$  is accepted. A plot of arrival time versus P/V is obtained as in Figure 2.10 and the CCS are calculated.

## **2.5. Travelling Wave based IM-MS - Synapt HDMS**

Travelling wave ion mobility mass spectrometry (TW IM-MS) is a relatively new IMS method implemented in the Synapt HDMS mass spectrometer (Waters, Manchester, UK) (14). Despite a successful commercialisation in 2006, its fundamentals have been explained only qualitatively. Shvartsburg has explored the TW IM-MS operation using theoretical models and numerical simulations of ion dynamics to formally understand the relationship between collision cross section and mobility in a travelling wave device (17), but as yet there is no adoption of such an approach by the IM-MS community. In the absence of a physical model for TW IM-MS, the mobilities (and hence collision cross sections) of ions need to be extracted *via* an external empiric calibration using a set of standards previously measured in helium by drift tube ion mobility mass spectrometry (DT IM-MS) (18-20). However in the short time Synapt technology has been available; it has made an attractive platform for applications to structural biology analysis (21-24) and complex mixture separations (25,26).

### **2.5.1. Instrument Layout and Operation**

Ions are produced in the n-ESI source via a 'Z-spray' and transferred through the T-wave ion guide into the quadrupole analyser. Three travelling wave ion guides (TWIGs), which form the Tri-wave, are placed between the quadrupole and time-of-flight mass analyser. Ions arriving in the Tri-wave device are stored in the Trap T-wave and then released into the IMS T-wave where the mobility separation occurs. A schematic representation of this instrument is shown in Figure 2.13. The mobility-separated packets of ions are then transferred via the Transfer T-wave to an orthogonal acceleration ToF mass analyser.

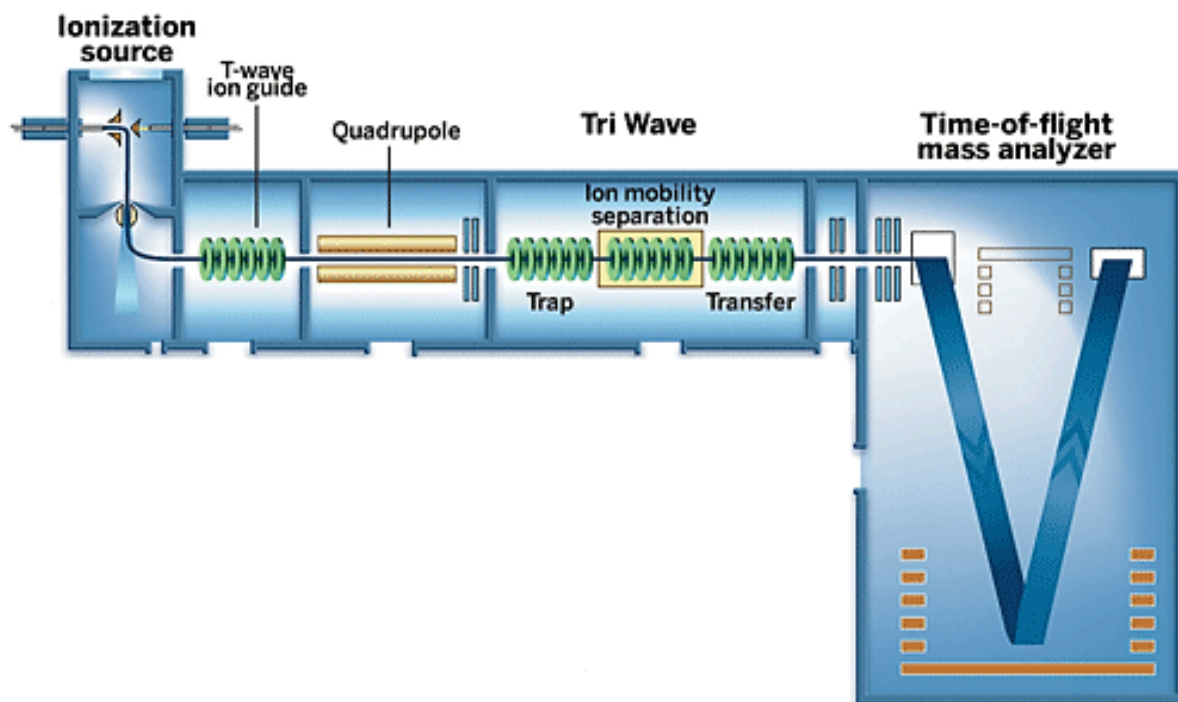
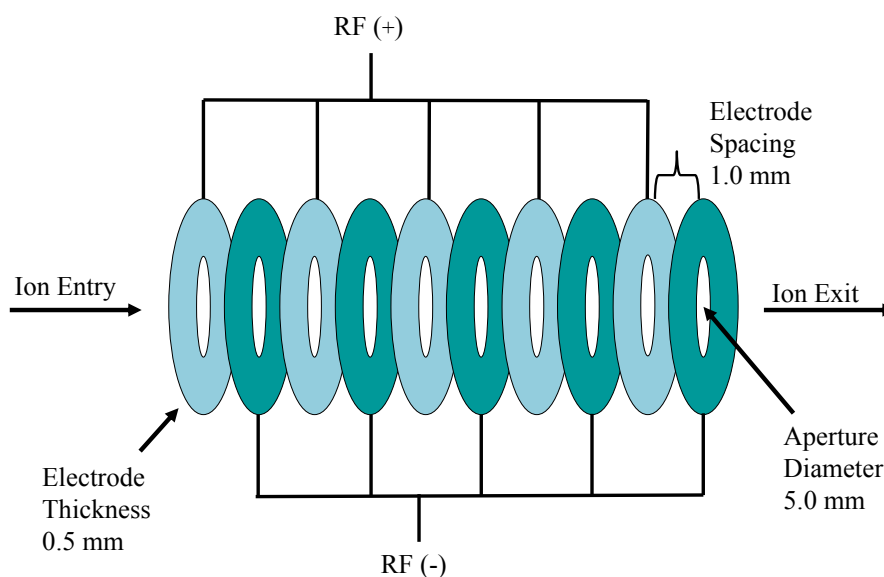


Figure 2.13: Schematic of Synapt HDMS instrument. Image taken from Waters website (27).

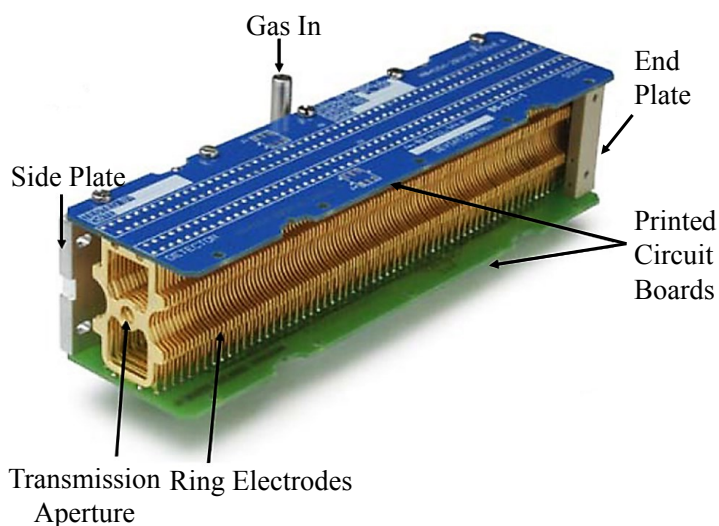
### 2.5.1.1. Travelling Wave Ion Guide Topology

The TWIG device comprises a series of ring electrodes arranged orthogonally to the ion transmission axis, as shown in Figure 2.14 (28). Opposite phases of an RF voltage are applied to consecutive electrodes to provide a radial confinement for ions (29). To propel ions through device a transient direct current (DC) pulse is sequentially superimposed on the top of RF potential. This provides a moving electric field of ‘travelling waves’ on which ions can ‘surf’, reducing their residence time in the device.



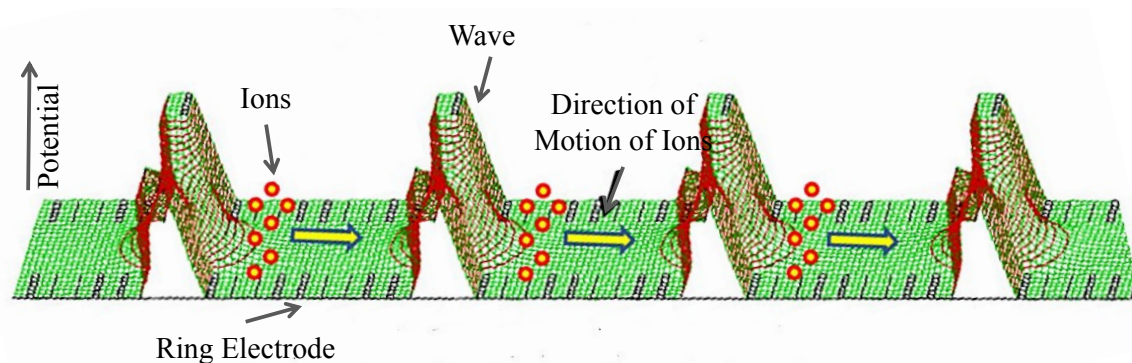
**Figure 2.14: Schematic diagram of an RF-only stacked ring ion guide. Image taken from Giles *et al.* (28).**

The mobility TWIG device (also referred as the travelling wave IM separator TWIMS) consists of 122 electrodes supported between two printed circuit boards (PCBs). A photograph of the T-wave device is shown in Figure 2.15, which has an overall length of 185 mm.



**Figure 2.15: Photograph of T-wave device. A T-wave device consists of a sequence of ring-shaped electrodes, supported on PCBs that deliver both RF and DC voltages. Image taken from the Waters website (27).**

The entrance and exit apertures of the TWIMS are 2 mm in diameter. The principle of operation is illustrated in Figure 2.16 by SIMION study, ions roll over the wave top. Mobility separation is achieved because ions of high mobility roll over the wave less often than species of low mobility so the transit time through the device is shorter. The IMS cell is operated at up to 1 mBar nitrogen gas.



**Figure 2.16: Traveling wave propelling ions in the Synapt system. SIMION plots show a packet of ions surfing on the front of the travelling wave as it passes along the SRIG. Image taken from the Waters website (27).**

The Trap and Transfer TWIGs are 100 mm long and consist of 33 electrode pairs with RF applied. Typically these cells are operated at argon pressures in the low  $10^{-2}$  mBar range.

### 2.5.1.2. Instrument Tuning Parameters and Typical Pressures

Typical instrument settings used in the presented work are shown in Table 2.4.

Instrument Parameters	Settings
Capillary / kV	1.0-2.0
Cone / V	60
Extractor	2.0
Source Temperature / °C	30
Nanoflow Gas Pressure / Bar	0.2
Aperture 1	1.0
Pre-filter	2.0
Ion Energy	2.0

<b>Instrument Parameters</b>	<b>Settings</b>
Trap Collision Energy	5.0
Transfer Collision Energy	3.0
Trap Gas Flow / mL/min	1.50
Source Gas Flow / mL/min	0.00
IMS Gas Flow / mL/min	10.00
Detector	1800
Pusher Interval / $\mu$ s	90
Pusher Width	4
Acceleration 1	70.0
Acceleration 2	200.0
Aperture 2	70.0
Transport 1	70.0
Transport 2	70.0
Steering	0.0
Tube Lens	75
Pusher	925.0
Puller	647.0
TOF / kV	9.10
Reflectron / kV	2.16
Trap DC Entrance	3.0
Trap DC Bias	13.0
Trap DC Exit	0.0
IM DC Entrance	2.0
IM DC Exit	2.0
Transfer DC Entrance	1.0
Transfer DC Exit	1.0
IMS Wave Velocity / m/s	250
IMS Wave Height	13.0
Transfer Wave Velocity / m/s	247
Transfer Wave Height	4.0
Trap Release Time / $\mu$ s	500
Trap Trap Height	30.0
Trap Extract Height	10.0

**Table 2.4:** Typical settings for Synapt HDMS in positive polarity. Values are shown in volts, unless stated otherwise.



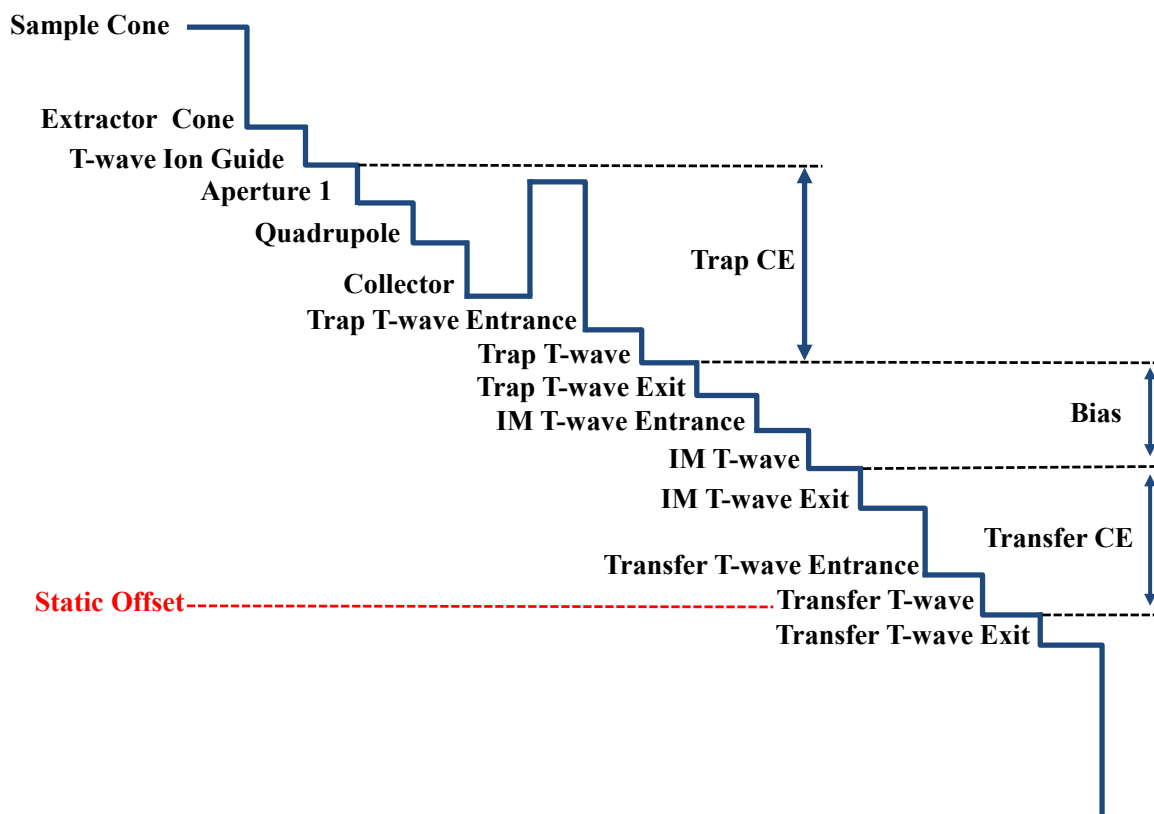


Figure 2.17: Schematic of the potential gradient along the Synapt HDMS instrument and the potential differences defining the Trap and Transfer collision energies (CE) and the Bias potential. All potentials are referenced to the Static Offset voltage typically set to 120 V. Image and caption adapted from Michaelevski *et al.* (30).

Typical operating vacuum readings attained in Synapt HDMS during mobility experiment are tabulated in a Table 2.5.

Region	Pressure / mBar
Backing	4.38
Source	$3.53 \times 10^{-3}$
Quadrupole	$3.42 \times 10^{-5}$
Trap	$1.85 \times 10^{-2}$
IMS	1.05
TOF	$1.12 \times 10^{-6}$

Table 2.5: Typical Synapt instrument pressures.

### 2.5.2. Calculation of CCS using TW IM-MS data

The mobility of ions through T-wave devices is not as directly related to collision cross section as in DT IM-MS instruments. Despite this, with a careful use of standards previously measured in helium via DT IM-MS, it is possible to convert mobilities measured with nitrogen as a drift gas in Synapt instrumentation into mass selected ‘helium based’ collision cross sections. Several ion mobility cross-sectional calibration procedures have been described previously (18,21,22,31-35).

In this thesis, where TW IM-MS was employed (in Chapters 3 and 4), ion mobility drift times were calibrated using multiply charged ions of cytochrome c (equine heart) and myoglobin (0.2 mg/mL, in 50/50 H<sub>2</sub>O/acetonitrile and 0.1% formic acid). The mobility drift times of the calibrant ions were corrected for their mass-dependant flight time (the time from when the ion exits the transfer T-Wave until it reaches ToF and is proportional to the  $\sqrt{\frac{m}{z}}$  of the ion), using the following Equation 2.5:

$$t'_d = t_d - \frac{c\sqrt{\frac{m}{z}}}{100} \quad \text{Equation 2.5}$$

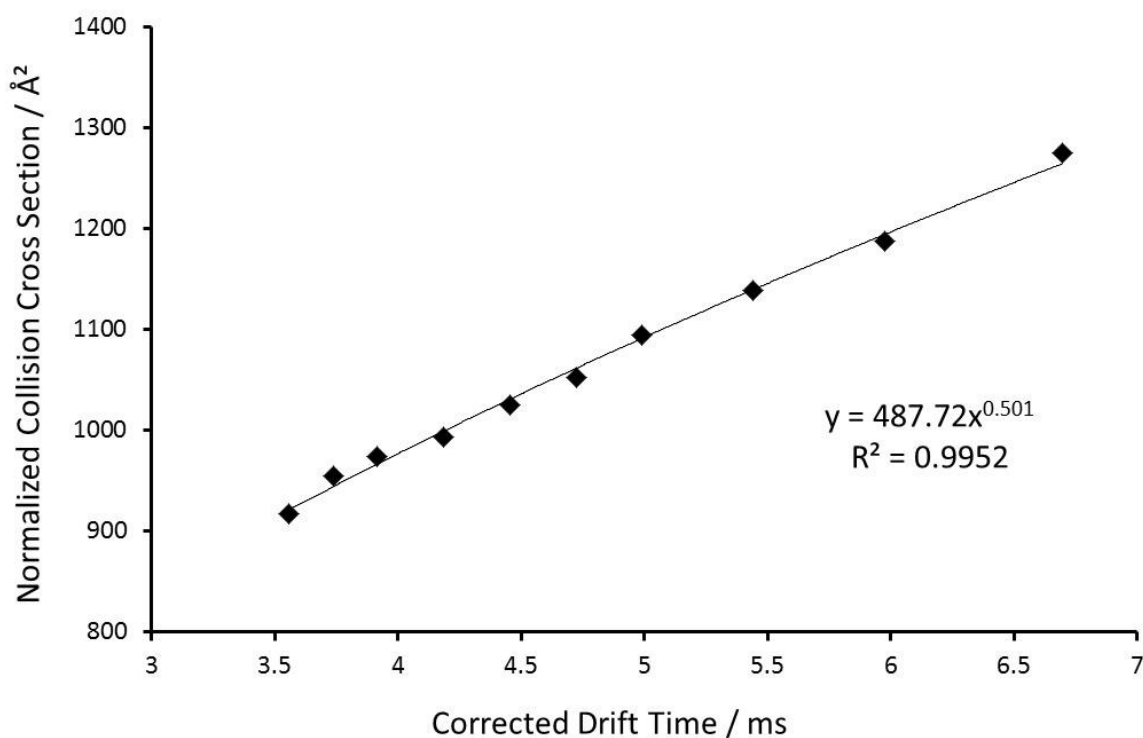
where  $t'_d$  is the corrected drift time in ms,  $t_d$  is the experimental drift time in ms,  $\frac{m}{z}$  is the mass-to-charge ratio of the observed ion and  $c$  is the Enhanced Duty Cycle (EDC) delay coefficient;  $c$  is instrument-dependant with a typical value between 1.4 and 1.6 (here 1.5). Using Clemmer’s Cross-Section Database (36), the calibrant CCS were corrected ( $\Omega'$ ) for both ion charge state ( $z$ ) and reduced mass ( $\mu$ ) to generate:

$$\Omega' = \frac{\text{published } \Omega}{z \sqrt{\frac{1}{\mu}}} \quad \text{Equation 2.6}$$

The corrected collision cross section values ( $\Omega'$ ) were plotted against the corrected mobility drift time ( $t'_d$ ) values of the calibrant ions and a power trend ( $y = Ax^N$ ) was fitted (plot is shown in Figure 2.18). The experimental CCS measurements for each

resolvable species at a given charge state were converted to estimated CCS by calculating:

$$\Omega' = t'_d{}^N \times A \times z \times \left( \sqrt{\frac{1}{\mu}} \right) \quad \text{Equation 2.7}$$



**Figure 2.18:** Normalized collision cross sections vs corrected mobility drift times for equine myoglobin ions,  $z = 13 - 22$  (0.2 mg/mL, in 50/50 H<sub>2</sub>O/acetonitrile and 0.1% formic acid). The plot is an example used to create a cross-section calibration for the Synapt instrument.

## 2.6. Mass Calibration

Calibration was performed using a 1 mg/mL solution of sodium iodide (Sigma Aldrich, Dorset, UK) dissolved in 50% isopropanol. The metal salt calibrant provides a series of monoisotopic cluster peaks corresponding to Na<sub>n+1</sub>I<sub>n</sub> ions, covering appropriate  $m/z$  range.

## 2.7. Optimization of experimental conditions

The goal of native mass spectrometry is to preserve quaternary protein structures to investigate protein complex topology and dynamics. The experimental parameters (voltages and pressures) were screened systematically to achieve an optimum signal and most importantly to avoid any instrument-induced structural transitions. Injection energy monitoring, where the potential pushing ions into the drift cell region is varied as well as heating of the drift cell (thermal unfolding) was used to explore the conformational space of proteins, beyond native conditions. The full tuning of MoQToF instrument parameters was performed for every experiment, including the drift cell voltages (L1, L2, L3, L4 and H1, TH1), sample and extractor cone potential. In Synapt HDMS instrument, particular attention was given to IMS wave height, IMS wave velocity and bias potential settings, as well as sample and extractor cone values. For both instruments used the source pressure was elevated systematically to obtain an improved transition of ions. The capillary potential was adjusted to provide optimum signal. The precise value of voltage used depended mainly on the solution composition and spray geometry but voltages were typically in the range of 1.0-2.0 kV. In case of samples prone to aggregation a nitrogen backing gas was used to maintain a stable spray. Unless stated otherwise the source block was heated 80°C to aid desolvation process.

Aqueous ammonium acetate salt concentrations were used in range of 10 mM to 500 mM, with neutral pH. From many biophysical studies it is clear that quaternary protein structures can often be preserved under these conditions. During the ESI process, the volatile buffer easily desolvates, rendering 'naked' protein ions, albeit substantially less charged than in organic ESI solvents as the surface is more compact in folded species. Salting out of the protein sample solutions was one of the essential steps in obtaining a clean mass spectrum. Dialysis method, either slide-A-lyzer dialysis cassettes (Thermo Scientific, Inc) or micro bio-spin chromatography columns (Bio-Rad Laboratories, Inc), was used for the removal of unwanted salt. Dialysate (buffer) volume, number of dialysate changes, time, dialysate agitation

(stirring) and temperature were the factors that affected the dialysis rate and were optimised for every protein sample.

In the subsequent results chapters the final optimised experimental conditions are specified in the relevant methodology section.

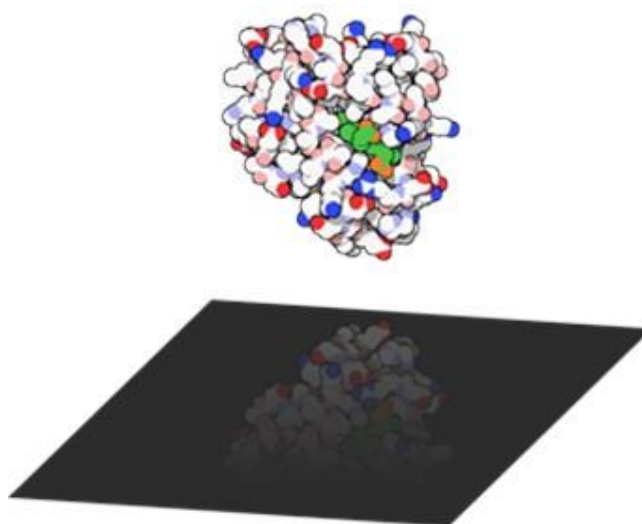
## **2.8. Estimation of collision cross sections from PDB structures**

The mobility and buffer gas specific CCS could be accurately calculated for any set of coordinates (available from RCSB Protein Data Bank) for subsequent comparison with experimental values. This would involve evaluation of orientationally averaged cross sections with an approximate but appropriate treatment of ion-buffer gas collisions. Indeed this can be accomplished, and Jarrold (37-39) and Bowers (40) have made significant breakthroughs in this direction. Currently, MOBCAL (37-39,41) developed by Shvartsburg and Jarrold is widely used to determine the theoretical CCS of biological molecules. This open source software program is based on three different treatments of the ion-buffer gas collision which calculate rotationally averaged cross sections from input coordinate files, derived from X-ray crystallography or NMR studies or from MD simulations. With the exception of some recent work by Shvartsburg (42), all computations rely on averaging the orientation of a set of coordinates and calculating the interaction with a chosen buffer gas. The contained models in MOBCAL are: the projection approximation (PA) (40), exact hard sphere scattering method (EHSS) (37,43), trajectory method (TM) (39) and recently developed a novel projected superposition approximation (PSA) (44-46). The above four methods determine the CCS with helium as the *in silico* buffer gas, which to date somewhat curtails the use of other gases in experiments.

In this thesis only EHSS or TM values are reported and compared with experimental data.

### 2.8.1. Projection Approximation Method

A particularly simple approach is the projection method (PA) (40) which has been employed to obtain rotationally averaged cross sections of proposed molecular structures for over 80 years. The CCS is determined by averaging the geometric projection areas over all possible orientations (Figure 2.19). However this method ignores the long-range interactions and all the details of the scattering process between the ion and buffer gas. The CCS obtained from crystal structure coordinates of some monomeric peptides and small proteins using the PA method were found to be on average up to 20% smaller than those obtained from the other approaches, EHSS and TM (this difference was marginally greater for the larger molecules) (47). In order to address the inadequacy of the PA approach, Wytenbach (40) implemented the use of scaled size parameters for all of the atoms in the molecule, however this still does not address the effects of multiple collisions, and hence for ions of masses greater than 2 kDa, or for particularly crenulated conformations of smaller molecules such as those found for crown ethers, this method will underestimate CCS and is of limited value for larger biomolecules.



**Figure 2.19:** The projection approximation method finds the average ‘shadow’ as a test conformer is rotated through all possible orientations. Image was adapted from the website of the Indiana University Bloomington (36).

### **2.8.2. Exact Hard Sphere Scattering Method**

In the exact hard sphere scattering (EHSS), a CCS is calculated by averaging the momentum transfer cross section which is related to the scattering angles between the incoming and departing gas atom trajectory. This model takes into account scattering and the collision process but does not consider the effects of long range potentials between the buffer gas and the molecular ion. Overall the EHSS model was found to give values that are slightly higher, especially for smaller molecules, than those determined by TM. The EHSS model overestimates the cross sections for small ions simply because it (as well as PA) in MOBCAL has been parameterized for fullerenes and other large ions of similar size, and small ions form shallower ion molecule potentials that are effectively manifested as smaller collision radii, finally the long range interactions included in TM appear to be of less importance for larger ions, but TM properly accounts for the size dependent variation of the interaction potential (47).

### **2.8.3. Trajectory Method**

The trajectory method (TM) (39) is considered to be the most reliable and accurate method for calculating the CCS especially for larger ions. This approach takes into account the long-range interactions and close collisions between the ion and buffer gas atom, as well as the effects of multiple collisions. To perform this calculation, the effective potential must be defined and then the trajectories are run within this potential to obtain the scattering angles.

### **2.8.4. Projected Superposition Approximation Method**

In the framework of a novel projected superposition approximation (PSA), molecular CCS are computed as a projection approximation method, however modified to account for collective size and shape effects (46). The PSA method is a much faster

than the trajectory method as recently demonstrated by Bleiholder and Bowers (44,45), reaching increased efficiency and no loss in accuracy.



## 2.9. References

1. Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D. and Alice, M.B. (1968) Molecular Beams of Macroions. *The Journal of Chemical Physics*, **49**, 2240-2249.
2. Yamashita, M. and Fenn, J.B. (1984) Electrospray ion source - another variation on the free-jet theme. *The Journal of Physical Chemistry*, **88**, 4451-4459.
3. Fenn, J.B. (2003) Electrospray wings for molecular elephants (Nobel lecture). *Angewandte Chemie International Edition* **42(33)**, 3871-3894.
4. Kebarle, P. and Tang, L. (1993) From ions in solution to ions in the gas phase - the mechanism of electrospray mass spectrometry. *Analytical Chemistry*, **65**, 972A-986A.
5. Iavarone, A.T. and Williams, E.R. (2003) Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *Journal of the American Chemical Society*, **125**, 2319-2327.
6. Fernandez de la Mora, J. (2000) Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica Chimica Acta*, **406**, 93-104.
7. Iribarne, J.V. and Thomson, B.A. (1976) On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, **64**, 2287-2294.
8. Dongré, A.R., Jones, J.L., Somogyi, Á. and Wysocki, V.H. (1996) Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model. *Journal of the American Chemical Society*, **118**, 8365-8374.
9. Wilm, M. and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source. *Analytical Chemistry*, **68(1)**, 1-8.
10. Juraschek, R., Dülcks, T. and Karas, M. (1999) Nanoelectrospray-more than just a minimized-flow electrospray ionization source. *Journal of the American Society for Mass Spectrometry*, **10**, 300-308.
11. Fenn, J., Mann, M., Meng, C., Wong, S. and Whitehouse, C. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science*, **246**, 64-71.
12. Ganem, B., Li, Y.T. and Henion, J.D. (1991) Detection of noncovalent receptor-ligand complexes by mass spectrometry. *Journal of the American Chemical Society*, **113**, 6294-6296.
13. McCullough, B.J., Kalapothakis, J., Eastwood, H., Kemper, P., MacMillan, D., Taylor, K., Dorin, J. and Barran, P.E. (2008) Development of an ion mobility quadrupole time of flight mass spectrometer. *Analytical Chemistry*, **80**, 6336-6344.
14. Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinou, K., Bateman, R.H., Bowers, M.T. and 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**, 1-12.

15. Kemper, P.R. and Bowers, M.T. (1990) A hybrid double-focusing mass spectrometer-high-pressure drift reaction cell to study thermal energy reactions of mass-selected ions. *Journal of the American Society for Mass Spectrometry*, **1**, 197-207.
16. Moore, J.H., Davis, C.C. and Coplan, M.A. (2002) *Building Scientific Apparatus* 3rd ed. Westview Press, Oxford.
17. Shvartsburg, A.A. and Smith, R.D. (2008) Fundamentals of traveling wave ion mobility spectrometry. *Analytical Chemistry*, **80**, 9689-9699.
18. Ruotolo, B.T., Benesch, J.L.P., Sandercock, A.M., Hyung, S.-J. and Robinson, C.V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nature Protocols*, **3**, 1139-1152.
19. Leary, J.A., Schenauer, M.R., Stefanescu, R., Andaya, A., Ruotolo, B.T., Robinson, C.V., Thalassinos, K., Scrivens, J.H., Sokabe, M. and Hershey, J.W. (2009) Methodology for measuring conformation of solvent-disrupted protein subunits using T-WAVE ion mobility MS: an investigation into eukaryotic initiation factors. *Journal of the American Society for Mass Spectrometry*, **20**, 1699-1706.
20. Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C.V. and 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**, 9557-9565.
21. Ruotolo, B.T., Giles, K., Campuzano, I., Sandercock, A.M., Bateman, R.H. and Robinson, C.V. (2005) Evidence for macromolecular protein rings in the absence of bulk water. *Science*, **310**, 1658-1661.
22. Ruotolo, B.T., Hyung, S.J., Robinson, P.M., Giles, K., Bateman, R.H. and Robinson, C.V. (2007) Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes. *Angewandte Chemie International Edition*, **46**, 8001-8004.
23. Smith, D.P., Giles, K., Bateman, R.H., Radford, S.E. and Ashcroft, A.E. (2007) Monitoring copopulated conformational states during protein folding events using electrospray ionization-ion mobility spectrometry-mass spectrometry. *Journal of the American Society for Mass Spectrometry*, **18**, 2180-2190.
24. Riba-Garcia, I., Giles, K., Bateman, R.H. and Gaskell, S.J. (2008) Evidence for structural variants of a- and b-type peptide fragment ions using combined ion mobility/mass spectrometry. *Journal of the American Society for Mass Spectrometry*, **19**, 609-613.
25. Vakhrushev, S.Y., Langridge, J., Campuzano, I., Hughes, C. and Peter-Katalinic, J. (2008) Ion mobility mass spectrometry analysis of human glycourinome. *Analytical Chemistry*, **80**, 2506-2513.
26. Bagal, D., Zhang, H. and Schnier, P.D. (2008) Gas-phase proton-transfer chemistry coupled with TOF mass spectrometry and ion mobility-MS for the facile analysis of poly(ethylene glycols) and PEGylated polypeptide conjugates. *Analytical Chemistry*, **80**, 2408-2418.
27. <http://www.waters.com>.
28. Giles, K., Pringle, S.D., Worthington, K.R., Little, D., Wildgoose, J.L. and Bateman, R.H. (2004) Applications of a travelling wave-based radio-

- frequency-only stacked ring ion guide. *Rapid Communications in Mass Spectrometry*, **18**, 2401-2414.
29. Gerlich, D. (1992) Inhomogeneous RF-fields - a versatile tool for the study of processes with slow ions. *Journal of Advances in Chemical Physics*, **82**, 1-176.
  30. Michaelevski, I., Kirshenbaum, N. and Sharon, M. (2010) T-wave ion mobility-mass spectrometry: basic experimental procedures for protein complex analysis. *Journal of Visualized Experiments*, 10.3791/1985.
  31. Williams, J.P. and Scrivens, J.H. (2008) Coupling desorption electrospray ionisation and neutral desorption/extractive electrospray ionisation with a travelling-wave based ion mobility mass spectrometer for the analysis of drugs. *Rapid Communications in Mass Spectrometry*, **22**, 187-196.
  32. Williams, J.P., Bugarcic, T., Habtemariam, A., Giles, K., Campuzano, I., Rodger, P.M. and Sadler, P.J. (2009) Isomer separation and gas-phase configurations of organoruthenium anticancer complexes: ion mobility mass spectrometry and modeling. *Journal of the American Society for Mass Spectrometry*, **20**, 1119-1122.
  33. Scarff, C.A., Thalassinou, K., Hilton, G.R. and 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**, 3297-3304.
  34. Thalassinou, K., Grabenauer, M., Slade, S.E., Hilton, G.R., Bowers, M.T. and Scrivens, J.H. (2008) Characterization of phosphorylated peptides using traveling wave-based and drift cell ion mobility mass spectrometry. *Analytical Chemistry*, **81**, 248-254.
  35. Smith, D., Knapman, T., Campuzano, I., Malham, R., Berryman, J., Radford, S. and Ashcroft, A. (2009) Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *European Journal of Mass Spectrometry*, **15**, 113-130.
  36. <http://www.indiana.edu/~clemmer>.
  37. Shvartsburg, A.A. and Jarrold, M.F. (1996) An exact hard-spheres scattering model for the mobilities of polyatomic Ions. *Chemical Physics Letters*, **261**, 86-91.
  38. Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C. and Jarrold, M.F. (1996) Structural information from ion mobility measurements: effects of the long-range potential. *The Journal of Physical Chemistry*, **100**, 16082-16086.
  39. Shvartsburg, A.A., Schatz, G.C. and Jarrold, M.F. (1998) Mobilities of carbon cluster ions: critical importance of the molecular attractive potential. *Chemical Physics*, **108**, 2416-2423.
  40. Wyttenbach, T., Helden, G.v., Batka Jr, J.J., Carlat, D. and Bowers, M.T. (1997) Effect of the long-range potential on ion mobility measurements. *Journal of the American Society for Mass Spectrometry*, **8**, 275-282.
  41. Clemmer, D.E. and Jarrold, M.F. (1997) Ion mobility measurements and their applications to clusters and biomolecules. *Journal of Mass Spectrometry*, **32**, 577-592.

42. Shvartsburg, A.A., Mashkevich, S.V., Baker, E.S. and Smith, R.D. (2007) Optimization of algorithms for ion mobility calculations. *The Journal of Physical Chemistry A*, **111**, 2002-2010.
43. Shvartsburg, A.A. and Smith, R.D. (2008) Optimum waveforms for differential ion mobility spectrometry (FAIMS). *Journal of the American Society for Mass Spectrometry*, **19**, 1286-1295.
44. Anderson, S.E., Bleiholder, C., Brocker, E.R., Stang, P.J. and Bowers, M.T. (2012) A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (III): application to supramolecular coordination-driven assemblies with complex shapes. *International Journal of Mass Spectrometry*, **330–332**, 78-84.
45. Bleiholder, C., Contreras, S., Do, T.D. and Bowers, M.T. (2012) A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (II). Model parameterization and definition of empirical shape factors for proteins. *International Journal of Mass Spectrometry*, **345-347**, 89-96.
46. Bleiholder, C., Wyttenbach, T. and Bowers, M.T. (2011) A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method. *International Journal of Mass Spectrometry*, **308**, 1-10.
47. Jurneczko, E. and Barran, P.E. (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst*, **136**, 20-28.

# 3

## **p53 DNA-binding domain** ***in vacuo***

*The tumour suppressor p53 is at the hub of a plethora of signalling pathways that maintain the integrity of human genome and regulate the cell cycle. It is therefore not surprising that this is mirrored in the intricate structure of p53. The central core of protein, consisting of the DNA-binding domain, is the most highly conserved domain and is required for sequence-specific DNA binding. Herein, mass spectrometry and ion mobility mass spectrometry studies on the conformational diversity of wild-type p53 DNA-binding domain are reported.*

### 3. p53 DNA-binding domain *in vacuo*

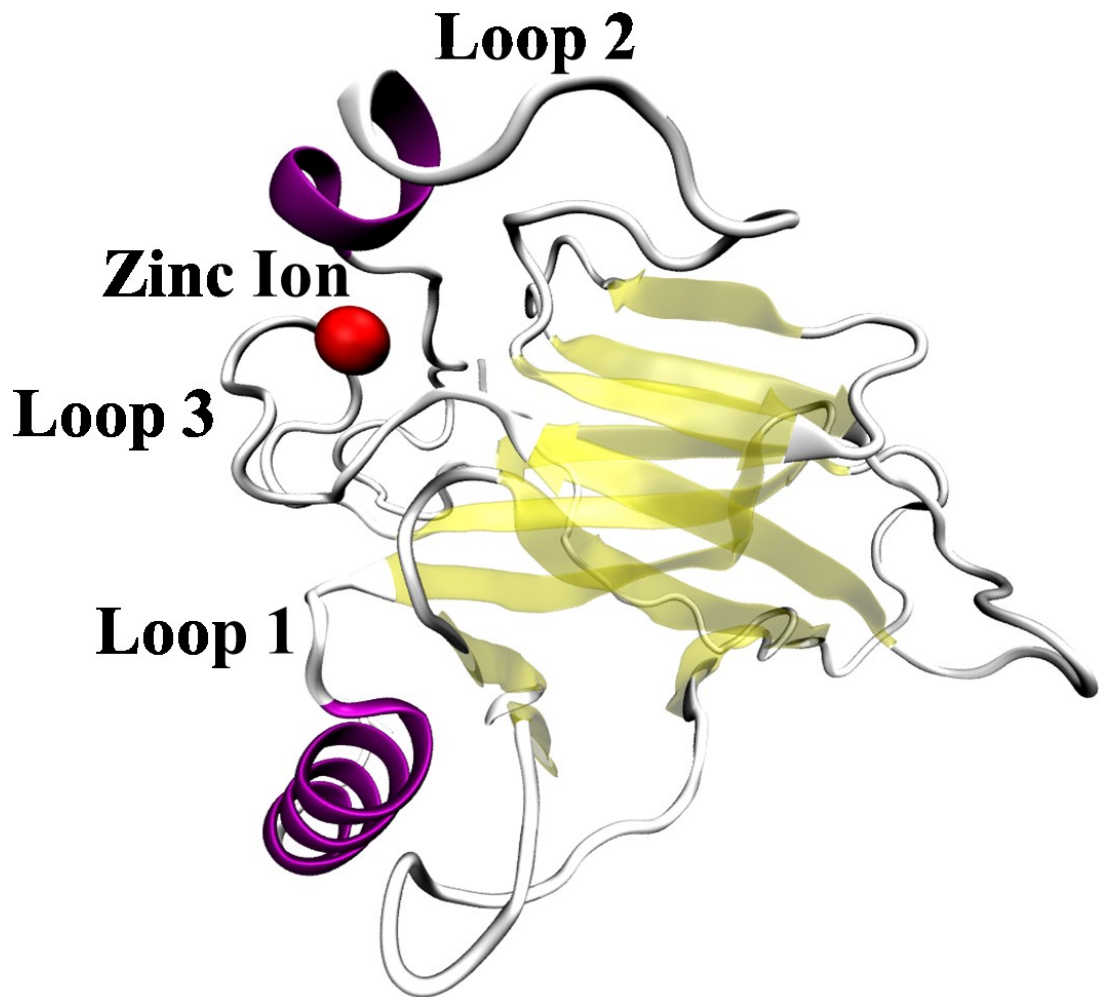
#### 3.1. Introduction

The central p53 domain is responsible for specific DNA-binding and it is probably the most interesting and important independent domain of the p53 tumour suppressor protein. This domain, also commonly known as ‘the core’ due to its protease resistance (1,2), contains the most evolutionary conserved sequences of the protein (3). Ever since the first structure of the p53 core domain in a complex with a 21-base-pair DNA was determined from high-resolution X-ray diffraction data in 1994 (4), a multitude of structures of the core domain of p53 have been elucidated: X-ray crystal and NMR solution structure of p53 core domain in DNA-free form (5,6), structures of p53 in a complex with DNA (7,8), several mutants (9-12), and p53 core bound to signalling proteins (13-15).

The structure of DNA-binding core domain (residues 94-292) consists of two anti-parallel  $\beta$ -sheets of four and five twisted strands, forming a ‘ $\beta$ -sandwich’ (Figure 3.1). This compact, barrel-like structure provides the basic platform for a loop-sheet-helix motif which docks to the DNA major groove and for two large loops, L2 (residues 163-195) and L3 (residues 236-251). The architecture of the L2/L3 region is stabilised by a zinc ion, which is tetrahedrally coordinated by a histidine and three cysteines residues (Cys176, His179, Cys238 and Cys242) (4-6). Mutations of these corresponding residues in mouse p53 inhibit DNA binding (16). The presence of single zinc ion molecule was found to be essential for retaining the WT conformation of p53. Hainaut *et al.* (17) and others (18,19) established that the use of copper and cadmium instead of zinc ion removes DNA-binding.

Human p53 core domain is only marginally stable and rapidly unfolds just above body temperature with a half-life of 9 minutes (20-22). Evidence suggest that the low thermodynamic and kinetic stability of p53 may be a result of an adaptive evolutionary process (5,23,24), to allow for a rapid inter conversion between folded

and unfolded states providing an additional layer of control of cellular protein levels.



**Figure 3.1:** The DNA-binding core domain of p53 (residues 94-312) showing a predominantly sheet-like structure (yellow), alpha helices (purple), loop regions (white) and coordinated zinc ion is shown in red. PDB 2FEJ file (5) was used to generate image in VMD software.

Furthermore, the NMR structure of p53 deposited by Wang *et al.* (6) reveals a number of buried polar groups that could potentially be key elements contributing to p53's inherent instability. Increasing the stability of the p53 protein has been deemed necessary for some biophysical and structural investigations, in which a stabilised quadruple mutant has been used (11,25).

The core domain of p53 is also the main target for mutations as over 80% of the missense tumour-derived mutations are mapped to this region (26).

## 3.2. How p53 binds DNA

p53 is known to regulate the gene transcription by binding to over 100 different, naturally occurring DNA-binding sites or response elements (27). The emerging picture is that p53 is involved in these promiscuous interactions with genomic DNA *via* both its DNA-binding domain and C-terminal regulatory domain. This chapter will solely focus on what is known about the involvement and regulation of the p53 core domain in sequence-specific DNA binding. Isolated p53-DBD binds specifically to double-stranded DNA sequence of two tandem decameric elements PuPuPuC(A/T)(A/T)GPyPyPy (Pu = A/G, Py = T/ C) that can be separated by 0 -13 nucleotide base pairs (28). Four core p53-DBDs bind to these response elements in a highly cooperative manner to give a complex with 4:1 stoichiometry, even in the absence of the tetramerization domain responsible for oligomerization of the full-length protein. Rippin *et al.* (29) report this phenomenon to be partially mediated by core-core interactions. The binding can also induce bending and twisting of the DNA (30,31).

X-Ray crystallographic measurements provide detailed information on the structural architecture of p53 tetramers bound to DNA (7). Two core domains bind to a half-site DNA, forming a symmetrical dimer (e.g. the green and grey molecules shown in Figure 3.2). The core-core interface is formed by residues located in the L3 region and a helical fragment within L2, which includes Pro177, His178, Arg181, Met243, and Gly244, in accord with NMR results (29,32). Both hydrophobic and water mediated polar contacts are an integral part of the protein-DNA interface. p53 molecules self-assemble on DNA half-sites to form a ‘dimer of dimers’, stabilized by core-core and base-pair stacking interactions (7). The nature of the core-core interface along the DNA fluctuates among the different DNA complexes, reflecting the transitory mode of this interface region and p53’s capability to bind to response elements with different spacer lengths between the two half-sites (33). For instance a 10-base pair spacer would place p53-DBD dimers on the same face of the DNA-double helix but a significant bending of DNA would be required for inter-dimer



interactions to occur, whereas a 5-base pair spacer would place p53 dimers on opposite faces of the DNA (34).



**Figure 3.2:** Image of human p53-DNA binding domains (residues 94-293) bound to two palindromic DNA half-sites. Representation of PDB 2AC0 file (7) rendered orthographically using VMD software.

The DNA major-groove contacts are facilitated by residues: Ala276, Cys277 and Arg280 from the carboxy-terminal helix and also Lys120 from the L1 region (7).

The vast amount of research probing the functions of the p53-DBD has provided many insights into possible modes of regulation of DNA binding. Multiple protein-protein interactions have also been mapped to this domain and may directly or indirectly affect the ability to interact sequence-specifically with DNA (35).

### 3.3. Methodology

#### 3.3.1. Protein Expression and Purification

Wild-type (WT) human p53-DBD (amino acid residues 94-312) was expressed and purified by Dr Penka V. Nikolova (King's College London, UK) and Dr David J. Clarke (University of Edinburgh, UK) according to the method outlined below. Cloning of the protein was performed using pRSETA expression vector (Invitrogen) which was modified so that it did not contain the His tag (20,36,37). The protein was expressed at 37°C in *Escherichia coli* C41 cells and grown up to an  $OD_{600}$  of 0.8. The temperature was reduced to 22°C before cell cultures were induced with 1 mM isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG) and grown overnight. Cells were harvested by centrifugation and lysed using Bug Buster Protein Extraction Reagent. Benzonase nuclease and EDTA-free protease inhibitor tablets were also used during this process. The soluble fraction was loaded on to an SP-Sepharose column eluted with NaCl gradient followed by gel-filtration chromatography using an HP26/60 Superdex 200 column.

#### 3.3.2. Sample Preparation

WT p53 DNA-binding domain protein samples were flash frozen in liquid nitrogen and stored in a -80°C freezer. Prior to mass spectrometry and ion mobility mass spectrometry measurements, protein samples were thawed and dialysed in 50 mM ammonium acetate for 2 h at 4°C using slide-A-lyzer dialysis cassettes (Thermo Scientific, Inc). Protein concentrations were measured spectrophotometrically (NanoDrop Spectrophotometer ND 1000 Thermo Scientific, USA) by using an extinction coefficient of  $\epsilon_{280} = 17,130 \text{ cm}^{-1} \text{ M}^{-1}$  calculated by the method of Gill and von Hippel (38). WT human p53-DNA binding domain amino acid sequence is shown below:

SSSVPSQKTY	QGSYGFR LGF	LHSGTAKSVT	CTYSPALNKM	FCQLAKTCPV
QLWVDSTPPP	GTRVRAMAIY	KQSQHMTEVV	RRCPHHERCS	DSDGLAPPQH
LIRVEGNLRV	EYLDDRNTFR	HSVVVPYEPP	EVGSDCTTIH	YNYMCNSSCM
GGMNRRPILT	IITLEDSSGN	LLGRNSFEVR	VCACPGRDRR	TEENLRKKKG
EPHHELPPGS	TKRALPNNT			

### **3.3.2.1. DNA Oligonucleotides**

Double-stranded DNA fragments used in this study were constructed by annealing purified synthetic oligonucleotides purchased from ATDBio (University of Southampton, UK). DNA oligomers used in this study contained the following palindromic sequences: GGAACATGTTCC and GAACATGTTCTGAACATGTTCC consensus binding site. Annealing was performed by heating the oligonucleotides at 95°C for 5 minutes and gradual, slow cooling to room temperature to ensure an efficient hybridization. The final concentration of the double-stranded DNA used for the binding experiments is based on the assumption that all of DNA had annealed completely, however trace amounts of unbound DNA/hairpin structures might be still present.

### **3.3.3. Mass Spectrometry and Ion Mobility Mass Spectrometry**

The travelling wave based ion mobility mass spectrometry (TW IM-MS) experiments were performed on a Synapt HDMS system (39) (Waters Corporation, Manchester, UK) described in detail in Chapter 2, Section 2.5.1. Samples were introduced into the source region by nano-electrospray ionisation (n-ESI). The instrument was operated in a positive ionisation mode with a capillary voltage of 0.8 -1 kV, cone voltage of 65 V and source temperature of 30 °C. Source backing pressure was elevated to 4 mBar for improved transmission of ions. Travelling wave heights of 13, 14 and 15 V and 250 m/s were employed. The ion mobility separator containing nitrogen gas was operated at a pressure of 1 mBar. During data acquisition, the parameters of the instrument were optimised to obtain the optimum ion mobility separation. Equine heart cytochrome c and myoglobin was analysed under denaturing conditions to create a calibration curve for cross-section measurements versus values obtained from drift tube ion mobility mass spectrometry (DT IM-MS) studies. Calibration procedure is described in Chapter 2, Section 2.5.2.

Drift tube ion mobility mass spectrometry (DT IM-MS) experiments were performed on an in-house modified quadrupole time-of-flight mass spectrometer, the MoQToF

(40). More detailed description of this instrument can be found in Chapter 2, Section 2.4.1. Ions were produced by positive n-ESI using Z-spray source, within a spray voltage range 1.0 to 1.8 kV and a source temperature of 80°C. The drift cell was filled with helium buffer gas at the average pressure of 3.5 - 4.0 Torr (4.7 - 5.3 mBar) and a temperature between 298 - 302 K. The electric potential across the cell was varied from 60 to 15 V and ATDs were recorded at eight drift voltages. The mobility of the ion of interest was calculated from a linear plot of average arrival time versus pressure/drift voltage ( $P/V$ ) and from this the rotationally-averaged collision cross section (CCS) for each resolvable species at a given charge state were obtained.

### **3.3.4. Theoretical Measurements of collision cross sections**

Using the input coordinate files from published X-ray crystallography measurements of p53 deposited by Wang *et al.* (6) (PDB: 2OCJ) the rationally-averaged CCS p53-DBD monomer, dimer and extended linear form of monomer were calculated. Projection Approximation (PA) (41), Exact Hard Sphere Scattering (EHSS) (42) and Trajectory Method (TM) (43) were utilised, all implemented in MOBCAL. The TM provides the most accurate estimate of the CCS and has been used here for a comparison with the experimental values of CCS. We have not employed it for the fully extended form of p53-DBD since the size of this molecule was too large for our compilation of the MOBCAL software. The less accurate EHSS method gives qualitatively similar results (44) and so the two values are comparable.

Using methods established by Jarrold and co-workers (42,43), CCS from coordinates obtained from the X-ray crystallography measurements of p53-DBD dimer cross-linked to 16-base-pair DNA fragment (deposited by Ho *et al.* (6)) were also calculated.

The MOBCAL calculations were performed by Dr Massimiliano Porrini.

### 3.3.5. Production of Waterfall Plots for data obtained on TW IM-MS

The waterfall plots for WT p53-DBD were produced in Origin 8.5.1 software (OriginLab, Northampton, MA, USA). Multi-curve waterfall charts were created by plotting estimated CCS values ( $x$ ), calculated as described in Chapter 2, Section 2.5.2, against peak intensity values ( $y$ ) and charge ( $z$ ). The intensities of the peaks were obtained from the Mass Lynx software (version 4.1, Waters, Manchester, UK).

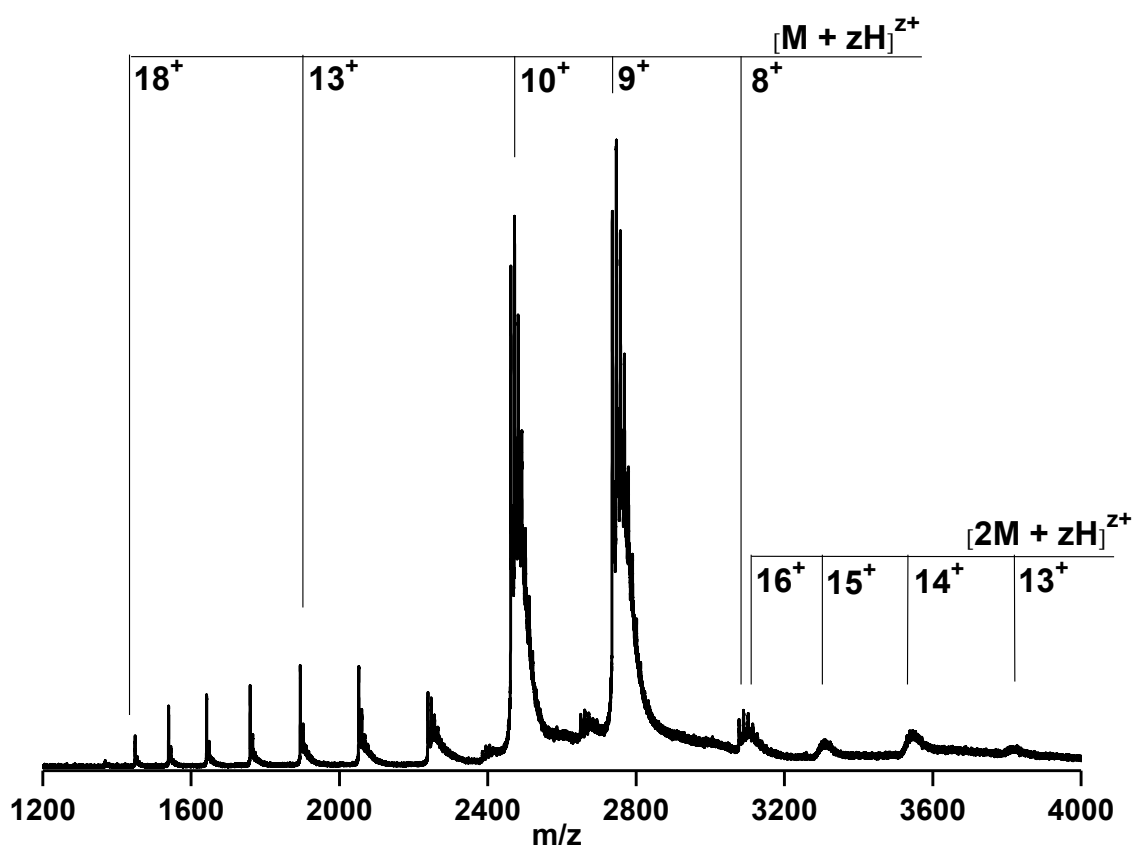
## 3.4. Insights into Structure(s) and Dynamics of WT p53-DBD by MS and IM-MS

This chapter focuses on the isolated core DNA-binding domain of p53, as a monomer, as well as in partnership with DNA. IM-MS is used to probe the conformational landscapes of this protein. The experimental approach provides the general shape and population of conformations via the measurement of their  $m/z$  and collision cross sections. The conformational spread of p53-DBD under buffered and denaturing solution conditions is quantified. The extent of changes in conformational families associated with binding of the isolated p53-DBD to two different DNA constructs is also monitored.

### 3.4.1. MS of WT p53-DBD in buffered solution conditions

From the buffered solution, a charge state distribution of wild-type (WT) p53 DBD monomer is observed with  $8 \leq z \leq 18$  and a few additional multimeric species (Figure 3.3). Each of the monomer charge states is attributable to the zinc bound form of the protein (calculated  $M_R = 24615.5$  Da, observed  $M_R = 24612.4$  Da) which suggests that all conformations present in solution retain the zinc binding site, and that it is not perturbed by the n-ESI desolvation process or by any subsequent coulombically driven unfolding. The most abundant species in the mass spectrum correspond to monomeric  $[M+9H]^{9+}$  and  $[M+10H]^{10+}$  indicative of a compact form

of the protein with a limited number of charge-accessible residues available for protonation. Peaks also display phosphate adducts at low charge density. Two unique  $m/z$  are present for a dimeric form of p53,  $[2M+zH]^{z+}$  where  $z = 13$  and  $z = 15$ , representing higher order multimeric association in solution, again zinc is bound stoichiometrically. Higher charged species, where  $11 \leq z \leq 18$  are observed at low abundance, evidence of a small population of more unfolded states of p53. Even under ‘near native’ conditions p53 displays a wide charge state distribution ( $8 \leq z \leq 18$ ), which is a signature behaviour of IDPs *in vacuo* as reported by mass spectrometry (45).



**Figure 3.3:** n-ESI spectrum of 95  $\mu$ M p53 DBD obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

### 3.4.2. MS of WT p53-DBD in denaturing solution conditions

At low pH (Figure 3.4) the observed spectrum displays a wider charge state envelope over the acquired mass range (400 - 4000  $m/z$ ). Under these conditions the most dominant species of p53 observed is a monomer. The deconvoluted mass ( $M_R = 24547.9$  Da) is less than the mass of the 'native' p53 monomer ( $M_R = 24612.4$  Da). This is due to the loss of the tetrahedrally coordinated zinc atom ( $M_R = 65.4$  Da) which destabilises the structure of protein. The observed charge states ranged from  $9 \leq z \leq 34$  attributable to facile protonation of solvent exposed basic residues. These sites are proposed to be basic residues - there are 19 arginine, 8 lysine and 9 histidine residues in the p53 molecule, as well as at the N-terminus where proton addition could potentially occur.

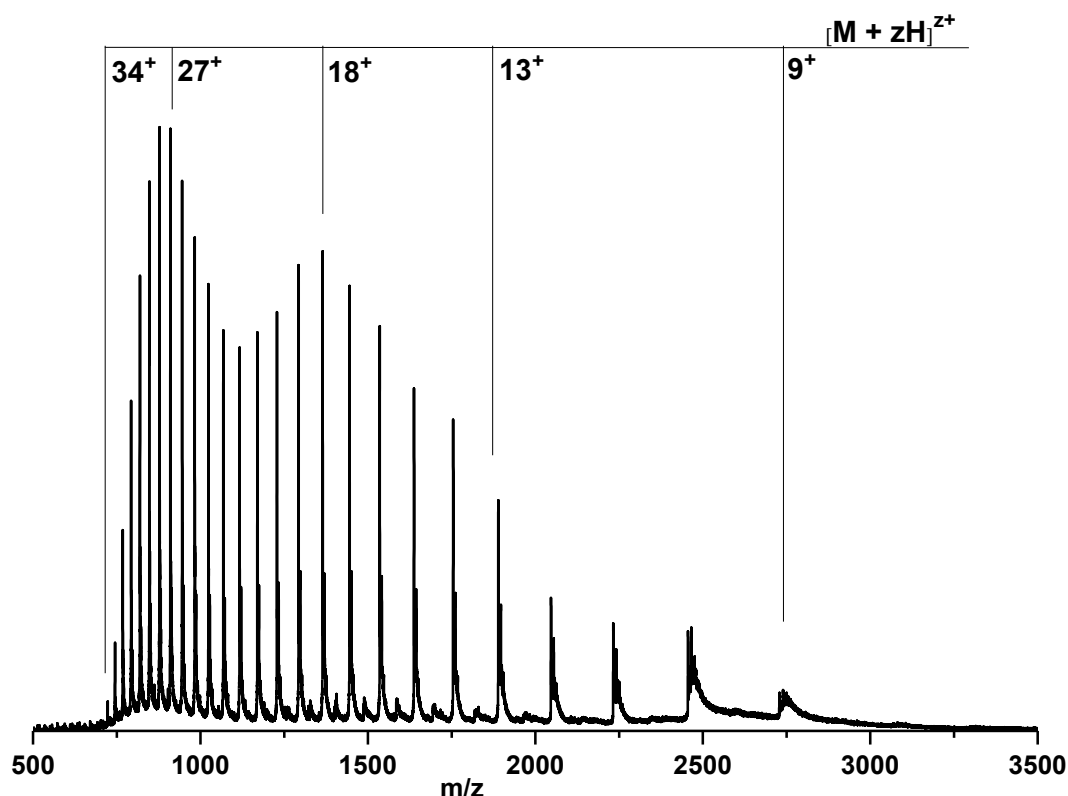


Figure 3.4: n-ESI spectrum of 95  $\mu$ M p53 DBD obtained via the TW IM-MS from pH 1.5 solution (50 mM ammonium acetate, containing 10% by volume propan-2-ol and 1% formic acid).

The spectrum displays a trimodal charge distribution with centres at  $z = 27$ , 18 and 11, suggesting at least three conformational families are present in the solution and transferred into the gas phase. The distribution ranging from  $22 \leq z \leq 34$  is more populated than the one ranging from  $12 \leq z \leq 22$ , and so it can be inferred that the denatured protein favours more extended conformations. In summary, as Figures 3.2 and 3.1 are compared, an addition of acid (lower pH) causes a shift in the most intense peaks to higher charge (lower  $m/z$ ), and a decrease in intensity of lower charged (higher  $m/z$ ) as the folded conformation is lost and the unfolded state is predominantly populated interactions.

### 3.4.3. MS of WT p53-DBD in high buffer solution conditions

Under higher buffer conditions (500 mM ammonium acetate) the most dominant species of p53 detected is zinc bound monomer (Figure 3.5).

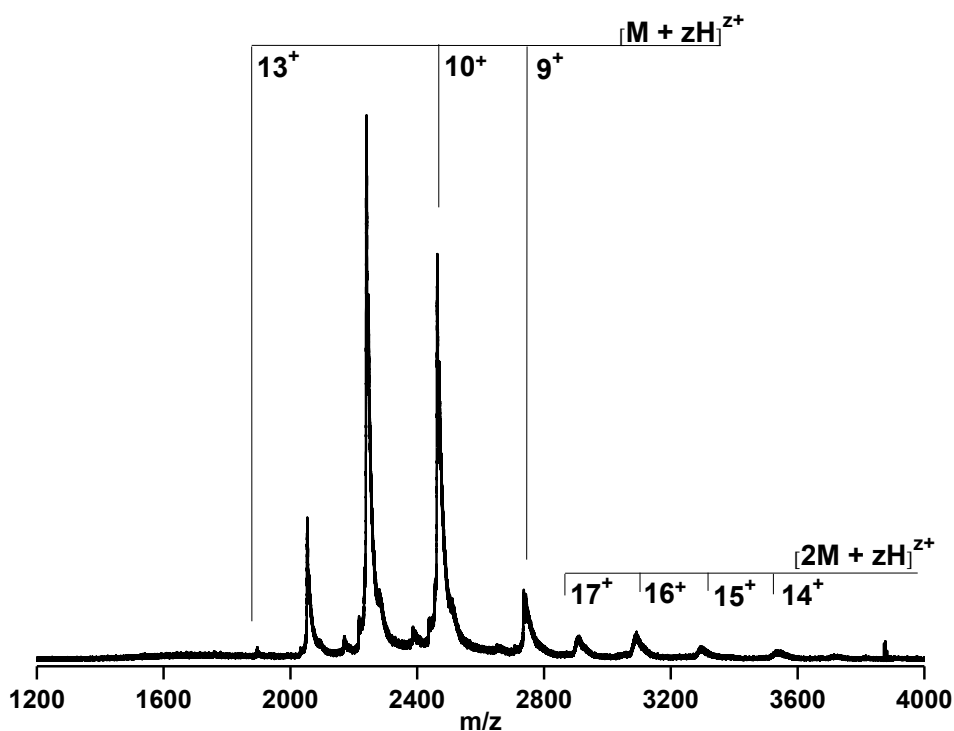


Figure 3.5: n-ESI spectrum of 95  $\mu$ M p53 DBD obtained via the DT IM-MS from a solution at high buffer concentration (500 mM ammonium acetate, containing 10% by volume propan-2-ol).



Nonetheless the observed mass spectrum displays a much tighter charge state envelope for the monomeric ensemble over the acquired mass range. There is a clear depopulation in the higher charge states of p53 monomer. The highest charge state corresponds to  $z = 13$  in contrast to  $z = 18$  observed in the spectrum sprayed from the lower buffer solution conditions (Figure 3.3). This is not surprising as the high concentration of buffer tends to stabilize the protein's structure driving p53 molecule to a more compact conformational isomer (with fewer charges). Additionally dimer population  $[2M+zH]^{z+}$  (again zinc is bound stoichiometrically) with the distribution ranging from  $14 \leq z \leq 17$  is more intense in higher buffer solution conditions. A new unique  $m/z$  peak for a dimeric form of p53 is detected, where  $z = 17$ .

#### **3.4.4. Conformational Landscapes occupied by WT p53-DBD**

Along with mass spectrometry, ion mobility measurements have been performed on the p53 DNA-binding domain sprayed from two different solutions. Collision cross-sections obtained are tabulated in Table 3.2, 3.3 and 3.4 and plotted as a function of charge state in Figure 3.6 and 3.7. The data taken at pH 1.5 (Figure 3.6) shows a smooth, near linear increase in cross-section with respect to charge, which can be explained by the range of dynamic states of the protein in solution, resulting in differently populated protonated states, which have little structure, and are easily subject to coulombically driven unfolding once in the gas phase. Faull *et al.* (46) have also used IM-MS to investigate the conformations of p53 in the presence and absence of zinc, showing that removal of this zinc ion disrupts the structure of the DNA-binding domain.

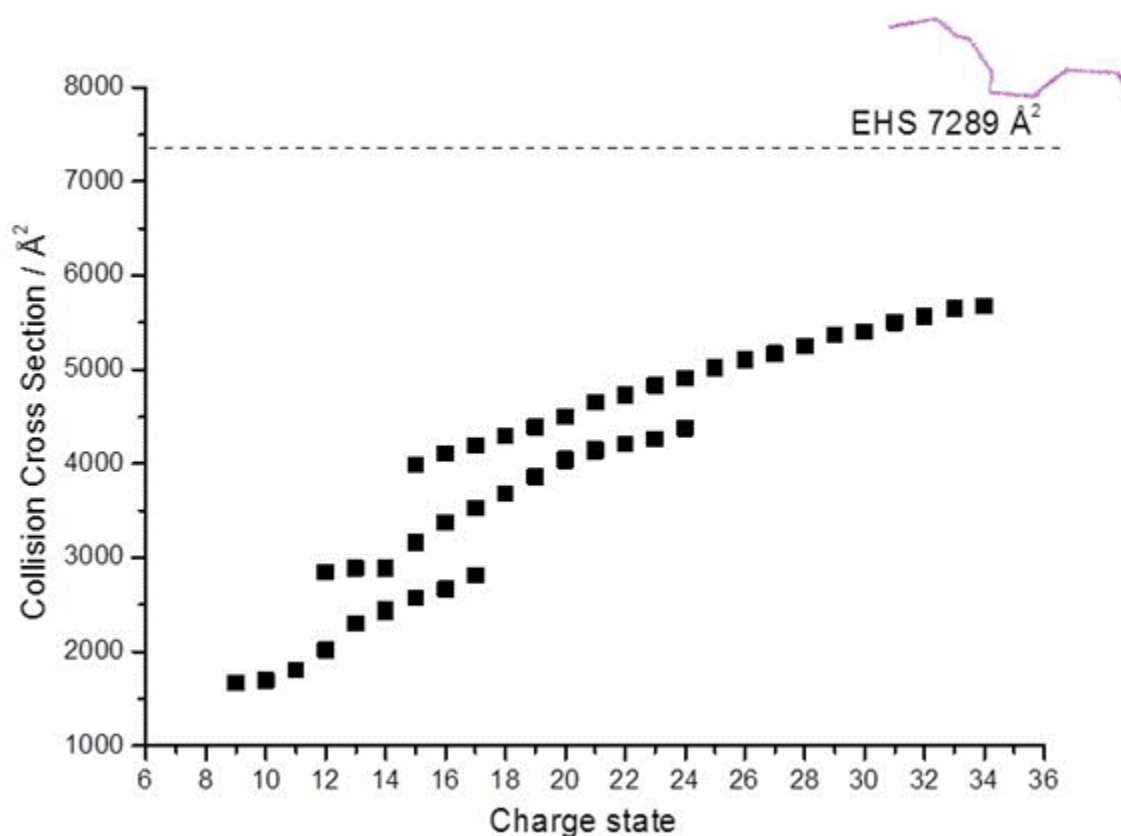


Figure 3.6: Collision cross sections versus charge obtained via the TW IM-MS of 95  $\mu\text{M}$  p53 DBD at low pH (50 mM ammonium acetate, containing 10% by volume propan-2-ol and 1% formic acid). The graph also shows the calculated CCS value for coordinates expected from a fully extended form of the p53 DBD. This modelled structure was built using *xleap* within the AMBER molecular modelling package (47) and was found to be 7289  $\text{\AA}^2$ .

Here in the absence of solvent at data taken at pH 1.5, each additional proton added to the flexible protein can cause local unfolding due to unfavorable interactions with other proximal proton(s). However it also displays some regions of stability seen as ‘shelving in the cross section values ( $z = 8$  to  $10$  and  $z = 12$  to  $14$ ). For each of the three lowest charge states, ( $z = 9, 10, 11$ ) a single conformation was resolved with CCS of 1671  $\text{\AA}^2$ , 1697  $\text{\AA}^2$  and 1805  $\text{\AA}^2$  respectively. At these low charge states, the structure is most likely compact but loss of the zinc atom indicates that there is much less secondary structure than the conformers reported (see below) from solutions buffered at pH = 6.8. For  $12 \leq z \leq 24$  multiple unfoldomers for each charge state were observed.

Using methodologies established by Jarrold and co-workers (42,43) CCS were calculated from coordinates obtained from the X-ray crystal structure of p53 deposited by Wang *et al.* (6) for p53-DBD monomer, dimer and extended monomer. The results are tabulated in Table 3.1.

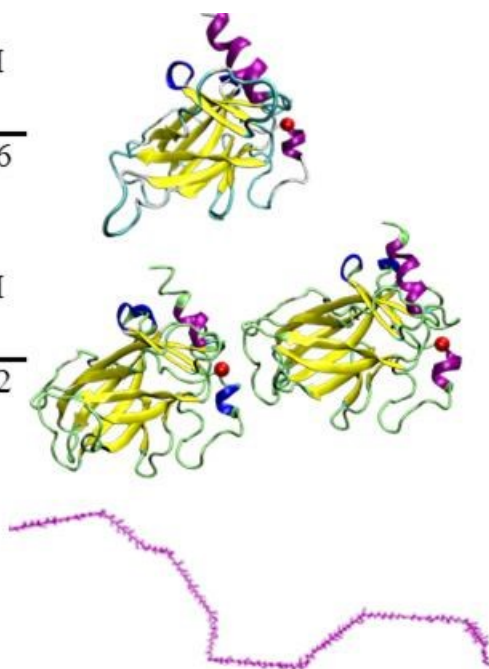
(a)	Method	PA	EHSS	TM
	CCS	2996	3861	3886

(b)	Method	PA	EHSS	TM
	CCS	1657	2101	2082

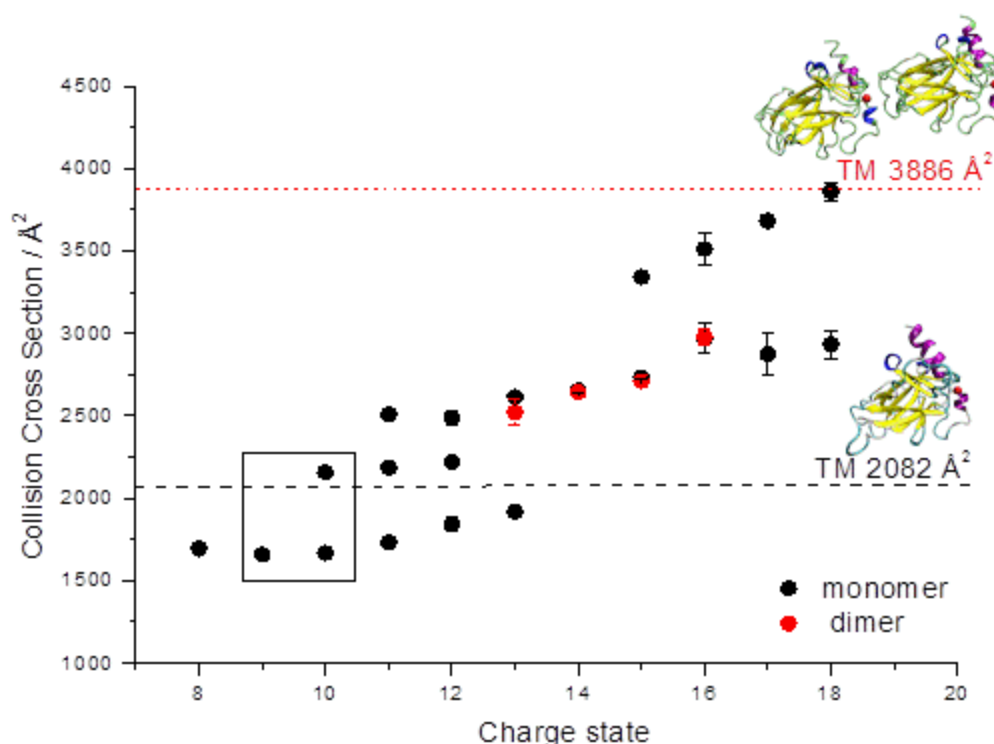
(c)	Method	PA	EHSS
	CCS	6304	7289



**Table 3.1:** Theoretical calculations of collision cross sections for (a) p53-DBD monomer (b) p53-DBD dimer and (c) p53 DBD extended monomer. Here values are shown using the Projection Approximation, Exact Hard Sphere Scattering and Trajectory Method. TM was not used here for the fully extended form of p53 since the size of this molecule was too large for the compilation of the MOBCAL software. The less accurate Exact Hard Sphere Scattering (EHSS) method gives qualitatively similar results (44) and so the two values are comparable.

The calculated CCS for coordinates expected from a fully extended form of the p53 DBD was built using *xleap* within the AMBER molecular modelling package (47) and found to be 7289 Å<sup>2</sup> (highlighted in Figure 3.6). Even the highest charge state experimentally observable has a significantly smaller CCS (by 22%), suggesting that the DBD of p53 retains some secondary structure, or at least intermolecular interactions, all be they transient.

From buffered conditions (Figure 3.7), there are two dominant charge states ( $z = 9$  and 10). The  $z = 9$  species provides a single resolvable conformer ( $1657 \text{ \AA}^2$ ), indicative of a highly compact form of the protein, whereas for  $z = 10$  two species are resolved, with the most populated being ( $1669 \text{ \AA}^2$ ) and a lower intensity extended form ( $2157 \text{ \AA}^2$ ). These conformers are not baseline resolved which could be due to the resolution of the Synapt HDMS instrument, or alternatively that they are in conformational equilibrium on the timescale of these IM-MS experiments (5 - 10 ms).



**Figure 3.7:** Collision cross sections versus charge obtained via the TW IM-MS of  $95 \mu\text{M}$  p53 DBD from buffered solution conditions (50 mM ammonium acetate, containing 10% by volume propan-2-ol).

The low CCS for  $z = 8-10$  are smaller than those calculated from the coordinates obtained from the crystal structure deposited by Wang *et al.* (6), here using the Trajectory Method (43) thus indicating that the protein has collapsed somewhat in comparison to the X-ray structure. This follows the general trend for monomeric proteins previously reported (44), and here reveals conformers which are smaller than the equivalent charge state species obtained from the low pH solution,

suggesting that the bound zinc has a stabilizing effect. The second much lower populated conformational family which starts with the more extended conformer for  $z = 10$  does coincide with the value calculated from crystal structure coordinates. The dimers that are present following n-ESI have significantly lower (by 23%) CCS than that obtained from the crystal structure; this form is clearly not stable in the gas phase under our experimental conditions.

For monomeric p53 DBD, with increasing charge, a pathway of unfolding is clearly visible, the CCS increase. However there is far more shelving than found at pH 1.5 suggesting a more concerted unfolding and again more stability to the protein. The calculated CCS for each charge state are higher than that the equivalent charge state conformers obtained from the physiological pH solution. The data in Figure 3.7 provide evidence for several stable states on the unfolding pathway. With increasing charge ( $z = 10$  to 13 and  $z = 15$  to 18) at least two conformers can be observed. Although p53 has unfolded to a more extended structure it still retains some non-covalent interactions that sustain the protein in these different conformations for the duration of the mobility experiment.

Arrival Time Distributions (ATD) for each charge state, showing the abundance of the conformation within the protein under buffered conditions is presented in Figure 3.8. For  $11 \leq z \leq 18$ , multiple unfoldomers (Un) are observed in each charge state, and the number of resolvable Un conformers is at least four, albeit at low intensity (Figure 3.9). The lack of baseline resolution in the arrival-time distributions suggests interconverting conformers for these extended forms over the timescale of our experiment. The CCS value of the second, much less populated family of conformers, which starts with the more extended conformer (X1) for  $z = 10$  (2157 Å<sup>2</sup>) coincides with the CCS value calculated from crystal structure coordinates (2072 Å<sup>2</sup>), but more extended forms are also present, suggesting unfoldomers (Un).

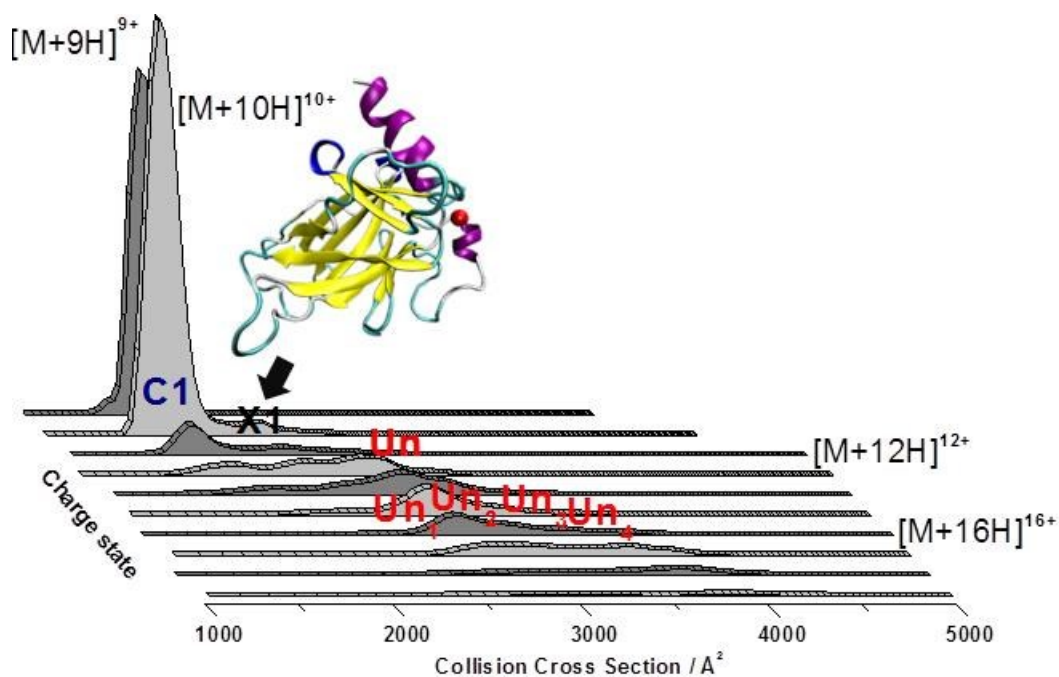


Figure 3.8: Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of p53 DBD sprayed from buffered solution conditions (95  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via the TW IM-MS. The x, y and z axis show the collision cross section (in  $\text{\AA}^2$ ), charge state (range  $9 \leq z \leq 18$ ), and the relative intensity, respectively.

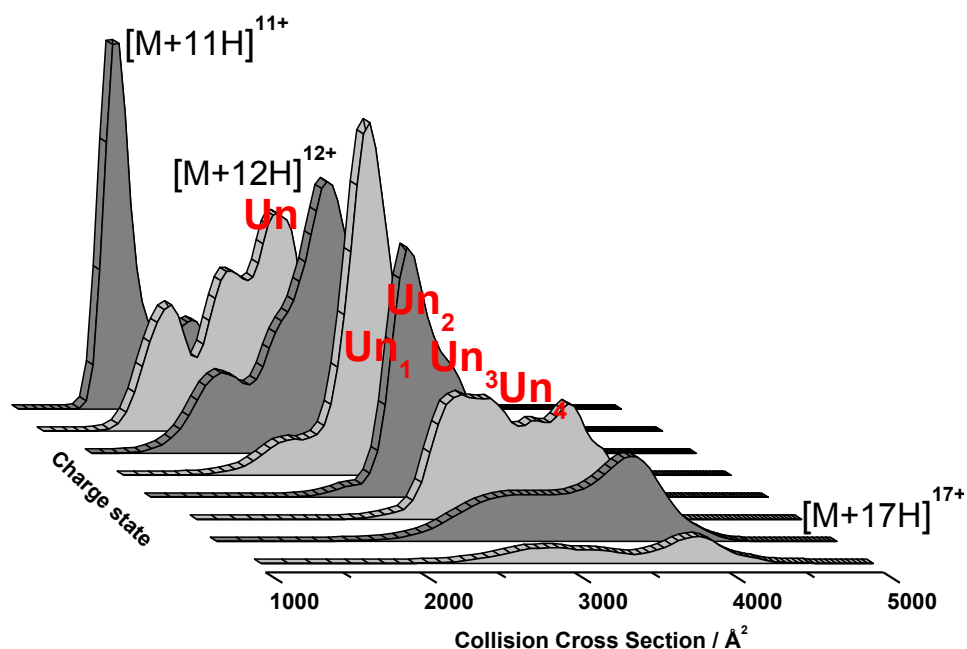


Figure 3.9: Enlargement of the ATDs for charge state range between  $11 \leq z \leq 18$ . Here a range of extended conformers for higher charge states is observed.

The ATDs for the wild type p53-DBD were generated for the non-adducted species. In the mass spectrum of the p53-DBD in buffered solution conditions (Figure 3.3) low charged populations display multiple phosphate adducts, in particular  $[M+9H]^{9+}$  and  $[M+10H]^{10+}$  are the heavily adducted species. The adducted populations could potentially show different conformational properties than the ‘naked’ protein. The comparison of ATDs for a non-phosphorylated protein and a p53-DBD with two phosphate adducts is presented in Figure 3.10. The conformational phenotypes appear to be comparable and the difference in CCS in only a marginal one, in a range of 1 - 2 %. Nonetheless, for the provided ATDs (p53-DBD and mutants presented in the subsequent chapter) a mass window corresponding to a ‘naked’ protein (non-adducted) was used to generate ion mobility profiles.

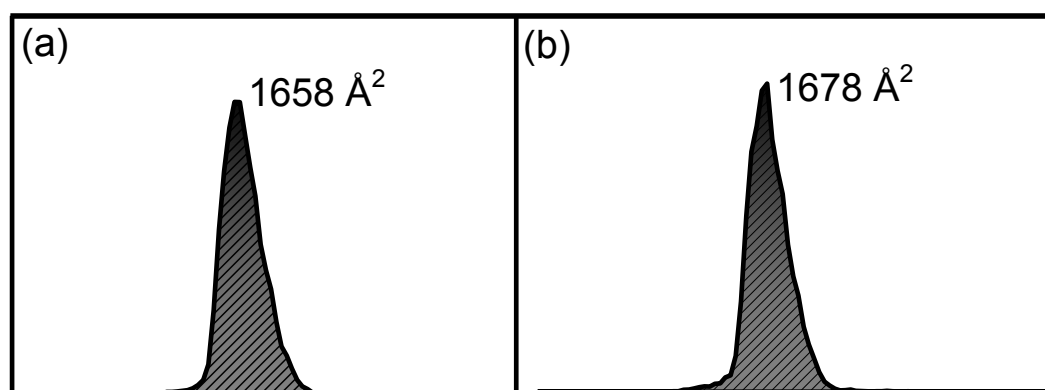


Figure 3.10: Comparison of ATDs for  $[M+9H]^{9+}$  species (a) non-phosphorylated (b) and doubly phosphorylated. Data were acquired via the TW IM-MS at 13 V IMS wave height from buffered solution conditions (95  $\mu$ M p53-DBD solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol),

#### 3.4.4.1. Collision Cross Sections of WT p53-DBD

Charge state	Collision Cross Section / Å <sup>2</sup>		
9	1671±18,		
10	1697±13,		
11	1804±24,		
12	2022±52,	2847±1	
13	2296±70,	2888±16	
14	2438±91,	2892±54	
15	2570±61,	3163±32,	3991±0
16	2668±0,	3374±11,	4110±24

Charge state	Collision Cross Section / $\text{\AA}^2$		
17	2810±0,	3524±27,	4196±22
18		3680±34,	4296±20
19		3860±73,	4389±44
20		4035±96,	4500±67
21		4143±89,	4649±2
22		4213±66,	4732±18
23		4260±80,	4834±18
24		4373±27,	4905±10
25			5020±27
26			5101±28
27			5170±23
28			5252±15
29			5366±6
30			5398±24
31			5499±35
32			5568±31
33			5654±43
34			5671±9

Table 3.2: Collision cross sections of p53 DBD obtained via the TW IM-MS from pH 1.5 solution (95  $\mu\text{M}$  solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol and 1% formic acid). Values are an average of three replicates and stated with the standard deviation of the mean.

Charge state	Collision Cross Section / Å <sup>2</sup>			
8	1695±14,			
9	1659±1,			
10	1669±17,	2157±28		
11	1732±21,	2187±22,	2508±26	
12	1842±42,	2220±9,	2491±42	
13	1918±26,		2614±23	
14			2655±12,	
15			2733±7,	3341±12
16			2968±93,	3511±100
17			2875±123,	3681±37
18			2933±83,	3861±56

Table 3.3: Collision cross sections of monomeric p53 DBD obtained via the TW IM-MS from pH 6.8 solution (95  $\mu\text{M}$  solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol). Values are an average of three replicates and stated with the standard deviation of the mean.



Charge state	Collision Cross Section / $\text{\AA}^2$
13	2520 $\pm$ 79
14	2644 $\pm$ 24
15	2712 $\pm$ 22
16	2979 $\pm$ 53

Table 3.4: Collision cross sections of p53 DBD dimer obtained via the TW IM-MS from pH 6.8 solution (95  $\mu\text{M}$  solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol). Values are an average of three replicates and stated with the standard deviation of the mean.

### 3.5. Interactions of WT p53-DBD with its consensus DNA binding site

The following section will focus on the use IM-MS to probe the involvement and regulation of the p53-DBD in sequence-specific DNA binding. In order to do this, a 12-mer double-stranded DNA (5'GGAACATGTTCC CCTTGTACAAGG'3 containing two consensus binding sites  $M_R = 7290.8$  Da) and a 20-mer ds DNA (5'GAACATGTTCTGAACATGTTCTTGTGTACAAGCTTGTACAAG'3 containing four consensus binding sites  $M_R = 1223.2$  Da) was used to examine the binding capacity of the WT p53.

#### 3.5.1. WT p53-DBD dimer bound to ds 12 mer DNA

As seen in previous section, in the absence of DNA p53-DBD exists mainly as a free monomer and is also capable of weak monomer-monomer interactions. In the presence of DNA, p53 possesses a remarkable self-organizing ability (mass spectrum presented in Figure 3.11). The WT p53-DBD consistently binds consensus double-stranded 12 mer DNA as a 2:1 complex only ( $z = 12 - 15$ ).

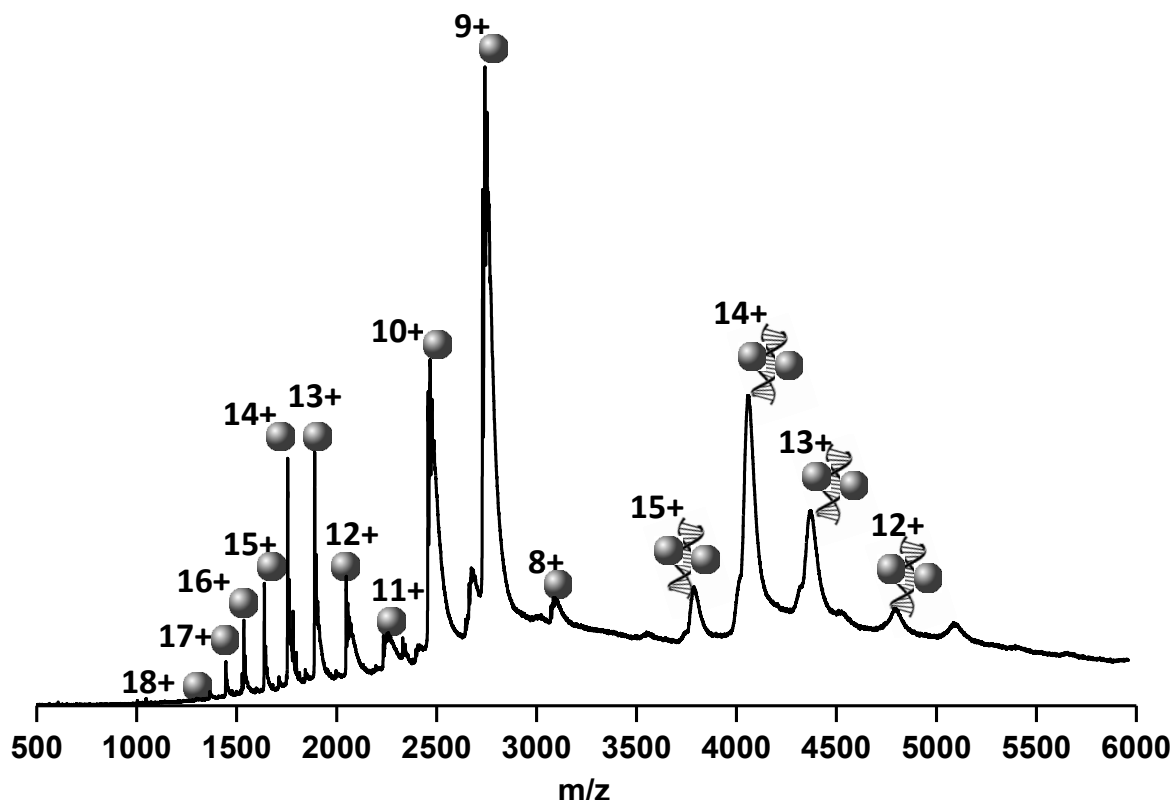


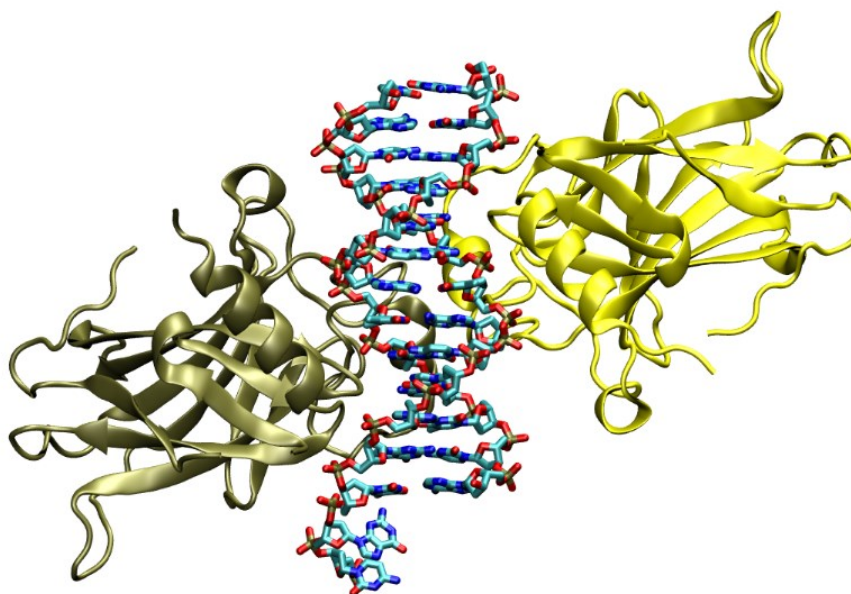
Figure 3.11: n-ESI spectrum of 20  $\mu\text{M}$  p53 DBD with a 10  $\mu\text{M}$  double-stranded 12-base-pair DNA fragment in a ratio 2:1 (protein to DNA, based on a protein monomer). Mass spectrum was obtained from a solution buffered with ammonium acetate and recorded via the DT IM-MS. The single spheres on the left represent a p53 monomer, whereas the spheres with DNA spirals correspond to the dimeric p53-DNA complex.

In mobility space, the 2:1 p53-DBD/DNA complex exists only as a single conformation, with CCS values only minimally higher (mass corrected) compared to the dimeric form without DNA (CCS values are tabulated in Table 3.5). For instance, for the  $z = 14$  complex, a CCS of  $3194 \text{ \AA}^2$  was measured, and for the  $z = 14$  dimer, a CCS of  $2644 \text{ \AA}^2$  was measured), thus suggesting a tight association of this complex.

Charge state	Collision Cross Section p53 dimer / $\text{\AA}^2$	Collision Cross Section p53 - DNA complex / $\text{\AA}^2$
13	2520	3113
14	2644	3194
15	2712	3191

**Table 3.5:** Collision cross sections of the dimeric p53-DNA complex obtained via the DT IM-MS from buffered solution conditions with ammonium acetate. (20  $\mu\text{M}$  p53 DBD with a 10  $\mu\text{M}$  double-stranded 12-base-pair DNA fragment in a ratio 2:1; protein to DNA, based on a protein monomer).

CCS values measured across the observable charge states ( $z = 13 - 15$ ) for the complex are unchanging suggestive of a stable structure in the gas phase. Using methods established by Jarrold and co-workers (42,43), the CCS from coordinates obtained from the X-ray crystal structure of p53-DBD dimer cross-linked to 16-base-pair DNA fragment (Figure 3.12) were calculated (deposited by Ho *et al.* (8)).



**Figure 3.12:** X-Ray crystal structure of p53-DBD dimer cross-linked to 16-base-pair DNA fragment. This structure misses some protein residues (92-96, 114-118 and 290-292) and a few basis in the DNA (helix C: G1 helix D: G1, C2 and C16). Representation of PDB 2AC0 file (7) rendered orthographically using VMD software.

The results are tabulated below in the Table 3.6:

Method	PA	EHSS	TM
CCS	3110	4020	3991.5

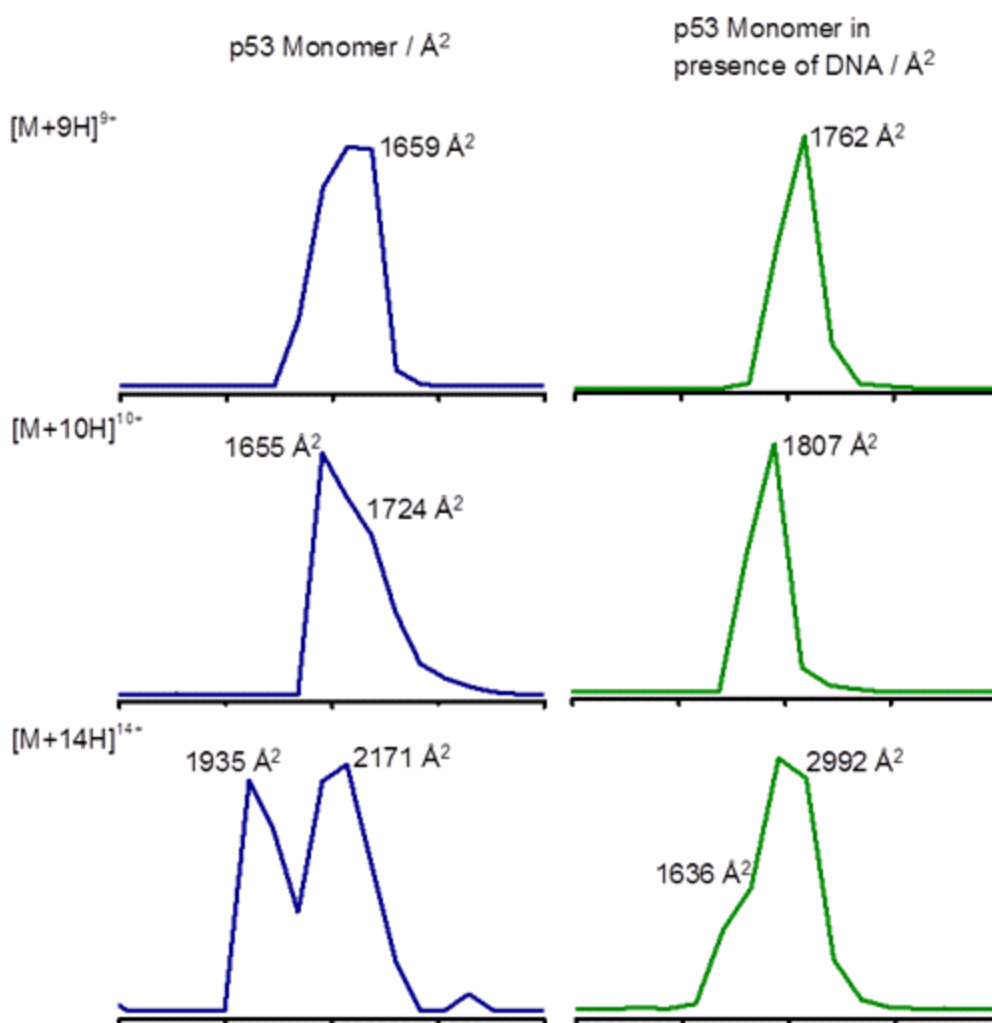
**Table 3.6: Theoretical measurements of collision cross sections for p53-DBD dimer cross-linked to 16-base-pair DNA fragment. Here values are shown using the Projection Approximation, Exact Hard Sphere Scattering and Trajectory Method.**

All experimentally measured CCS for p53-DBD/DNA complex are smaller (by 25%) than those calculated from the coordinates obtained from the crystal structure deposited by Ho *et al.* (8), here using the Trajectory Method (43) thus indicating to some extent the structural collapse in comparison to the X-ray structure. This again follows the common tendency for proteins we have previously described (44).

It is interesting to note that in the mass spectrum of p53-DBD with a double-stranded 12-base-pair DNA fragment, a charge state distribution of the monomeric species is also observed. Overall the charge state distribution is relatively similar to the one of p53-DBD sprayed from buffered solution conditions with no DNA present. The most abundant species in the spectrum represent  $[M+9H]^{9+}$  and  $[M+10H]^{10+}$  suggesting a more compact conformational family. Higher charged species, where  $11 \leq z \leq 18$ , are also observed at lower abundance (but at marginally higher intensity than those observed in the MS in the absence of DNA), indicative of a population of more extended states of p53.

In the mobility view, there are some important differences in the conformational landscapes of monomeric p53 observed, when the DNA-target site is available versus no DNA present. This is illustrated in the ATDs produced for p53 in Figure 3.13. Namely the  $[M+10H]^{10+}$  ion exists as a single conformational isomer with a CCS value of  $1807 \text{ \AA}^2$  in the presence of DNA oligomers but as two distinct conformational families, both of smaller CCS values without DNA ( $1655 \text{ \AA}^2$  and  $1724 \text{ \AA}^2$  respectively). Further, for the  $z = 9$  species, a CCS of  $1762 \text{ \AA}^2$  was

measured in the presence of DNA, and for the  $z = 9$  monomer, a CCS of 1659 Å<sup>2</sup> was measured without DNA, which is evident to some degree of preference for more extended states of p53 in the presence of DNA-binding target.



**Figure 3.13:** Comparison of ATDs for the three most intense charge states  $[M+9H]^{9+}$ ,  $[M+10H]^{10+}$ ,  $[M+14H]^{14+}$  monomeric form of p53-DBD without DNA versus DNA present. ATDs were obtained from 20 μM p53-DBD solution and from 20 μM p53-DBD with a 10 μM double-stranded 12-base-pair DNA fragment in a ratio 2:1 (protein to DNA, based on a protein monomer), respectively. Both were acquired via the DT IM-MS from buffered solution conditions.

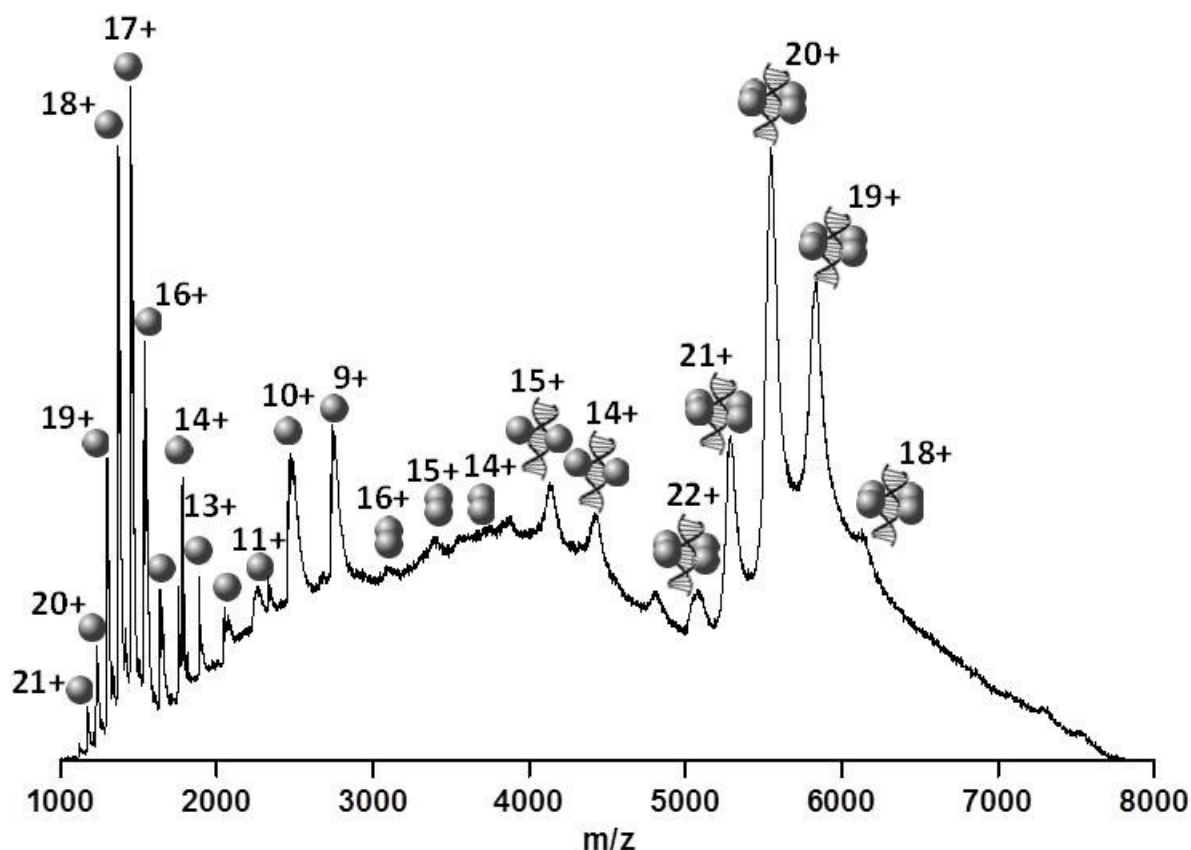
Here it appears that if a DNA-target site is available the equilibrium between the p53 monomeric conformational isomers is shifted towards more extended structures. This depopulation of the more compact form of p53 can be interpreted in at least two

different ways. First, the compact forms of p53 have been used to form dimeric species required for the p53-DNA complex. Secondly, in contrast in order to counteract the reduction in abundance of particular p53 conformers (chaperoned by the presence of DNA-target) and to maintain the equilibrium between the different dynamic states of p53, concentrations of other p53 conformers would change to favour the formation of the conformer with highest affinity for the specific DNA-binding site - and here this would suggest the more extended or unfolded populations are favoured by DNA. And finally, dissociation of p53 protein from DNA complex during the desolvation process cannot be ruled out.

This short argument supports the model for the formation of specific p53-DNA-binding complexes through a conformational diversity consistent with the hypothesis on protein structure known as the 'new view' (48,49). In contrast to the traditional paradigm of 'one sequence, one structure, one function' and a view on proteins as structurally restricted, the new view proposes the existence of proteins as an ensemble of conformational isomers (48,50). Intrinsically disordered proteins - and p53 is one of them - are the most extreme manifestation of protein-conformational diversity (51-53).

### **3.5.2. WT p53-DBD tetramer bound to ds 20 mer DNA**

In the presence of a longer strand of DNA, the observed spectrum (Figure 3.14) displays a wider charge state envelope over the acquired mass range (1000 - 8000  $m/z$ ). However the charge state distribution for p53-DNA complex is very narrow compared to monomeric p53. The interaction is highly cooperative and relies on the recognition of DNA sequences by the core domain and on the tetramerization of the protein, which is heavily mediated by DNA.



**Figure 3.14:** n-ESI spectrum of 20  $\mu$ M p53 DBD with a double-stranded 5  $\mu$ M 20-base-pair DNA fragment in a ratio 4:1 (protein to DNA, based on a protein monomer). Mass spectrum was obtained from a solution buffered with ammonium acetate and recorded via the DT IM-MS. The single spheres on the left represent a p53 monomer, double spheres correspond to the dimeric form of p53 whereas the spheres with DNA spirals correspond to the dimeric p53-DNA complex.

The WT p53-DBD binds consensus double-stranded 20 mer DNA mainly as a 4:1 complex where  $z = 18 - 22$ . But the dimeric species of p53-DBD bound to DNA also exists where  $z = 14$  and 15, which are likely intermediate states between transition from monomer (and dimer) to native tetramer. Pagel *et al.* (54) reports qualitatively similar findings where MS of the core-DNA complex shows three species populated simultaneously: monomeric p53-DBD, dimeric core bound to DNA and finally tetramer p53-DBD bound to DNA.

In the presence of ds 20mer DNA, the observed ion envelope for monomer is significantly much more abundant in comparison to charge states distribution with a shorter DNA fragment (with two consensus sites) or in the absence of DNA. The

observed charge states ranged from  $9 \leq z \leq 21$ . In particular the distribution ranging from  $15 \leq z \leq 21$  is more populated than the one ranging from  $9 \leq z \leq 14$  and so it can be inferred that p53 under these conditions favours more extended conformations. It would be interesting to see what further IM-MS measurements would reveal. Especially if the increase in the abundance of higher charge states observed in MS would be paralleled by the presence of more unfolded conformations in the ATDs produced in the mobility experiments.

### **3.6. Summary**

Mass spectrometry provides a compelling, detailed insight into the conformational flexibility of this tumour suppressor protein. The plasticity of p53 DNA-binding domain is reflected in the existence of more than one conformation, independent of any conformational changes prompted by binding. In a complex with DNA, p53-DBD binds all available consensus binding sites in both DNA constructs. WT p53 binds to ds 12mer DNA as a dimer and as ‘dimer of dimers’ to 20mer DNA, forming a 58 kDa and a 110 kDa complex. The interaction is cooperative, and is dependent on the recognition of DNA-target by the p53 and on the tetramerization of the protein, chaperoned by the presence of DNA. Conformational diversity of p53 is critical for the regulation of different cellular functions such as apoptosis and growth arrest. However, in terms of the design of molecules as effective drugs, the conformational spread of p53 is understandably an obstacle. A treatment that rectifies one function of the protein impaired by a mutation may not be enough to treat the disease.



### 3.7. References

1. Bargonetti, J., Manfredi, J.J., Chen, X., Marshak, D.R. and Prives, C. (1993) A proteolytic fragment from the central region of p53 has marked sequence-specific DNA-binding activity when generated from wild-type but not from oncogenic mutant p53 protein. *Genes & Development*, **7**, 2565-2574.
2. Pavletich, N.P., Chambers, K.A. and Pabo, C.O. (1993) The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes & Development*, **7**, 2556-2564.
3. Yang, A., Kaghad, M., Caput, D. and McKeon, F. (2002) On the shoulders of giants: p63, p73 and the rise of p53. *Trends in Genetics*, **18**, 90-95.
4. Cho, Y., Gorina, S., Jeffrey, P. and Pavletich, N. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346-355.
5. Cañadillas, J.M.P., Tidow, H., Freund, S.M.V., Rutherford, T.J., Ang, H.C. and Fersht, A.R. (2006) Solution structure of p53 core domain: Structural basis for its instability. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 2109-2114.
6. Wang, Y., Rosengarth, A. and Luecke, H. (2007) Structure of the human p53 core domain in the absence of DNA. *Acta Crystallographica Section D Biological Crystallography*, **63**, 276-281.
7. Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E. and Shakked, Z. (2006) Structural basis of DNA recognition by p53 tetramers. *Molecular Cell*, **22**, 741-753.
8. Ho, W.C., Fitzgerald, M.X. and Marmorstein, R. (2006) Structure of the p53 Core Domain Dimer Bound to DNA. *Journal of Biological Chemistry*, **281**, 20494-20502.
9. Joerger, A.C., Ang, H.C., Veprintsev, D.B., Blair, C.M. and Fersht, A.R. (2005) Structures of p53 cancer mutants and mechanism of rescue by second-site suppressor mutations. *Journal of Biological Chemistry*, **280**, 16030-16037.
10. Joerger, A.C., Ang, H.C. and Fersht, A.R. (2006) Structural basis for understanding oncogenic p53 mutations and designing rescue drugs. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 15056-15061.
11. Joerger, A.C., Allen, M.D. and Fersht, A.R. (2004) Crystal structure of a superstable mutant of human p53 core domain - Insights into the mechanism of rescuing oncogenic mutations. *Journal of Biological Chemistry*, **279**, 1291-1296.
12. Tu, C., Tan, Y.H., Shaw, G., Zhou, Z., Bai, Y., Luo, R. and Ji, X. (2008) Impact of low-frequency hotspot mutation R282Q on the structure of p53 DNA-binding domain as revealed by crystallography at 1.54 angstroms resolution. *Acta Crystallographica Section D Biological Crystallography*, **64**, 471-477.
13. Gorina, S. and Pavletich, N.P. (1996) Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science*, **274**, 1001-1005.

14. Derbyshire, D.J., Basu, B.P., Serpell, L.C., Joo, W.S., Date, T., Iwabuchi, K. and Doherty, A.J. (2002) Crystal structure of human 53BP1 BRCT domains bound to p53 tumour suppressor. *Embo J.*, **21**, 3863-3872.
15. Joo, W.S., Jeffrey, P.D., Cantor, S.B., Finnin, M.S., Livingston, D.M. and Pavletich, N.P. (2002) Structure of the 53BP1 BRCT region bound to p53 and its comparison to the Brcal1 BRCT structure. *Genes & Development*, **16**, 583-593.
16. Rainwater, R., Parks, D., Anderson, M.E., Tegtmeyer, P. and Mann, K. (1995) Role of cysteine residues in regulation of p53 function. *Molecular Cell Biology*, **15**, 3892-3903.
17. Hainaut, P., Rolley, N., Davies, M. and Milner, J. (1995) Modulation by copper of p53 conformation and sequence-specific DNA binding: role for Cu(II)/Cu(I) redox mechanism. *Oncogene*, **10**, 27-32.
18. Meplan, C., Richard, M.J. and Hainaut, P. (2000) Metalloregulation of the tumor suppressor protein p53: zinc mediates the renaturation of p53 after exposure to metal chelators in vitro and in intact cells. *Oncogene*, **19**, 5227-5236.
19. Verhaegh, G.W., Richard, M.J. and Hainaut, P. (1997) Regulation of p53 by metal ions and by antioxidants: dithiocarbamate down-regulates p53 DNA-binding activity by increasing the intracellular level of copper. *Molecular Cell Biology* **17**, 5699-5706.
20. Bullock, A.N., Henckel, J., DeDecker, B.S., Johnson, C.M., Nikolova, P.V., Proctor, M.R., Lane, D.P. and Fersht, A.R. (1997) Thermodynamic stability of wild-type and mutant p53 core domain. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 14338-14342.
21. Friedler, A., Veprintsev, D.B., Hansson, L.O. and Fersht, A.R. (2003) Kinetic instability of p53 core domain mutants: implications for rescue by small molecules. *Journal of Biological Chemistry*, **278**, 24108-24112.
22. Ang, H.C., Joerger, A.C., Mayer, S. and Fersht, A.R. (2006) Effects of common cancer mutations on stability and DNA binding of full-length p53 compared with isolated core domains. *Journal of Biological Chemistry*, **281**, 21934-21941.
23. Khoo, K.H., Andreeva, A. and Fersht, A.R. (2009) Adaptive evolution of p53 thermodynamic stability. *Journal of Molecular Biology*, **393**, 161-175.
24. Khoo, K.H., Joerger, A.C., Freund, S.M. and Fersht, A.R. (2009) Stabilising the DNA-binding domain of p53 by rational design of its hydrophobic core. *Protein Engineering Design & Selection*, **22**, 421-430.
25. Nikolova, P.V., Henckel, J., Lane, D.P. and Fersht, A.R. (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14675-14680.
26. Olivier, M., Eeles, R., Hollstein, M., Khan, M.A., Harris, C.C. and Hainaut, P. (2002) The IARC TP53 database: new online mutation analysis and recommendations to users. *Human Mutations*, **19**, 607-614.
27. Harms, K., Nozell, S. and Chen, X. (2004) The common and distinct target genes of the p53 family transcription factors. *Cellular and Molecular Life Sciences*, **61**, 822-842.

28. Kern, S.E., Kinzler, K.W., Bruskin, A., Jarosz, D., Friedman, P., Prives, C. and Vogelstein, B. (1991) Identification of p53 as a sequence-specific DNA-binding protein. *Science*, **252**, 1708-1711.
29. Rippin, T.M., Freund, S.M.V., Veprintsev, D.B. and Fersht, A.R. (2002) Recognition of DNA by p53 Core Domain and Location of Intermolecular Contacts of Cooperative Binding. *Journal of Molecular Biology*, **319**, 351-358.
30. Balagurumoorthy, P., Sakamoto, H., Lewis, M.S., Zambrano, N., Clore, G.M., Gronenborn, A.M., Appella, E. and Harrington, R.E. (1995) Four p53 DNA-binding domain peptides bind natural p53-response elements and bend the DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 8591-8595.
31. Nagaich, A.K., Appella, E. and Harrington, R.E. (1997) DNA bending is essential for the site-specific recognition of DNA response elements by the DNA binding domain of the tumor suppressor protein p53. *Journal of Biological Chemistry*, **272**, 14842-14849.
32. Klein, C., Planker, E., Diercks, T., Kessler, H., Kunkele, K.P., Lang, K., Hansen, S. and Schwaiger, M. (2001) NMR spectroscopy reveals the solution dimerization interface of p53 core domains bound to their consensus DNA. *Journal of Biological Chemistry*, **276**, 49020-49027.
33. Joerger, A.C. and Fersht, A.R. (2008) Structural biology of the tumor suppressor p53. *Annual Review of Biochemistry*, **77**, 557-582.
34. Joerger, A.C. and Fersht, A.R. (2010) The tumor suppressor p53: from structures to drug discovery. *Cold Spring Harbor Perspectives in Biology*, **2**, a000919.
35. Hainaut, P. and Wiman, K.G. (2005) *25 years of p53 research*. Springer.
36. Patel, S., George, R., Autore, F., Fraternali, F., Ladbury, J.E. and Nikolova, P.V. (2008) Molecular interactions of ASPP1 and ASPP2 with the p53 protein family and the apoptotic promoters PUMA and Bax. *Nucleic Acids Research*, **36**, 5139-5151.
37. Nikolova, P.V., Wong, K.-B., DeDecker, B., Henckel, J. and Fersht, A.R. (2000) Mechanism of rescue of common p53 cancer mutations by second-site suppressor mutations. *Embo J.*, **19**, 370-378.
38. Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry*, **182**, 319-326.
39. Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinou, K., Bateman, R.H., Bowers, M.T. and 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**, 1-12.
40. McCullough, B.J., Kalapothakis, J., Eastwood, H., Kemper, P., MacMillan, D., Taylor, K., Dorin, J. and Barran, P.E. (2008) Development of an ion mobility quadrupole time of flight mass spectrometer. *Analytical Chemistry*, **80**, 6336-6344.

41. Wytttenbach, T., Helden, G.v., Batka Jr, J.J., Carlat, D. and Bowers, M.T. (1997) Effect of the long-range potential on ion mobility measurements. *Journal of the American Society for Mass Spectrometry*, **8**, 275-282.
42. Shvartsburg, A.A. and Jarrold, M.F. (1996) An exact hard-spheres scattering model for the mobilities of polyatomic Ions. *Chemical Physics Letters*, **261**, 86-91.
43. Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C. and Jarrold, M.F. (1996) Structural information from ion mobility measurements: effects of the long-range potential. *The Journal of Physical Chemistry*, **100**, 16082-16086.
44. Jurneczko, E. and Barran, P.E. (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst*, **136**, 20-28.
45. Jurneczko, E., Cruickshank, F., Porrini, M., Nikolova, P., Campuzano, I.D., Morris, M. and Barran, P.E. (2012) Intrinsic disorder in proteins: a challenge for (un)structural biology met by ion mobility-mass spectrometry. *Biochemical Society Transactions*, **40**, 1021-1026.
46. Faull, P.A., Florance, H.V., Schmidt, C.Q., Tomczyk, N., Barlow, P.N., Hupp, T.R., Nikolova, P.V. and Barran, P.E. (2010) Utilising ion mobility-mass spectrometry to interrogate macromolecules: Factor H complement control protein modules 10–15 and 19–20 and the DNA-binding core domain of tumour suppressor p53. *International Journal of Mass Spectrometry*, **298**, 99-110.
47. <http://ambermd.org/>.
48. James, L.C. and Tawfik, D.S. (2003) Conformational diversity and protein evolution – a 60-year-old hypothesis revisited. *Trends in Biochemical Sciences*, **28**, 361-368.
49. Vukojevic, V., Yakovleva, T. and Bakalkin, G. (2011), *p53*. Springer US, Vol. 1, pp. 127-141.
50. Frauenfelder, H., Sligar, S.G. and Wolynes, P.G. (1991) The energy landscapes and motions of proteins. *Science*, **254**, 1598-1603.
51. Wright, P.E. and Dyson, H.J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, **293**, 321-331.
52. Tompa, P. (2002) Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, **27**, 527-533.
53. Uversky, V.N. (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Science*, **11**, 739-756.
54. Pagel, K., Natan, E., Hall, Z., Fersht, A.R. and Robinson, C.V. (2013) Intrinsically disordered p53 and its complexes populate compact conformations in the gas phase. *Angewandte Chemie International Edition* **52**, 361-365.

# 4

## Common p53 Cancer Mutants

*It has become evident that p53 is a molecular node at the junction of an extensive and complex network of stress response pathways. Deregulation of this protein has a colossal effect on carcinogenesis, as mutated p53 can induce an amplified epigenetic instability of tumour cells, facilitating and accelerating the evolution of the tumour. In order to understand the mechanism how p53 is inactivated, it is necessary to understand its structure and how it responds to mutation. Such knowledge may ultimately provide basis for the rational design of novel therapeutics that may reverse the effects of mutation in the treatment of cancer. Herein, the conformational phenotypes exhibited by common cancer-associated mutations are surveyed. Whether the second-site suppressor from loop L1, H115N, could trigger conformational changes in p53 hotspot cancer mutations is also tested.*

## 4. Common p53 Cancer Mutants

### 4.1. Mapping Cancer Mutations on the DNA-binding Domain

Tumour suppressor p53 is the most mutated protein in human cancers. Mutations are found in almost every type of cancer. Malignancies with high mutational frequencies (40-50%) include ovarian, colorectal, esophageal, head and neck and lung cancers whereas tumours of the brain, breast, liver and stomach show intermediary mutation frequency (20-30%) (1). A distinctive feature of the p53 mutational map is an exceptionally high rate of occurrence of missense mutations which result in the expression of a mutant protein. Indeed, these single-point amino acid substitutions in p53 lead to abrogation of protein function, rather than deletions or nonsense mutations, as it is the case with most tumour suppressor proteins. The latest release of the International Agency for Research on Cancer (IARC R16, November 2012) TP53 mutation database reports 29575 cases of somatic mutations (2). These missense mutations can lead to loss or alteration of p53 binding activity to its downstream targets and may thus induce aberrant cell proliferation, with consequent malignant cellular transformation.

The majority of tumour-derived mutations (over 80%) target the core DNA-binding domain of p53 (3). There are at least three reasons to explain why the cancer-associated mutations are predominantly mapped on the DNA-binding domain of p53. Firstly, the oncogenic core domain mutants are known to have a phenomenal 'dominant-negative effect' on WT p53 if present (4,5). Since the transcriptional function of p53 relies on formation of homotetramers, the mutated p53 proteins may interfere with WT by forming hetero-oligomers (containing both mutant and WT subunits) malfunctioned in sequence specific DNA-binding (6,7). WT p53 core domains in a mixed tetramer may still be able to recognize DNA response elements, but binding cooperativity will be progressively lost upon amalgamation of inactive mutants (8,9). One of the key findings of the study by Chan *et al.* (10) study with the

R249S and R273H mutants, is that at least three DNA binding-defective mutants are required to inactivate the p53 tetramer.

A second rationale is the low thermodynamic stability of the p53 DNA-binding domain, which is just above 37°C (above this temperature p53 DNA-binding domain partially unfolds and forms aggregates) (11-13). The intrinsic low stability of core domain means that virtually every non-conservative substitution in the core domain will reduce the melting temperature below 37°C and have a profound effect on the amount of folded protein in the cell.

The third explanation is that many of the functionally important amino acid residues of this domain are arginines. Arginine codons are particularly prone to mutagenesis, as they contain CG dinucleotides, which, when damaged, are restored with lower fidelity than other dinucleotides (14).

Approximately 35% of missense mutations in this domain cluster within six so-called 'hotspot' codons detectable in almost every type of cancer (175, 245, 248, 273, and 282). Four of these codons correspond to arginine residues (175, 248, 273 282) involved in direct DNA contact (residues 248 and 273) or in stabilisation of the DNA-binding surface (residues 175 and 282) (3).

## **4.2. Structural Classes of p53 Core Domain Cancer Mutants**

### **4.2.1. DNA-contact Mutations**

DNA-contact mutations inactivate p53 by removing the critical DNA-contact amino acid residues, for instance Arg-248, Arg-273, and Arg-280. Two of the five most frequent cancer mutations, R273H and R273C have a minor effect on the thermodynamic and kinetic stability of the protein and the overall architecture of the DNA-binding surface is not perturbed (15,16). These findings are in accord with the results of Ang *et al.* (17) which show that the R273H mutation displays only residual DNA-binding activity *in vitro* and the binding selectivity is significantly impaired.

Crystallographic data confirmed that the mutation has not induced any major conformational changes in the adjacent side chains (Asp-281 and Phe-134) (15,16). A similar scenario with rather more localized structural changes can be expected for DNA-contact mutations Arg-248 and Arg-280.

#### **4.2.2. Structural Mutations in the L3 loop**

Another class of mutations corresponds to structural cancer hotspot mutations in the L3 loop, G245S and R249S. The G245S mutation prompts minor conformational changes that essentially affect residues in the core dimerization interface upon binding to a DNA half-site, which weakens DNA-binding, presumably due to reduced binding cooperativity (16,17).

The changes induced by the R249S mutation are much more pronounced as Arg-249 is essential for stabilizing the conformation of the L3 loop via its guanidinium group. This perturbation affects both DNA-contacts (via Arg-248) and the core dimerization interface, consequently resulting in substantially weakened DNA-binding in addition to a loss of protein stability (15,17).

#### **4.2.3. Disruption of the Zinc-binding Region**

Structural mutations in the zinc-binding region are highly destabilizing either due to a reduction in zinc-binding (C242S) or in structuring the zinc-binding site (R175H). Arg-175 is located in the L2 loop, mutations found in this codon result in complete abrogation of DNA-binding activity. Surprisingly the corresponding alanine mutant, R175A, is only moderately destabilised as compared to R175H and shows 70% WT activity in DNA-binding. Equally, it was found that the R175C mutant essentially maintains the WT cell cycle arrest and pro-apoptotic activity, whereas these functions are lost upon introduction of large bulky residues (e.g. for mutants R175W and R175Y) that presumably cause structural distortion and directly interfere with zinc-binding (18).



#### 4.2.4. Structural Mutations at the Periphery of the DBD

One of the most common mutational hot spots is at position Arg-282, located in the C-terminal helix H2, which is part of a loop-sheet-helix motif that binds to the major groove of DNA response elements. Joerger *et al.* (16) solved the high-resolution crystal structure of several oncogenic mutations, including R282W. The data reveal that introduction of the R282W hotspot mutation induces distinct, substantial structural distortions in the loop-sheet-helix motif but the overall fold of the protein is not compromised. Several stabilizing interactions are also lost which are reflected in the observed loss in the thermodynamic stability of the mutant. Importantly, the scaffold of the DNA-binding surface, including helix H2, is fundamentally preserved, which explains why the mutant is still able to bind to DNA at sub physiological temperatures both *in vitro* and *in vivo*, at least for high-affinity response elements (4,19).

Similarly, the structure of H168R cancer mutant positioned in the  $\beta$ -turn region of the L2 loop shows that the global fold of the protein is maintained but exhibits local structural perturbations (19,20).

#### 4.2.5. $\beta$ -sandwich Mutations

Y220C is the most frequent tumor mutation found outside the structural elements that form the DNA-binding surface (loops L2 and L3 and the loop-sheet-helix motif) and accounts for approximately 1.4% of somatic p53 missense mutations (2). This mutation creates a solvent-accessible crevice at the end of a  $\beta$ -sandwich but leaves the overall structure of the core domain intact, as reported by Joerger *et al.* (16).

Likewise, other  $\beta$ -sandwich oncogenic large-to-small substitutions, V143A and F270L form internal cavities in the hydrophobic core of the protein without collapse of the surrounding structure (16,19). Taken together, a common structural feature of the  $\beta$ -sandwich mutants appears to be that they induce only minor structural local

perturbations, although the effect on the stability of the core is generally much more severe than for the hotspot mutations in the DNA-binding area.

### **4.3. Rescue of p53 Mutants by ‘Conformational’ Drugs**

p53 is mutated or functionally inactivated in the vast proportion of human cancers, pharmacological restoration of WT p53 conformation and function by small molecules targeting mutant p53 proteins has been proposed as a tactic for the design of anti-cancer drugs (21-23). Over the past few years, a number of small molecules that reactivate mutant p53 have been identified and characterised. The mechanism of action of many compounds remains controversial and has not been fully established yet (reviewed by Bykov (24) and Wiman *et al.* (25). This section will describe only some of these molecules and discuss possible mechanisms for mutant p53 rescue.

A designed nine-residue peptide, CDB3, is derived from p53-binding protein 2 (53BP2 or ASPP), a known p53-binding protein that interacts with the p53 core domain and enhances p53-mediated transactivation (26). NMR measurements have mapped the CDB3 docking site to a surface region of p53 that partly intersects with the DNA-binding surface (including the loop-sheet-helix motif and parts of  $\beta$ -strand S8), although the exact binding mechanism of CDB3 is still unclear (27). The rationale behind this method is that a peptide binding with a higher affinity to the folded state of p53 rather than to the mutated form will shift the equilibrium towards the active WT conformation. *In vitro* studies revealed that CDB3 is able to induce the refolding of the denatured p53 core domain, as well as the re-establishment of sequence-specific DNA-binding to various p53 mutants (27). Remarkably, CDB3 was also shown to reinstate the WT conformation transcriptional activity of two hotspot p53 mutants, R273H and R175H, in human tumour cells (28).

Foster and coworkers screened a chemical library for compounds that could stabilize the folded conformation of the core domain and identified the first most promising small molecule, CP-31398 that can rescue mutant p53 (23). This Pfizer compound

was found to stabilize the core domain and enhance the transcriptional activity of p53, but the exact mechanism of action of CP-31398 remains still elusive (23). Further Rippin *et al.* (29) showed conflicting data. This NMR study failed to detect any binding of CP-31398 compound to the p53 core domain, rather, gel shift assays and NMR indicated that the compound binds DNA.

p53 reactivation and induction of massive apoptosis (PRIMA-1) and mutant p53 reactivation and induction of rapid apoptosis (MIRA-1) are two compounds that were identified in a cellular screen of a chemical library for compounds that suppress the growth of tumour cells (21,30). PRIMA-1 and MIRA-1 were found to restore sequence-specific DNA-binding and convert the mutant p53 conformation to WT, leading to transactivation of p53 target genes (21,30).

Restoring the DNA binding function of p53 mutant proteins which are highly expressed in tumours is a prominent goal in the field. With the emerging detailed structural data on mutant p53, rational drug design and virtual screening become progressively attractive methods. Nonetheless there is a whole spectrum of different p53 cancer mutations, and development of the 'Holy Grail', a generic drug that could stabilize a range of unstable oncogenic mutants, remains undoubtedly a challenge. Whether this is truly feasible, awaits future studies.

This chapter focuses on the use of IM-MS to probe the conformational diversity of common single oncogenic mutations of the p53 DNA-binding domain: R249A, R273H, K292I, A276Y, and R175H. The second-site suppressor from loop L1, the H115N mutant, is also examined, to determine whether it could trigger conformational changes in p53 hotspot cancer mutation, using a double mutant approach (R248Q/H115N, R249/H115N, and R273H/H115N). Locations of mutated residues used in the studies are depicted schematically on the three-dimensional structure of p53 in Figure 4.1. The sequence-specific DNA-binding properties of the H115 mutant are also revealed.

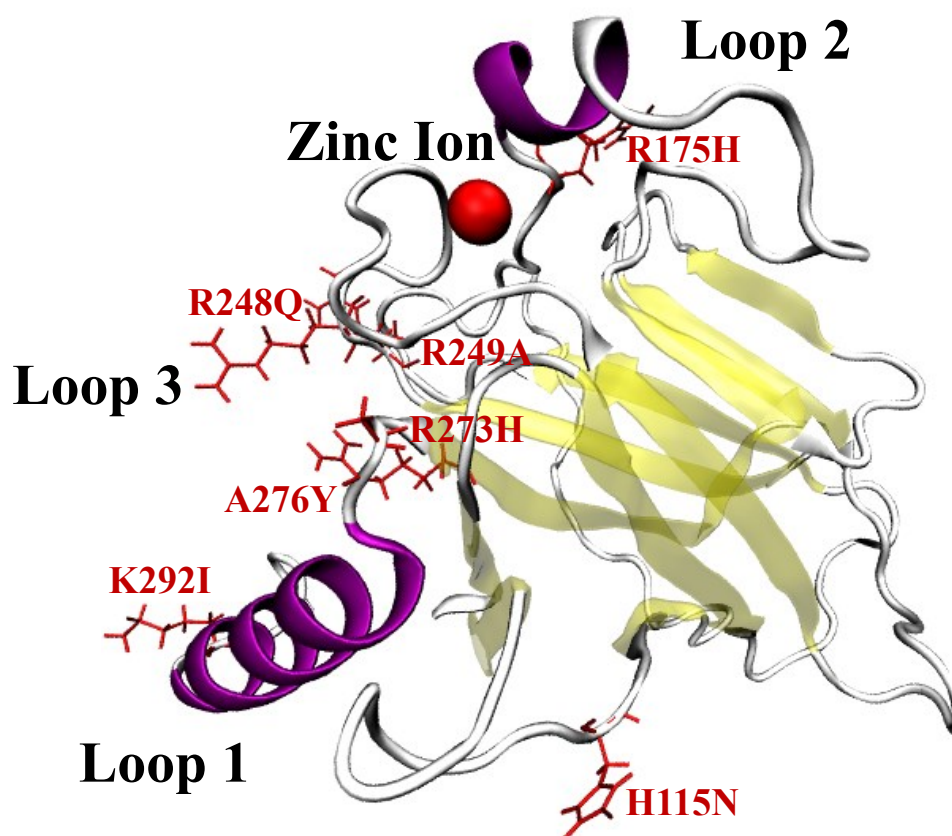


Figure 4.1: Structure of the DNA-binding core domain of p53 (residues 94 - 312). Locations of residues mutated in the experiments are labelled in red, the zinc atom is represented as a red sphere. PDB 2FEJ file (31) was used to generate image in VMD software.

## 4.4. Methodology

### 4.4.1. Site-directed Mutagenesis, p53 Expression and Purification

Single-point mutants of p53 DNA-binding domain (R249A, A276Y, K292I and H115N) and double mutants (R248Q/H115N, R249A/H115N, and R273H/H115N) were expressed and purified by Dr Penka V. Nikolova (King's College London, UK). R175H and R273H mutants were supplied by Dr David J. Clarke (University of Edinburgh, UK). All mutants were prepared according to the method outlined here. Cloning of the proteins was performed using pRSETA expression vector (Invitrogen)

which was modified so that it did not contain the His tag (20,32,33). p53 mutants were generated using the QuikChange® site-directed mutagenesis kit (Stratagene). Proteins were expressed at 37°C in *Escherichia coli* C41 cells and grown up to an optical density at 600 nm ( $OD_{600}$ ) of 0.8. The temperature was reduced to 22°C before cell cultures were induced with 1 mM isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG) and grown overnight. Cells were harvested by centrifugation and lysed using Bug Buster Protein Extraction Reagent. Benzonase nuclease and EDTA-free protease inhibitor tablets were also used during this process. The soluble fraction was loaded on to an SP-Sepharose column eluted with NaCl gradient followed by gel-filtration chromatography using a HP26/60 Superdex 200 column.

#### **4.4.2. Sample Preparation**

Protein samples were flash frozen in liquid nitrogen and stored in a -80°C freezer. Prior to mass spectrometry and ion mobility mass spectrometry measurements, protein samples were thawed and dialysed in 50 mM ammonium acetate for 2 h at 4°C (with an exception of R175H mutant which was dialysed in 100 mM ammonium acetate). Solutions were made up to contain 10% propan-2-ol to aid electrospray desolvation.

##### **4.4.2.1. DNA Oligonucleotides**

Double-stranded 12-mer DNA (5'GGAACATGTTCC CCTTGTACAAGG'3 containing two consensus binding sites  $M_R = 7290.8$  Da) used in this study was constructed by annealing purified synthetic oligonucleotides purchased from ATDBio (University of Southampton, UK). Annealing was performed by heating the oligonucleotides at 95°C for 5 minutes and gradual, slow cooling to room temperature to ensure an efficient hybridization. The final concentration of the double-stranded DNA used for the binding experiments is based on the assumption that all of DNA had annealed completely, however trace amounts of unbound DNA/hairpin structures might be still present.

#### **4.4.3. Mass Spectrometry and Ion Mobility Mass Spectrometry**

The travelling wave based ion mobility mass spectrometry (TW IM-MS) experiments were performed on a Synapt HDMS system (34) (Waters Corporation, Manchester, UK) described in detail in Chapter 2, Section 2.5.1. Samples were introduced into the source region by nano-electrospray ionisation (n-ESI). The instrument was operated in a positive ionisation mode with a capillary voltage of 0.8 -1 kV, cone voltage of 65 V and source temperature of 30 °C. Source backing pressure was elevated to 4 mBar for improved transmission of ions. Travelling wave heights of 13, 14 and 15 V and 250 m/s were employed. The ion mobility separator containing nitrogen gas was operated at a pressure of 1 mBar. During data acquisition, the parameters of the instrument were optimised to obtain the optimum ion mobility separation. Equine heart cytochrome c and myoglobin was analysed under denaturing conditions to create a calibration curve for cross-section measurements versus values obtained from drift tube ion mobility mass spectrometry (DT IM-MS) studies. The calibration procedure is described in Chapter 2, Section 2.5.2.

Drift tube ion mobility mass spectrometry (DT IM-MS) experiments were performed on an in-house modified quadrupole time-of-flight mass spectrometer, the MoQToF (35). A more detailed description of this instrument can be found in Chapter 2, Section 2.4.1. Ions were produced by positive n-ESI using Z-spray source, within a spray voltage range 1.0 to 1.8 kV and a source temperature of 80°C. The drift cell was filled with helium buffer gas at the average pressure of 3.5 - 4.0 Torr (4.7 - 5.3 mBar) and a temperature between 298 - 302 K. The electric potential across the cell was varied from 60 to 15 V and ATDs were recorded at eight drift voltages. The mobility of the ion of interest was calculated from a linear plot of average arrival time versus pressure/drift voltage ( $P/V$ ) and from this the rotationally-averaged collision cross section (CCS) for each resolvable species at a given charge state were obtained.

#### 4.4.4. Production of Waterfall Plots for data obtained via TW IM-MS

The waterfall plots for p53 mutants were produced in Origin 8.5.1 software (OriginLab, Northampton, MA, USA). Multi-curve waterfall charts were created by plotting estimated CCS values ( $x$ ), calculated as described in Chapter 2, Section 2.5.2, against peak intensity values ( $y$ ) and charge ( $z$ ). The intensities of the peaks were obtained from Mass Lynx software (version 4.1, Waters, Manchester, UK).

#### 4.4.5. Production of Waterfall Plots for data obtained via DT IM-MS

The waterfall plot for the R273H mutant was produced in Origin 8.5.1 software (OriginLab, Northampton, MA, USA). A multi-curve waterfall chart was created by plotting the average CCS ( $x$ ), against peak intensity values ( $y$ ) and charge ( $z$ ). The intensities of the peaks were obtained from the Mass Lynx software (version 4.1, Waters, Manchester, UK). CCS values were calculated from arrival time distribution using *Equation 4.1* below:

$$\Omega_{avg} = \frac{(18\pi)^{1/2}}{16} \left[ \frac{1}{m_b} + \frac{1}{m} \right]^{1/2} \frac{ze}{(k_B T)^{1/2}} \frac{1}{N} \frac{t_d V}{L^2} \quad \text{Equation 4.1}$$

where  $m$  and  $m_b$  are the masses of the ion and buffer gas,  $z$  is the ion charge state,  $e$  is the elementary charge,  $k_B$  is the Boltzmann constant,  $T$  is the gas temperature,  $N$  is the gas number density,  $L$  is the drift cell length,  $V$  is the voltage applied across the drift cell (here 60 V) and  $t_d$  is the drift time. This relationship uses the drift time ( $t_d$ ) of each ion:

$$\text{Drift Time} = \text{Arrival Time} - \text{Dead Time} \quad \text{Equation 4.2}$$

Dead time was calculated by taking an average value of the intercept across the arrival time distribution for each charge state.

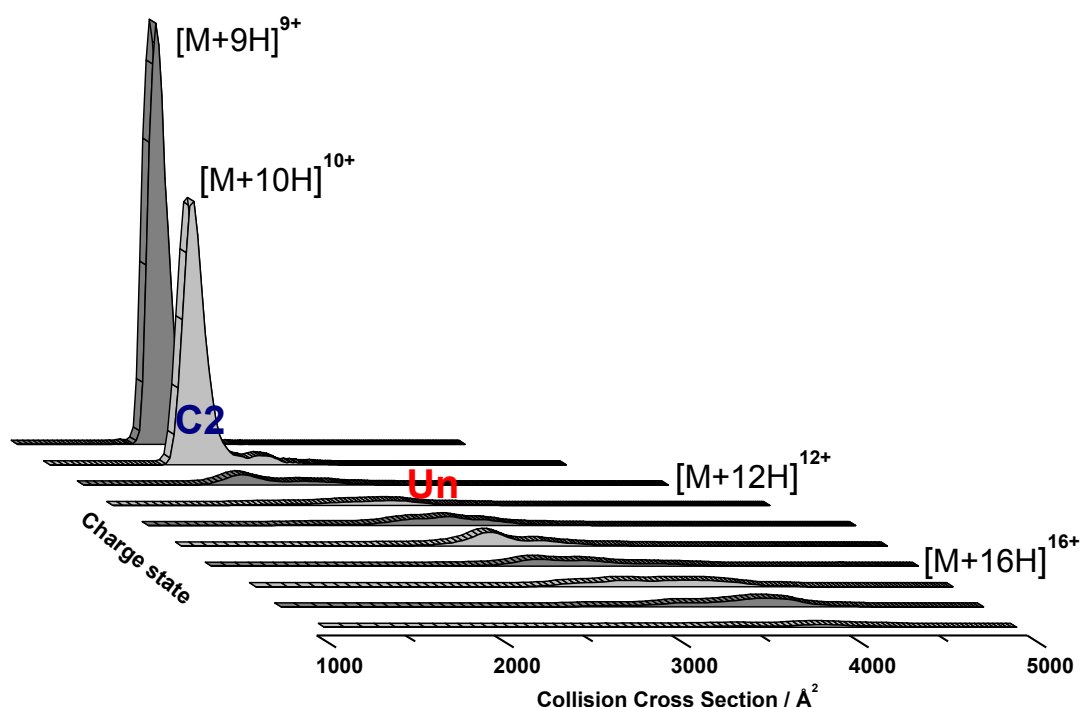
## **4.5. Insights into the Diverse Conformational Nature of Common p53 Cancer Mutants**

This section reports on IM-MS studies on the conformational diversity of common cancer-associated p53 mutants. The conformational phenotypes are defined and the variation in these, as exhibited by five single-point mutations (R249A, R273H, K292I, A276Y, and R175H), are monitored.

### **4.5.1. Structural Mutation in the L3 loop - R249A**

One of the most common mutational hot spots is at position arginine 249, which is replaced by serine, thus inducing substantial structural distortions either at the periphery of L2 loop or directly at the center of L3 loop of the DNA-binding surface (15). This perturbation indirectly affects DNA contacts (via Arg-248) and the core dimerization interface, resulting in destabilization of the protein and considerably impaired DNA-binding (13,15,17). The far-UV CD spectra revealed that R249S mutation has a similar conformation to WT p53-DBD (13). Here the related mutation R249A was examined with IM-MS (Figure 4.2 and Figure 4.3 for the mass spectrum). When compared to WT p53 (refer to Chapter 3, Section 3.4.4), it is clear that this single point mutation inflicts a dramatic change to the conformational space occupied by the protein, with the R249A mutant possessing noticeably less flexibility.





**Figure 4.2:** Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of R249A mutation sprayed from buffered solution conditions (63  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via the TW IM-MS. The x, y and z axis show the collision cross section (in Å<sup>2</sup>), charge state (range  $9 \leq z \leq 18$ ), and the relative intensity, respectively.

The majority of R249A presents as  $z = 9$  and 10 ions, both in a compact form with similar CCS (1789 Å<sup>2</sup> and 1818 Å<sup>2</sup>, respectively) as for the C1 form of the WT p53, although the width of the ATDs for the mutant is less than that for the WT, thus implying a more ‘molten globule-like’ compact conformer for this single point mutant, which is here denoted C2. The  $z = 9$  charge state is also more intense here compared to WT p53, which suggests a more compact structure. For  $11 \leq z \leq 18$ , multiple unfoldomers (Un) for each charge state are observed. However these extended conformers are less populated than in WT p53. Overall, the gas-phase structure of R249A favours a compact geometry, suggesting that the mutation conferred intrinsic stability to a collapsed form, which is significantly smaller (by 14%) than that found in WT p53.

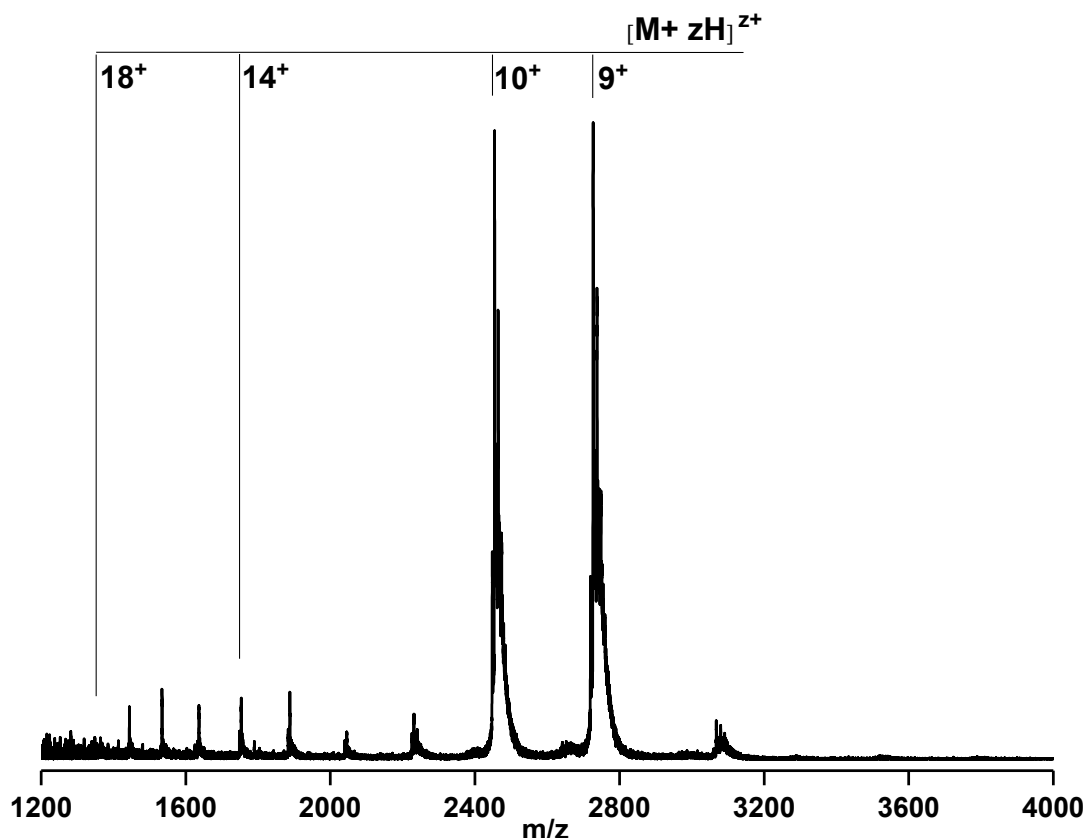
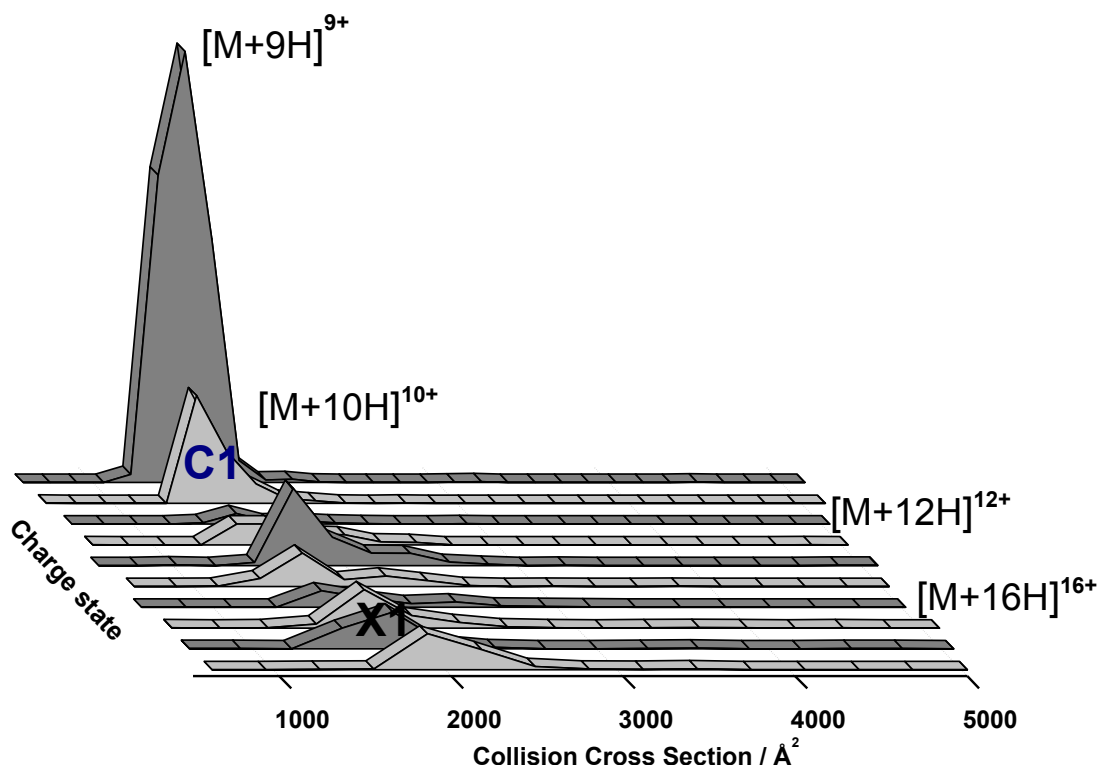


Figure 4.3: n-ESI spectrum of 63  $\mu$ M R249A mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

#### 4.5.2. DNA Contact Mutation - R273H

Cancer-associated mutants, such as R273H, are present at the p53 DNA-binding interface and are referred to as hotspot contact mutants. R273 is one of the most frequently mutated codons, representing about 20% of all cancer mutants in the p53 database (2). This mutation has a marginal effect on the thermodynamic stability of the protein, but affects the conformation required for DNA binding (11,17). The introduction of the imidazole ring associated with the histidine residue has been reported to perturb the local environment, and to break a salt bridge between R273 and the carboxylate group of D281 in helix 2 (36). Binding of full-length R273H to the *gadd45* recognition sequence is weakened approximately 1000-fold *in vitro* as a consequence of mutation, and binding specificity is largely lost (17). The residual

binding activity of R273H is too weak to perform a regular transcriptional function *in vivo* (4,37).



**Figure 4.4:** Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of R273H mutation sprayed from buffered solution conditions (100  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via the DT IM-MS.

Data for this mutation were obtained using DT IM-MS (Figure 4.4 and Figure 4.5 for the mass spectrum). The acquired mass spectrum has a high baseline and broad peaks in the high  $m/z$  range, a common appearance of unstable and aggregated species. Here the compact conformation is retained over a large range of charge states ( $9 \leq z \leq 16$ ) and the more extended forms (X1) are only present at the very high charge states. This remarkable stability of a compact conformation and the lack of unfoldomers indicate that this mutation favours more compact and rigid conformations.

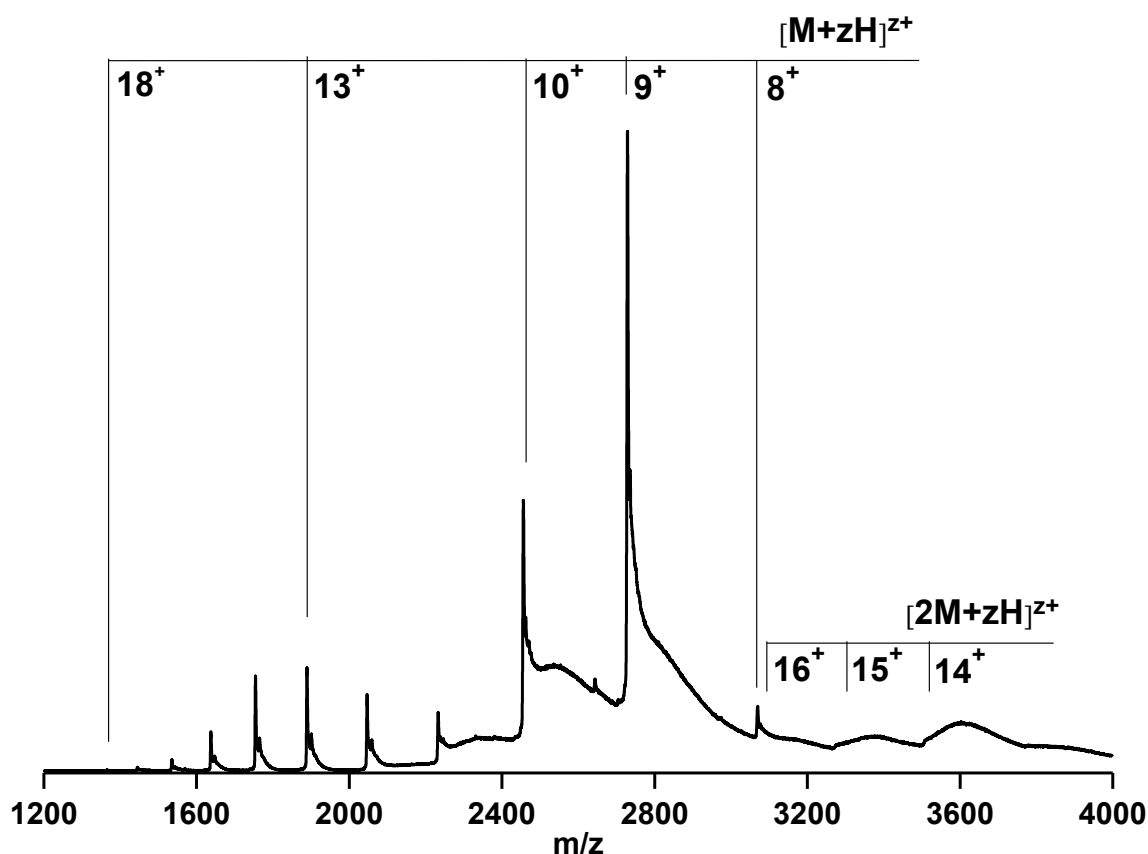


Figure 4.5: n-ESI spectrum of 100  $\mu$ M R273H mutation obtained via the DT IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

### 4.5.3. K292I Mutation

Cellular levels of p53 are controlled elaborately through ubiquitylation and proteasomal degradation, which are mediated by numerous E3 ubiquitin ligases. Interestingly, there are two sequentially located lysine residues in the p53-DBD, K291, and K292, which are required for polyubiquitination and degradation of p53 mediated by makorin ring finger protein 1 (MKRN1) (38). These two highly conserved residues are mutated in several different types of human tumors, suggesting their critical roles in p53 function (2). The conformational phenotype exhibited by one of these cancer-associated mutants, K292I, which is implicated in the inhibition of protein-protein interactions was examined (Figure 4.6 and Figure 4.7 for the mass spectrum).

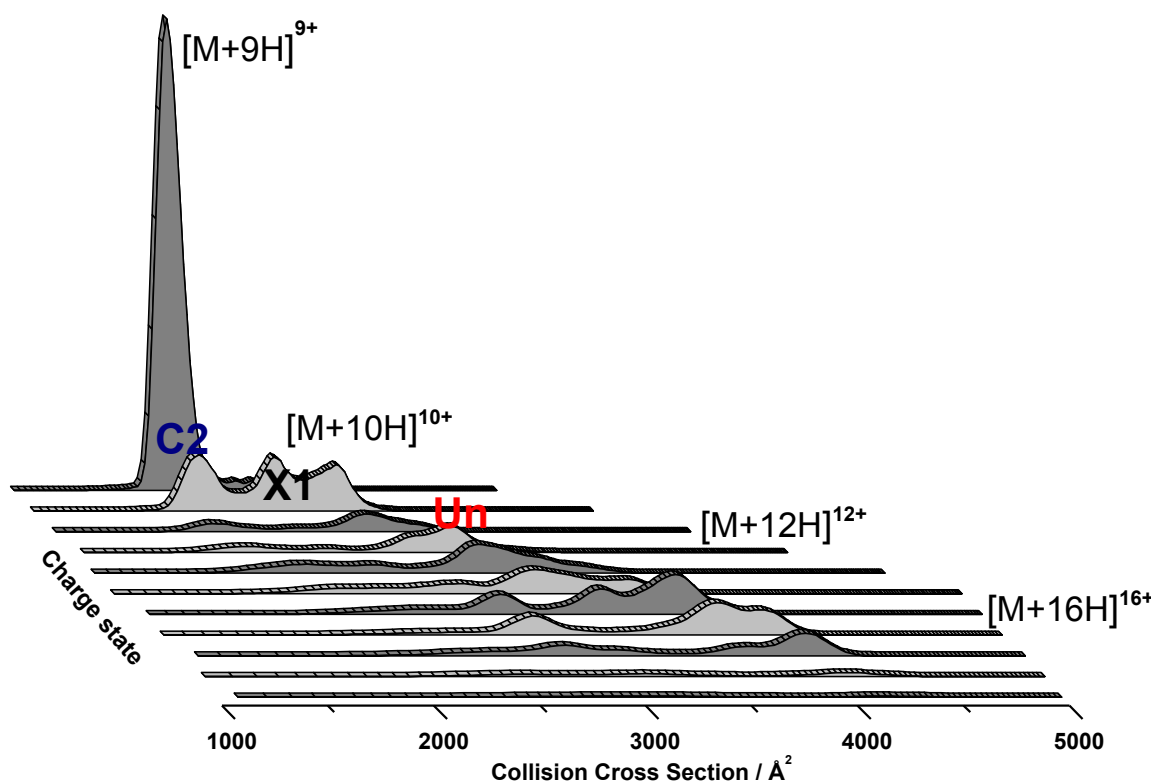


Figure 4.6: Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of K292I mutation sprayed from buffered solution conditions (75  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via the TW IM-MS.

Here a single compact conformation is only exhibited by  $z = 9$  (C2), gradually shifting to more extended structures (X1). For  $10 \leq z \leq 18$ , multiple unfoldomers are observed and their distribution is comparable to that of the WT p53, although the intensity varies.

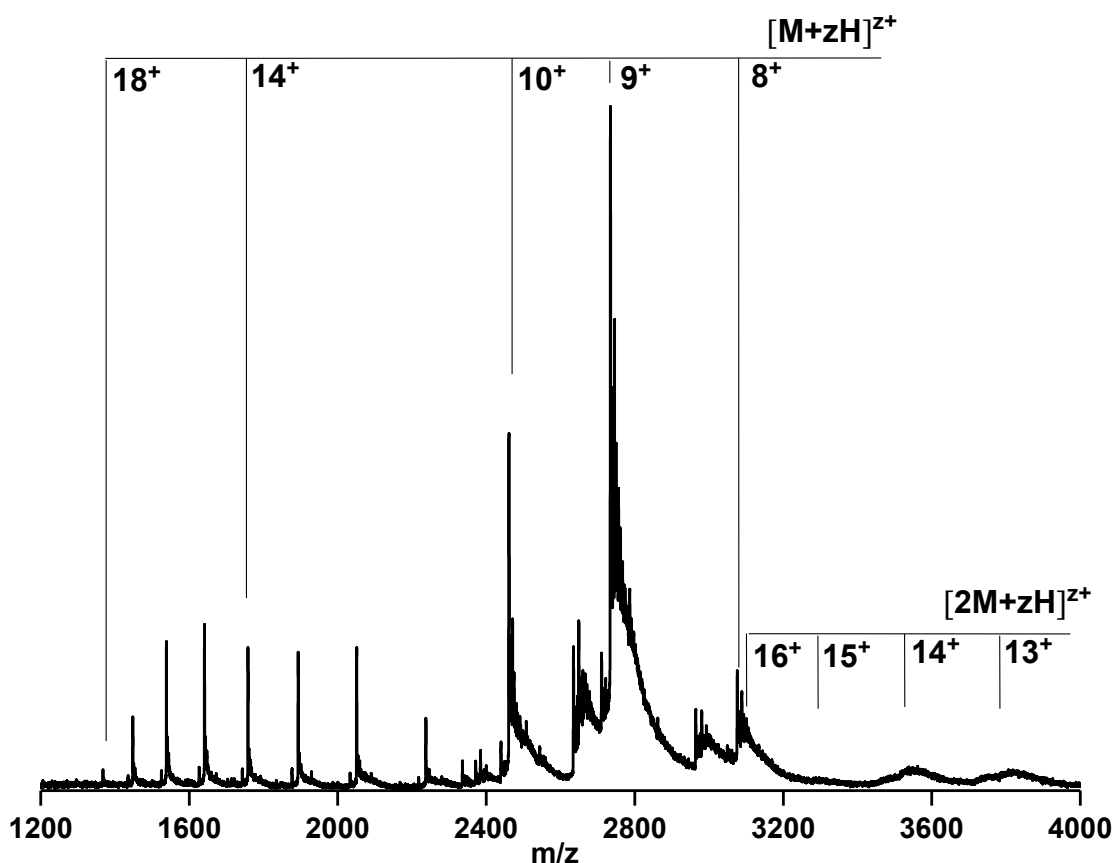


Figure 4.7: n-ESI spectrum of 75  $\mu$ M K292I mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

#### 4.5.4. A276Y Mutation

Another mutation, A276Y, is an example of a contact mutation as it directly binds to DNA (3,39). However, the substitution of the alanine residue with a larger aromatic amino acid tyrosine also has a pivotal effect on the function of protein. This mutation maps to the binding sites of protein kinases and Mdm2, a major regulator of p53 levels in the cell (40). In mobility space (Figure 4.8 and Figure 4.9 for the mass spectrum) the majority of the protein molecules are located in the  $z = 9$  and 10 region, and exist as a compact conformational family (C2).

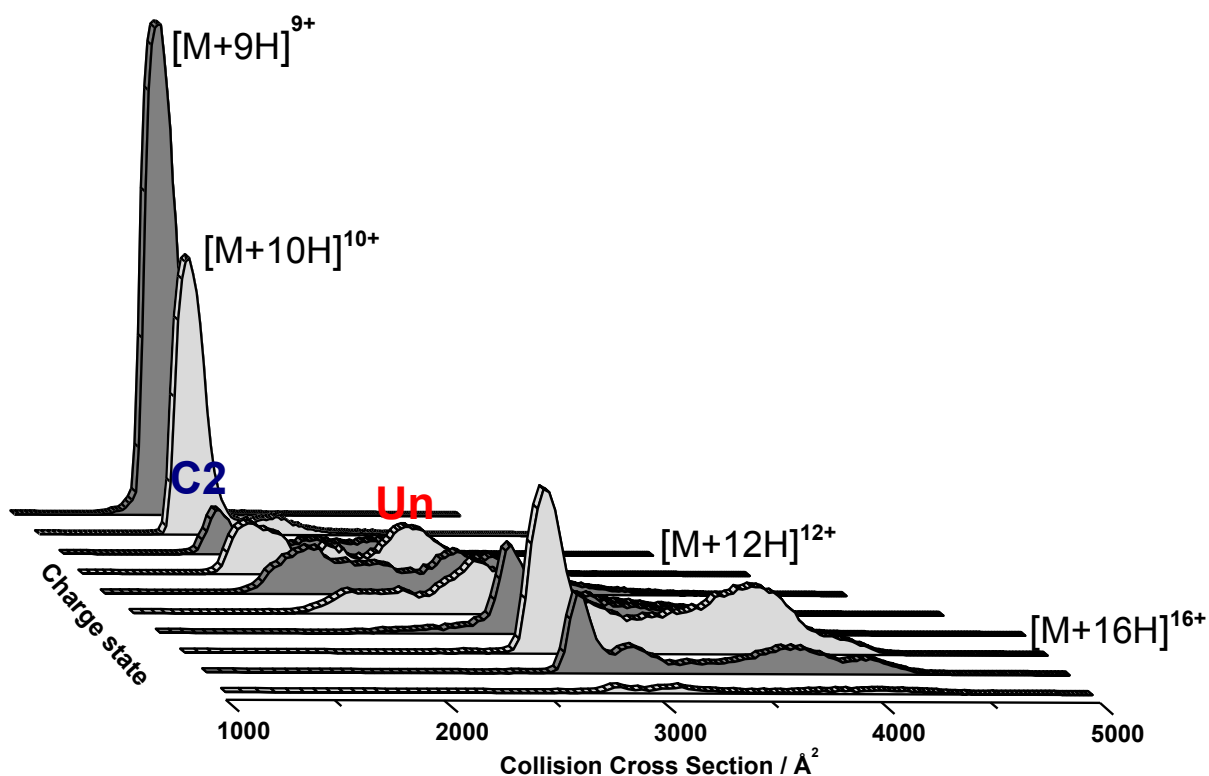


Figure 4.8: Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of A276Y mutant sprayed from buffered solution conditions (73  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via the TW IM-MS.

Remarkably, this compact form of the protein is retained over a wide range of charge states ( $9 \leq z \leq 14$ ) and its signal intensity is significantly higher than that of all previously shown mutants or WT p53 proteins. For  $10 \leq z \leq 18$ , multiple unfoldomers (Un) are observed. The abundance of the latter is spread over a very broad distribution of conformers; for instance for  $z = 17$ , at least four unfoldomers can be resolved, centered at  $2778 \text{ \AA}^2$ ,  $3049 \text{ \AA}^2$ ,  $3689 \text{ \AA}^2$ , and  $4046 \text{ \AA}^2$ . An extraordinary feature of this substitution is that both compact forms and more extended conformations are highly populated and stable over a longer mobility time, which is in marked contrast with other mutants or WT p53 proteins.

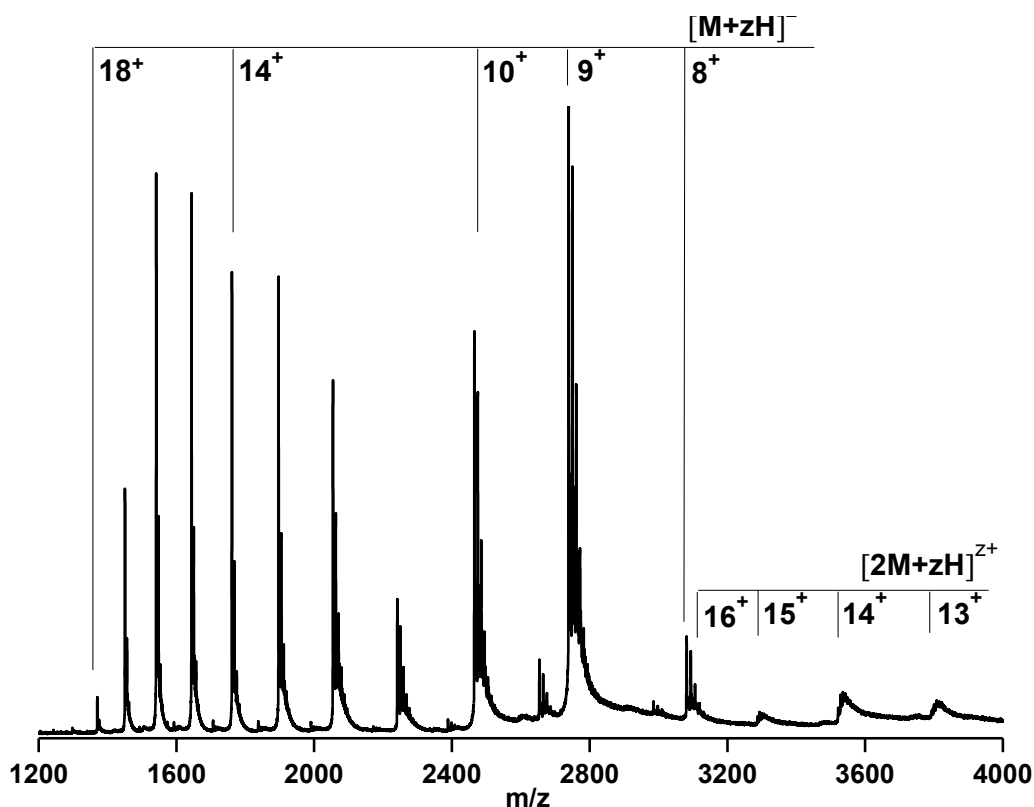


Figure 4.9: n-ESI spectrum of 73  $\mu$ M A276Y mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

#### 4.5.5. Disruptions of the Zinc-binding Region - R175H

Arg-175 is located in the loop L2 region and is the hotspot codon for one the most common oncogenic mutations, R175H. This mutant is likely to perturb the zinc-binding site (41) (Arg-175 is directly adjacent to three of the zinc ion ligands Cys-176, His-179, Cys-238), although there are no NMR or X-Ray crystallography records available for this mutant yet. The deleterious effect of this single point substitution appear to reside primarily in the fact that an introduction of a bulky histidine residue causes structural distortion and may directly interfere with zinc binding rather than in the mere loss of WT interactions (42). It results in a complete loss of function both at both physiological and sub physiological temperatures *in vitro* and *in vivo* (4,19,37). This mutation is highly destabilising, but in contrast to



the missense mutations found in the  $\beta$ -sandwich region of p53-DBD, DNA-binding activity is abrogated in the folded state (4).

The low thermodynamic stability and associated aggregation susceptibility of this mutation is revealed in the mass spectrum obtained (Figure 4.10). Mutation has a profound effect on the appearance of the spectrum. High baseline and broad peaks are a common trait of unstable and aggregated species (as well as the salt pick up). Overall the charge state distribution is qualitatively similar to WT p53-DBD, evidence that the native-like structure of the protein has not been completely lost. A charge state distribution of monomer is observed with  $9 \leq z \leq 17$  where the most abundant species correspond to  $[M+10H]^{10+}$  and  $[M+11H]^{11+}$ . The dimeric form  $[2M+zH]^{z+}$  is also significantly more pronounced for this mutant, and an ion envelope of  $14 \leq z \leq 17$  is observed. Unfortunately no mobility data was recorded for this mutant due to limited sample availability.

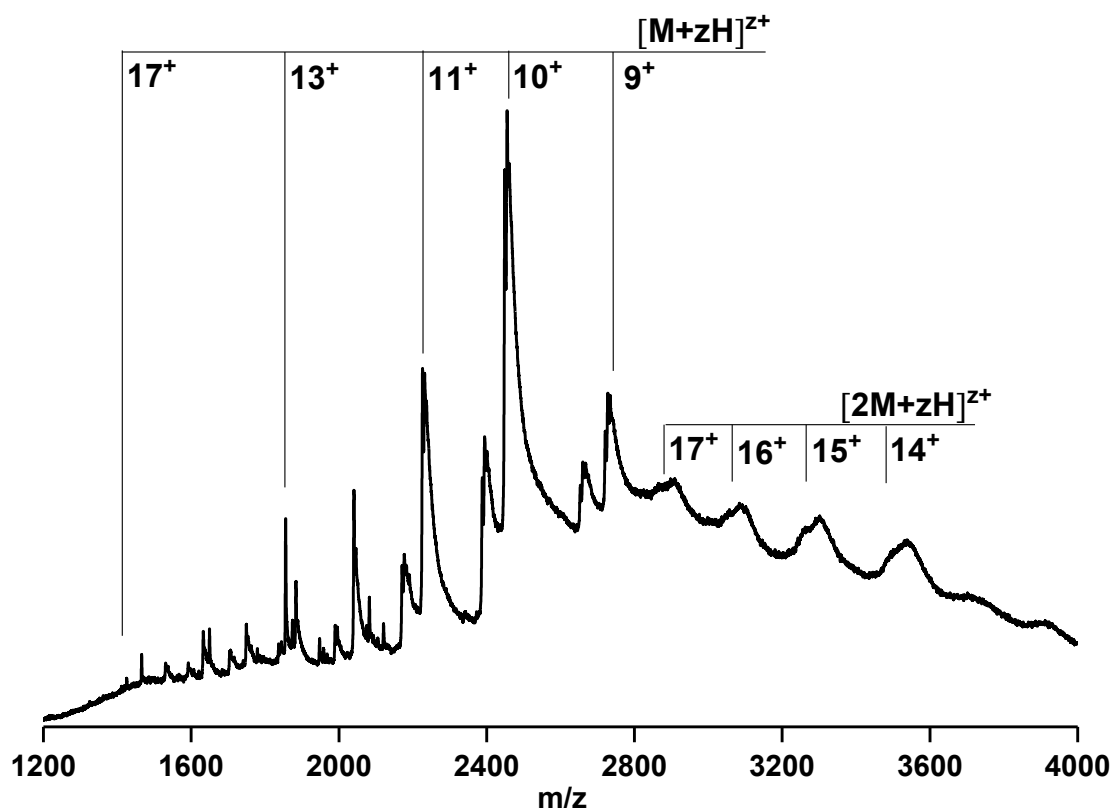


Figure 4.10: n-ESI spectrum of 5  $\mu$ M R175H mutation obtained via the DT IM-MS from buffered solution with 100 mM ammonium acetate.

## **4.6. Mechanism of Rescue of Common p53 Cancer Mutants by Second-site Suppressor Mutations**

The deleterious effects of many common tumorigenic mutations can be reversed by intragenic second-site suppressor mutations. Structure-based designs combined with yeast and mammalian cell-based screening assays have provided a number of second-site suppressors that restore function of various cancer-associated mutants (43-45). Global suppressors, N239Y and N268D, act as ‘global stability’ suppressors by increasing the overall thermodynamic stability of cancer mutants that can be positioned in different regions of p53 core domain (i.e. V143A, G245S, or even R175H) (20,43). These two mutations have also been used in Nikolova’s semi-rational design of a superstable quadruple mutant of human p53-DBD (46). Specific suppressors, a second class of second-site suppressor mutations, have been identified for the DNA-contact mutants and structural mutants with major distortions of the DNA-binding surface. S240R (44) and T284R (45) have been exploited as second-site suppressor in a double mutation approach for the classic DNA-mutants R273H and R273C. Remarkably in both cases, the introduction of an arginine residue close to the hotspot of the deleterious mutation potentially creates novel DNA contacts. The H168R mutation in combination with a second suppressor mutation, T123A or N239Y, was found to rescue a hotspot structural mutant R249S (43,44). The resultant core domain of the mutant binds the DNA response elements with WT affinity, although the thermodynamic stability of the protein is not fully repaired (15,20).

An even more detailed picture of possible ways to rescue mutant p53 will emerge from additional high resolution structures of cancer mutants in tandem with their respective second-site suppressors.

#### 4.6.1. Cold Spot Mutations used as Second-site Suppressors

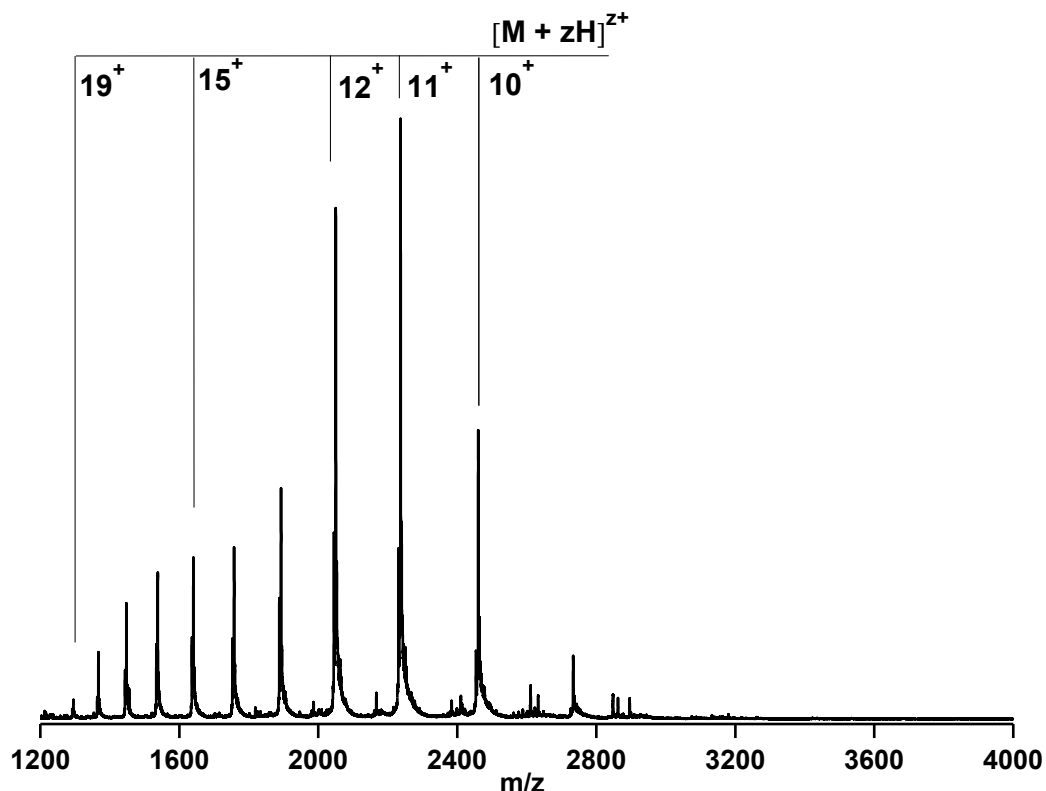
The term ‘mutational cold spot’ describes the region in the L1 loop (amino acids 112-124) in the core of p53 protein that is known for infrequent mutations in human tumours (47-50). Structurally, the L1 loop is a part of the loop-sheet-helix motif of the p53-DBD and contains only three amino acid residues that are in direct contact with DNA. Multiple studies have shown that the L1 loop is one of the most mobile and flexible regions of the core domain indicating an element of intrinsic disorder (15,16,39,42,51,52). Computational mutagenesis of L1 loop performed by Pan *et al.* (47) revealed that residues Leu-114, His-115, Ser-116 and Gly-117 make up the most flexible region of the loop. Further detailed molecular dynamics simulations which analysed human and worm forms of p53 showed that loop L1 has a pivotal role in stabilization of p53 structure (53). Several studies reported that some mutations in the L1 loop region, such as H115N, S121F and T123P display supra-wild-type DNA binding or altered binding-sequence specificity (48,50,54,55). Marabet *et al.* (13) assessed the potential role of H115N and S116M mutations as second-site suppressors on the p53 hotspot mutations in terms of their effect on thermal stability and DNA binding. The results showed, *inter alia*, that H115N is capable of rescuing the DNA-binding of one of the most common contact mutants, R248.

H115N is of a special interest here as IM-MS has been used as a tool to test whether this mutation could trigger conformational changes in a double mutant approach in three p53 cancer-associated mutations: R249A, R273H and R248Q. In addition the sequence-specific DNA-binding properties of this mutant were elucidated.

#### 4.6.2. Engineered Mutation in the L1 loop - H115N

IM-MS analysis provided a conformational fingerprint of the H115N p53 mutation, in particular a broader distribution of conformers relative to the WT p53 was observed. According to the mass spectrometry results (Figure 4.11), the protein has a similar range of charge states ( $9 \leq z \leq 19$ ) as the WT p53 protein, however the higher

charge states present with higher intensity. The most abundant peak corresponds to  $z = 11$ , higher by two charge states than for WT p53, which infers a less compact structure. The deconvoluted mass of the protein is 24591.8 Da where the reduction of mass is due to the change in amino acid residues from a histidine (137.1 Da) to an asparagine residue (114.1 Da).



**Figure 4.11:** n-ESI spectrum of 97  $\mu$ M H115N mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

In IM-MS, the conformational diversity of this engineered mutant is remarkable (Figure 4.12). For  $z = 9$  and 10, a C1-type conformer with a CCS averaging 1730  $\text{\AA}^2$  is observed, but for  $z = 11$ , the intensity is spread over a wide distribution of conformers from C1, with a CCS of 1827  $\text{\AA}^2$ , to X1 with a CCS centered on 2184  $\text{\AA}^2$ .

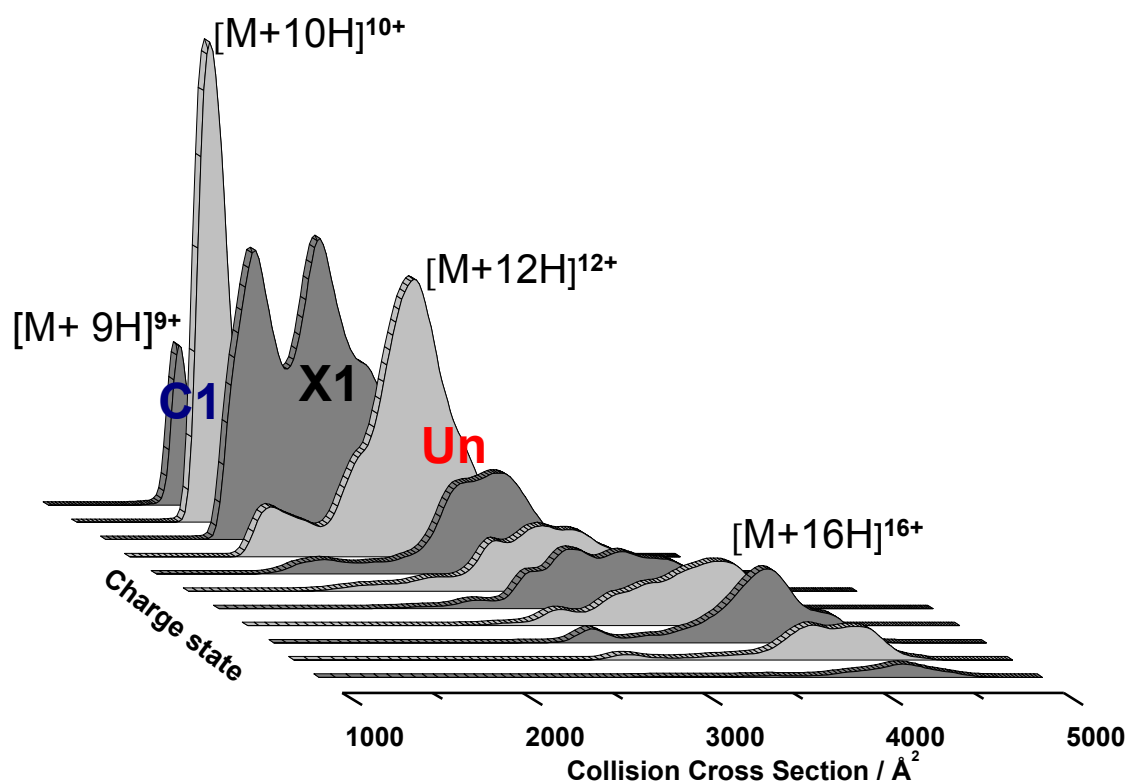
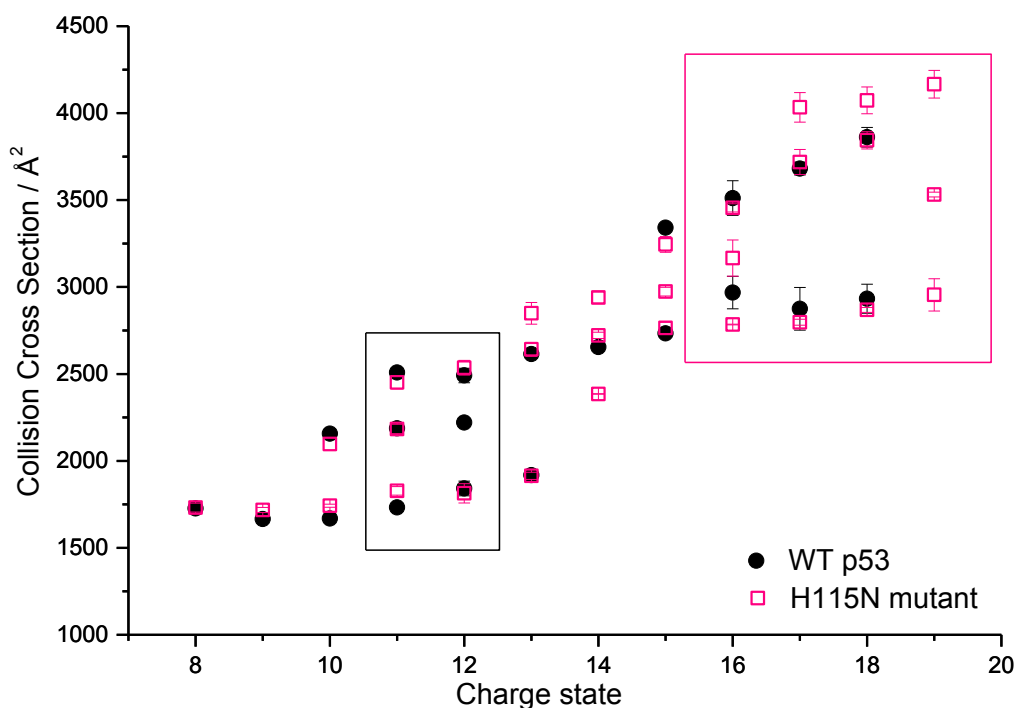


Figure 4.12: Waterfall plot representing the *in vacuo* 'conformational phenotype' of H115N mutation sprayed from buffered solution conditions (102  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via TW IM-MS.

Mass spectrometry data show that the intensity of the more extended populations is more pronounced, but IM-MS provides another dimension to this analysis, thus indicating that this mutant is intrinsically more plastic and significantly more disordered than the WT p53. This diverse conformational phenotype was also seen in the detailed far-UV-study at various temperatures, where H115N conformation was found to be different from the WT p53 conformation (13). Other results showed that the mutant H115N is stabilized intramolecularly and kept more rigid in a conformation that is more suitable for DNA binding, as hypothesized in the literature (48), potentially it might be suggestive of that the more extended conformations are more suitable for DNA-binding.

When we directly contrast the hard numbers for estimated CCS of H115N mutant with WT p53 (Figure 4.13), it can be clearly seen that the mutant favours more

extended conformations at higher charge values ( $17 \leq z \leq 19$ ). The CCS values are comparable in the more compact native-like conformations ( $8 \leq z \leq 11$ ).



**Figure 4.13: Collision cross sections versus charge of 97  $\mu$ M H115N mutation compared with 95  $\mu$ M WT p53 DBD obtained via the TW IM-MS from buffered solution conditions (50 mM ammonium acetate, containing 10% by volume propan-2-ol). The black box highlights the most intense charge states of H115N mutation.**

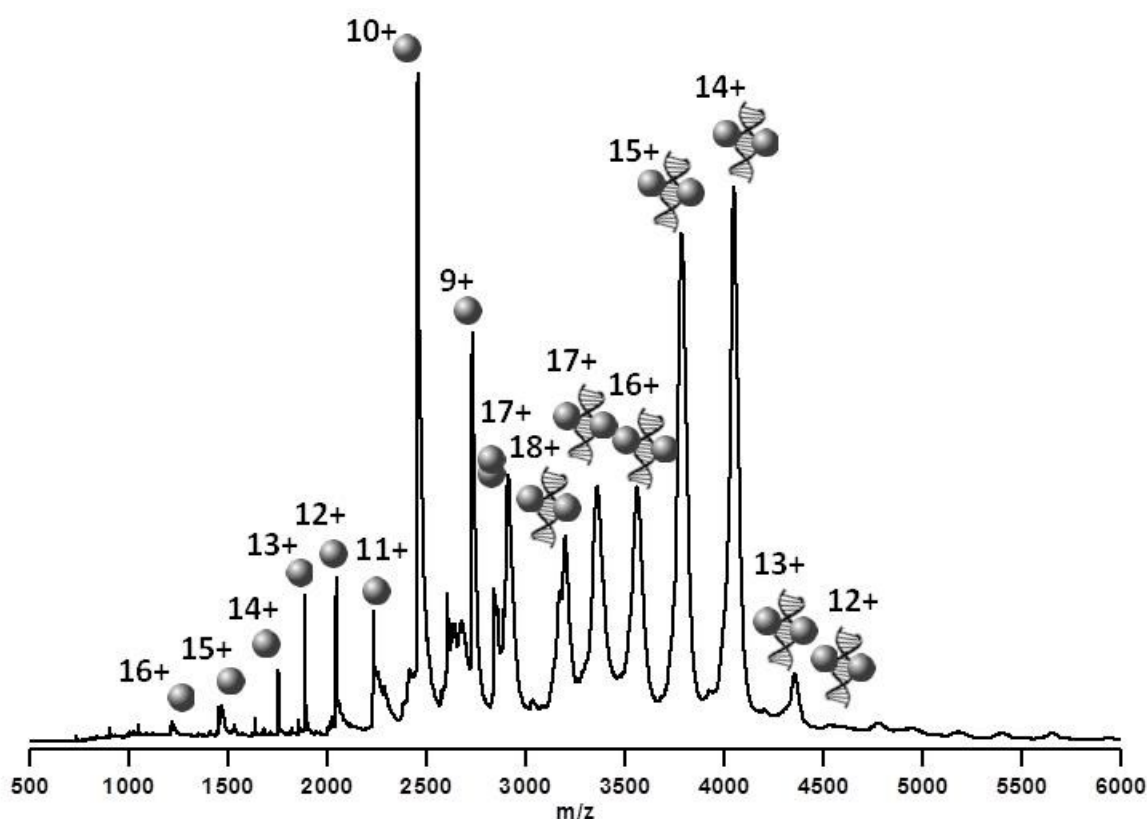
However the most abundant peak for the WT p53 is  $z = 9$  which adopts a single conformation, while the most abundant peak for the H115N mutant is  $z = 11$  which exhibits three distinct conformational isomers.

### 4.6.3. DNA-binding properties of H115N Mutant

Binding to DNA of purified H115N protein reported by Ahn *et al.* (48) was found to be superior to the WT p53 by several criteria including gel mobility shift assay, filter binding and DNase I footprinting, nonetheless H115N mutant was markedly impaired in causing apoptosis when cells were subjected to DNA damage. Remarkably H115N also displayed supra-wild-type non-specific DNA-binding. Based on the X-Ray crystallography structure of the core domain of p53, the H115

residue does not directly contact DNA (3). However, structural studies proposed that residues in loop L1 form part of the p53 dimer-dimer interface (56). Hence the H115N mutation might contribute to stronger interdomain interactions that stabilize cooperative binding of tetramer to DNA. And perhaps mutation of H115 allows for greater ease of forming the favourable conformation for DNA-binding. In contrast Merabet *et al.* (13) reported that mutants from the L1 loop (H115N and S116M) bind to the DNA with similar affinity to that of the wild-type p53.

When partnered with DNA, the H115N mutant possesses a remarkable self-organizing ability, a 2:1 protein/DNA complex is observed, thus implicating that the binding is strongly cooperative.



**Figure 4.14:** n-ESI spectrum of 97  $\mu$ M H115N mutant with a 14  $\mu$ M double-stranded 12-base-pair DNA fragment in a ratio 7:1 (protein to DNA, based on a protein monomer). Mass spectrum was obtained from a solution buffered with ammonium acetate and recorded via the DT IM-MS. The single spheres on the left represent a p53 monomer, whereas the spheres with DNA spirals correspond to the dimeric p53-DNA complex.

Here data does not support the superior DNA-binding ability of this mutant; on the contrary it shows rather decreased binding affinity. The p53/H115N-DBD/DNA complex has a wider charge-state distribution ( $z = 12 - 18$ ; Figure 4.14). The stoichiometry of DNA to protein had to be radically decreased (Figure 4.15) in order to remove the excess unbound DNA present in the mass spectrum, thus providing clear evidence of reduced DNA-binding activities for this mutation.

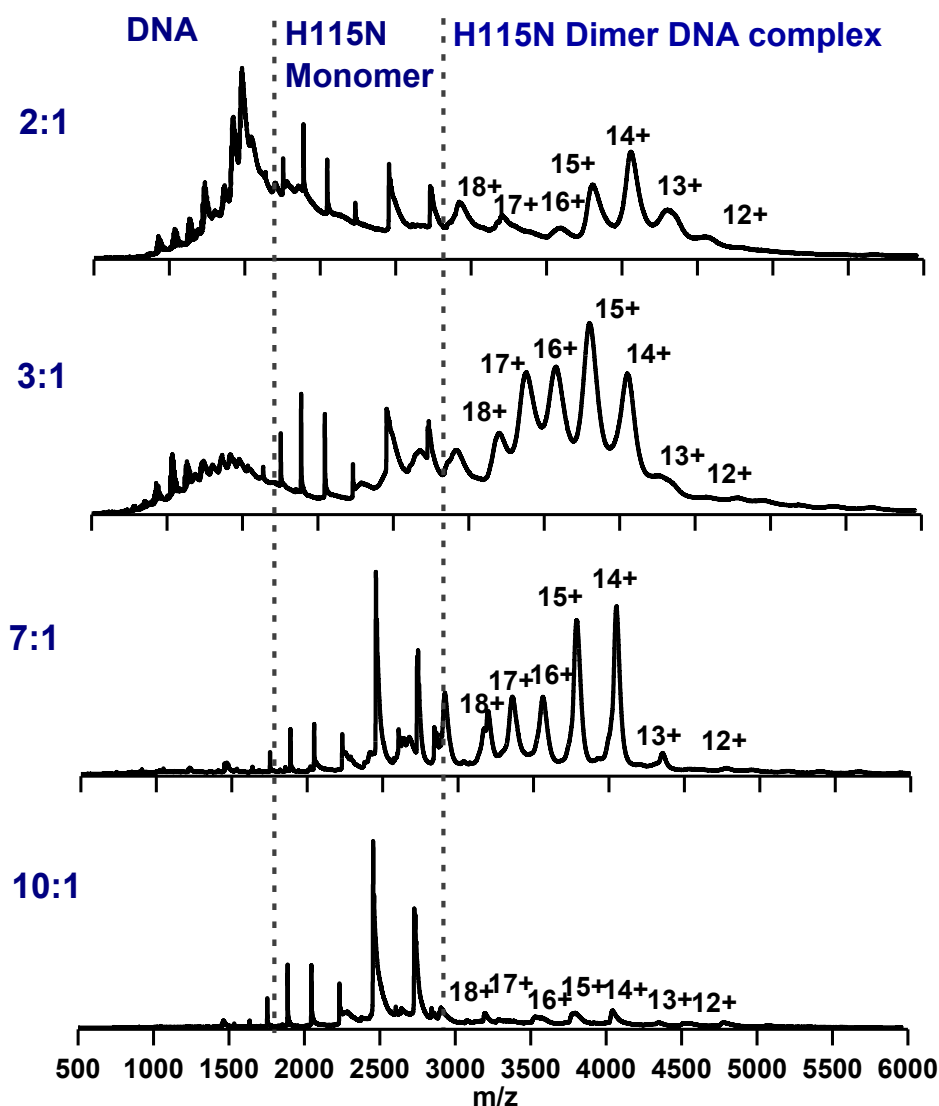


Figure 4.15: n-ESI mass spectra of H115N mutation with a double stranded 12-base-pair DNA fragment sprayed at different stoichiometric ratios (2:1, 3:1, 7:1, and 10:1; 97  $\mu\text{M}$  : 49  $\mu\text{M}$ , 97  $\mu\text{M}$  : 33  $\mu\text{M}$ , 97  $\mu\text{M}$  : 14  $\mu\text{M}$  and 97  $\mu\text{M}$  : 10  $\mu\text{M}$  ; protein to DNA, based on a protein monomer). All spectra were obtained from ammonium acetate buffered solution and recorded via the DT IM-MS.



IM-MS data for this mutation show that the 2:1 H115N DBD/DNA complex favours a single conformational state. The CCSs are centered around similar values as the corresponding charge states of the 2:1 WT p53 complex (CCS values are tabulated in Table 4.1).

Charge state	Collision Cross Section p53-DNA complex / $\text{\AA}^2$	Collision Cross Section H115N - DNA complex / $\text{\AA}^2$
13	3113	2736
14	3194	3123
15	3191	3246

**Table 4.1: Collision cross sections of the dimeric H115N-DNA complex (97  $\mu\text{M}$  H115N mutant with a 14  $\mu\text{M}$  double-stranded 12-base-pair DNA fragment in a 7:1 ratio; based on a protein monomer) versus dimeric WT p53-DNA complex (20  $\mu\text{M}$  p53 DBD with a 10  $\mu\text{M}$  double-stranded 12-base-pair DNA fragment in a ratio 2:1; protein to DNA, based on a protein monomer). Both obtained via the DT IM-MS from buffered solution conditions).**

In the n-ESI spectrum of H115N mutant with a double-stranded 12-base-pair DNA fragment (Figure 4.14), a charge state distribution for the monomeric population is also observed. The most abundant charge state is attributable to  $[\text{M}+10\text{H}]^{10+}$  ion which is one proton lower than detected in the absence of the DNA partner (Figure 4.11). Overall the lower charge states are more pronounced in the presence of DNA. There is also a drastic decrease in the abundance of ion envelope for the higher charged species, where  $11 \leq z \leq 16$ . From the MS data it can be inferred that more compact conformational family of H115N mutant is favoured in the presence of DNA ligand.

This strong preference towards more compact conformational isomers in the presence of DNA in MS is in accord with further mobility measurements. The differences in the conformational landscapes of monomeric H115N mutant observed when the DNA is available versus no DNA present are illustrated in Figure 4.16. In the absence of DNA partner, the H115N molecule assumes two distinct conformational isomers for  $z = 10$  with CCS values of  $1893 \text{ \AA}^2$  and  $2290 \text{ \AA}^2$ , but in the presence of ligand only a single conformational population was detected of a

lower CCS value of  $1678 \text{ \AA}^2$ . This is similar for the  $[M+11H]^{11+}$  ion which exists as a single conformational isomer with a lower CCS when DNA is added. In addition, for the  $z = 9$  species, a CCS of  $1763 \text{ \AA}^2$  was measured in the presence of DNA, and a CCS of  $1874 \text{ \AA}^2$  was measured without DNA, again strongly manifesting the reduction of more extended states of H115N mutant.

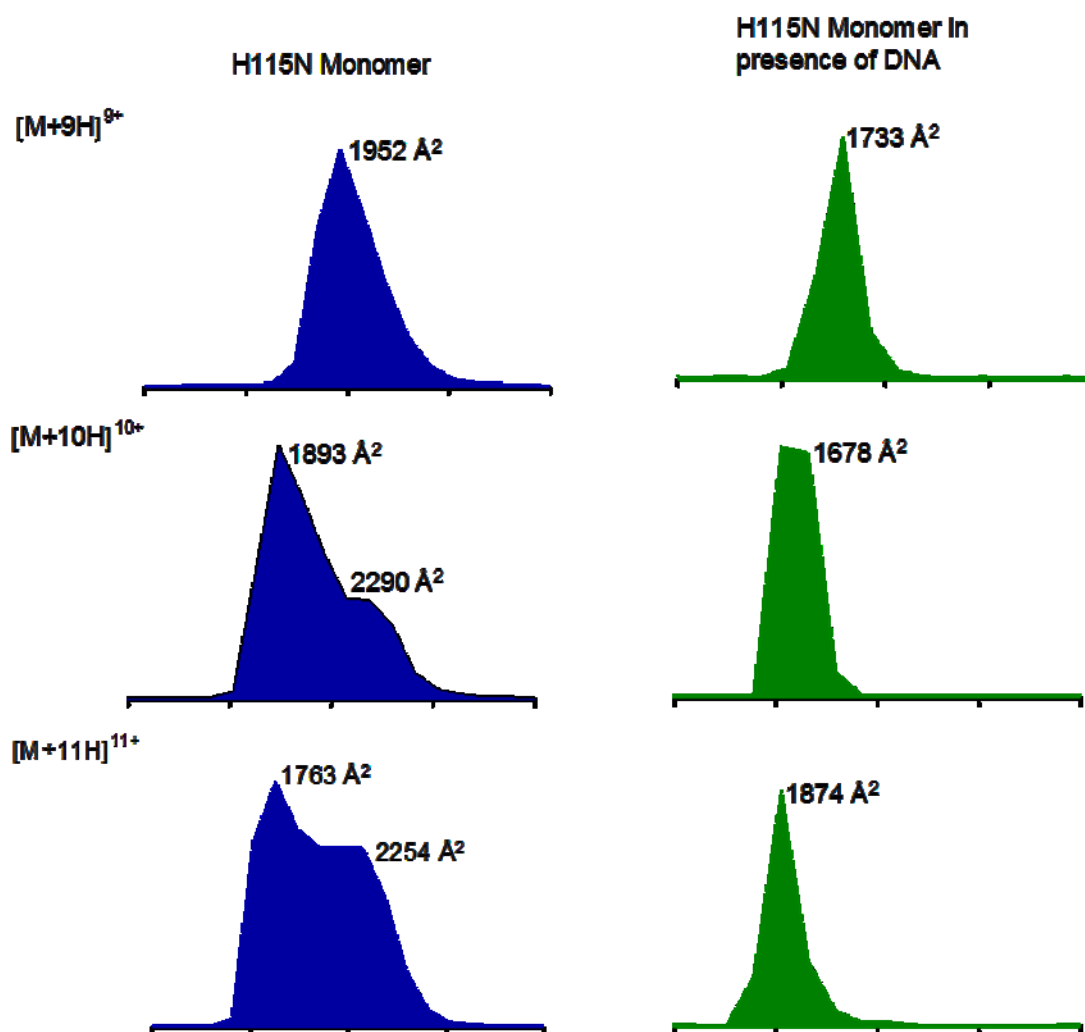


Figure 4.16: Comparison of ATDs for  $[M+9H]^{9+}$ ,  $[M+10H]^{10+}$  and  $[M+11H]^{11+}$  monomeric form of H115N without DNA versus DNA present. ATDs were obtained from  $97 \mu\text{M}$  H115N solution and from  $97 \mu\text{M}$  H115N mutant with a  $14 \mu\text{M}$  double-stranded 12-base-pair DNA fragment in a ratio 7:1 (protein to DNA, based on a protein monomer), respectively. Both were acquired via the DT IM-MS from buffered solution conditions.

The amount of conformational promiscuity and plasticity of the H115N mutant is significantly decreased in the presence of DNA. The conformational equilibrium is shifted towards more rigid and compact conformational isomers, possibly suggesting that the conformer with highest affinity for the DNA-target site is a compact one, which allows for greater ease of forming the favourable conformation for DNA-binding. Remarkably, depopulation in more extended conformations is in a direct contrast what was demonstrated for the WT p53 (refer to Chapter 2, Section 3.5.1).

The shift in the equilibrium towards the more compact conformational state (in the presence of DNA-target site) can be explained by at least three principally different possibilities. One plausible rationale is that the H115N mutant dynamically samples multiple monomeric structural isomers out of which functional states can be selected. In the presence of DNA, one of the transient conformers which enables binding with the highest affinity, would then be recruited. Addition of DNA would also shift the equilibrium of conformers, leading to higher occupancy of the active, here more compact conformer. An alternative hypothesis would be that the more extended species have formed dimeric species required for the H115N-DNA complex, leaving the residual compact conformational isomers. Thirdly, dissociation of H115N protein from DNA complex during desolvation processes cannot be ruled out.

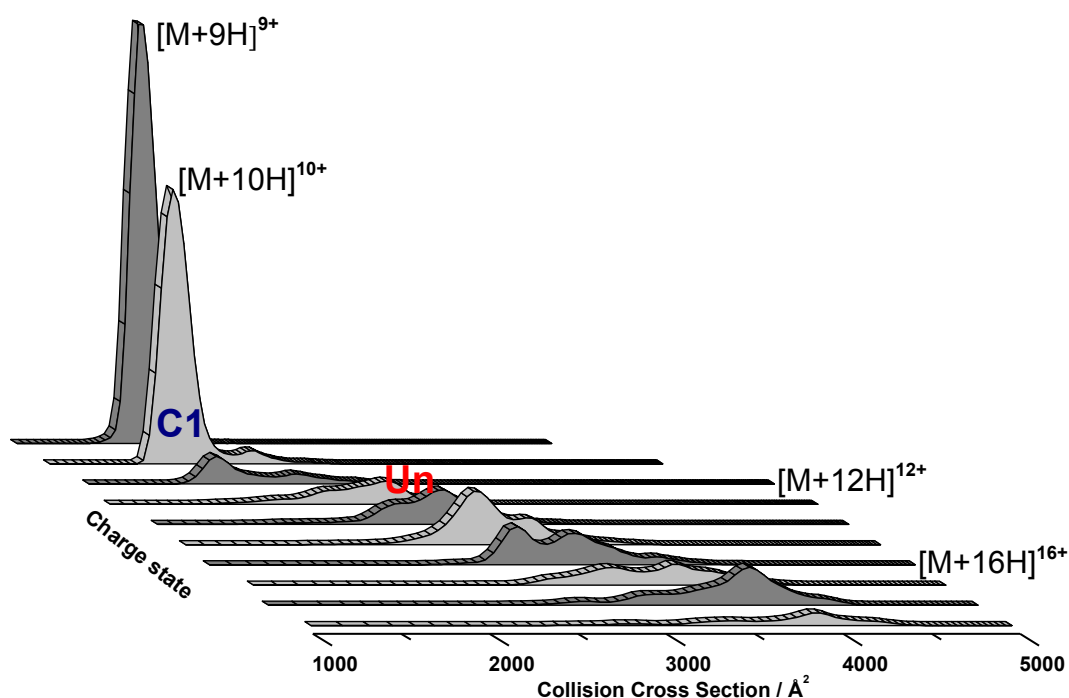
This first model poses a key element of the ‘new view’ (57,58), namely, of having a single protein oscillating between a set of structural isomers and forming of the specific p53-DNA-binding complexes through a conformational diversity.

This study exemplifies the usefulness of MS and IM-MS in exploring the dynamic landscapes of disordered proteins. The unique advantage of IM-MS over conventional ensemble measurements, where multiple conformations are often averaged and neglected, is that it is capable of segregating the heterogeneity of a complex ensemble, which is of key importance in assigning structure-function relationship to proteins that are conformationally mobile, such as IDPs. Single molecule techniques are able to capture many distinct dynamic states of protein molecules. Fluorescence techniques, such as fluorescence resonance energy transfer (FRET), have already revealed the existence of multiple conformations in

equilibrium of the full-length p53 protein (59). NMR and pre-steady-state kinetics can also provide essential details regarding protein dynamics and the binding process - specifics that cannot be obtained by crystal structure analysis alone. Such studies are, however, technically challenging and, unfortunately, unachievable in many cases.

#### 4.6.4. Double Mutant R249A/H115N

IM-MS data for the double mutant R249A/H115N are shown in Figure 4.17 (Figure 4.18 for the mass spectrum). Here the effect of the ‘second site suppressor’ H115N becomes more apparent.



**Figure 4.17:** Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of R249A/H115N double mutant sprayed from buffered solution conditions (83.2  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol).

In mobility space, the conformational occupancy of this double mutant appears highly similar to that presented by WT p53. The highly abundant low charge states show an increased width in their ATDs, which is indicative of a C1 conformer as opposed to the narrower C2 conformer, but remarkably the extended conformers Un

are noticeably more populated than in the WT p53. It appears that the double mutant presents a juxtaposition of the conformational space presented by each single mutant, providing not just a visual reference but more critically details on the conformational spread afforded by the double mutant. According to the variable temperature far-UV CD study this double mutant was classified to proteins with a conformation different from the WT p53-DBD, nonetheless the introduction of H115N has raised the thermal stability, by approximately 1.5 - 2 °C, from 36.5 to 38.5 °C (13), which might result in the conformational occupancy of this double mutant highly similar to that presented by WT p53 in the mobility profile.

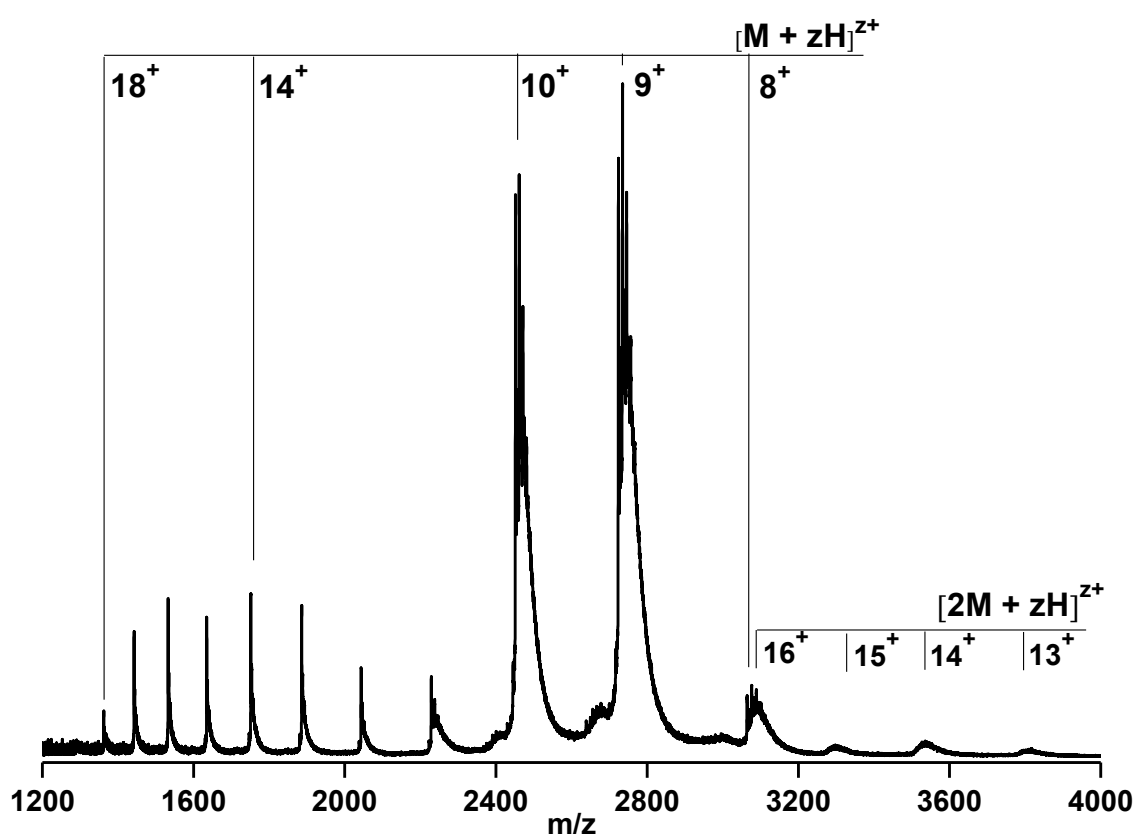


Figure 4.18: n-ESI spectrum of 83.2  $\mu\text{M}$  R249A/H115N double mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

#### 4.6.5. Double Mutant R273H/H115N

Finally the impact of H115N as a ‘second-site suppressor’ on the contact mutant R273H (Figure 4.19; Figure 4.20 for the mass spectrum) is considered. Again the conformational signature of the double mutant is highly comparable to that of the WT p53 and the plasticity of the H115N mutant has increased the population of the more extended forms of the protein. These results are certainly in accord with the far-UV CD study by Merabet *et al* (13) where the conformation of this mutant was found to be highly similar to the WT p53 as well as the overall thermal stability appeared to be greater than the cancer-associated mutation.

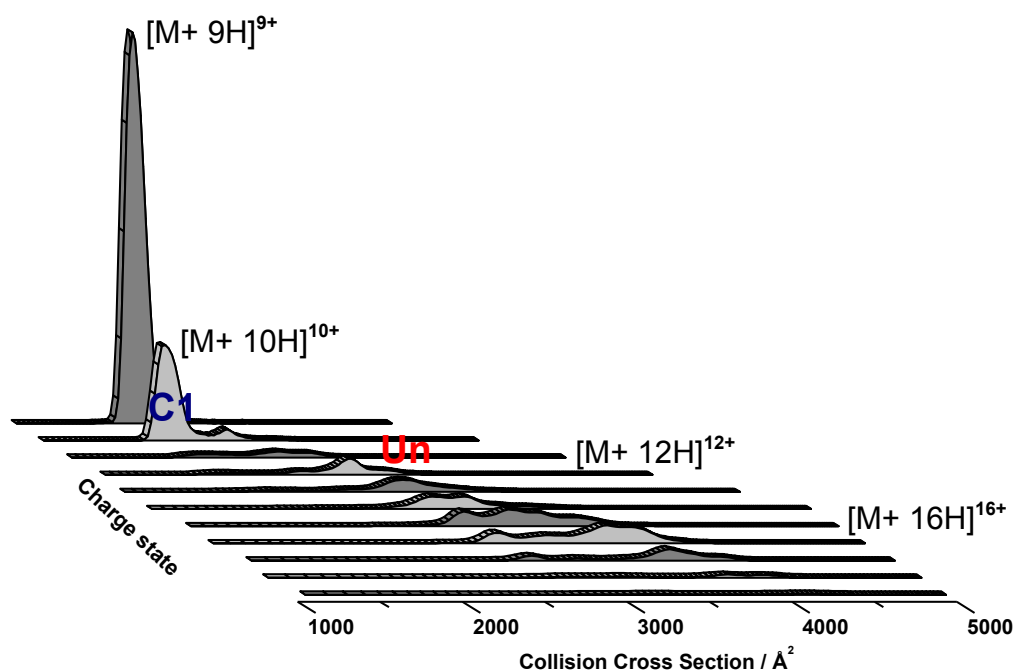


Figure 4.19: Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of R273H/H115N double mutant sprayed from buffered solution conditions (99  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol).

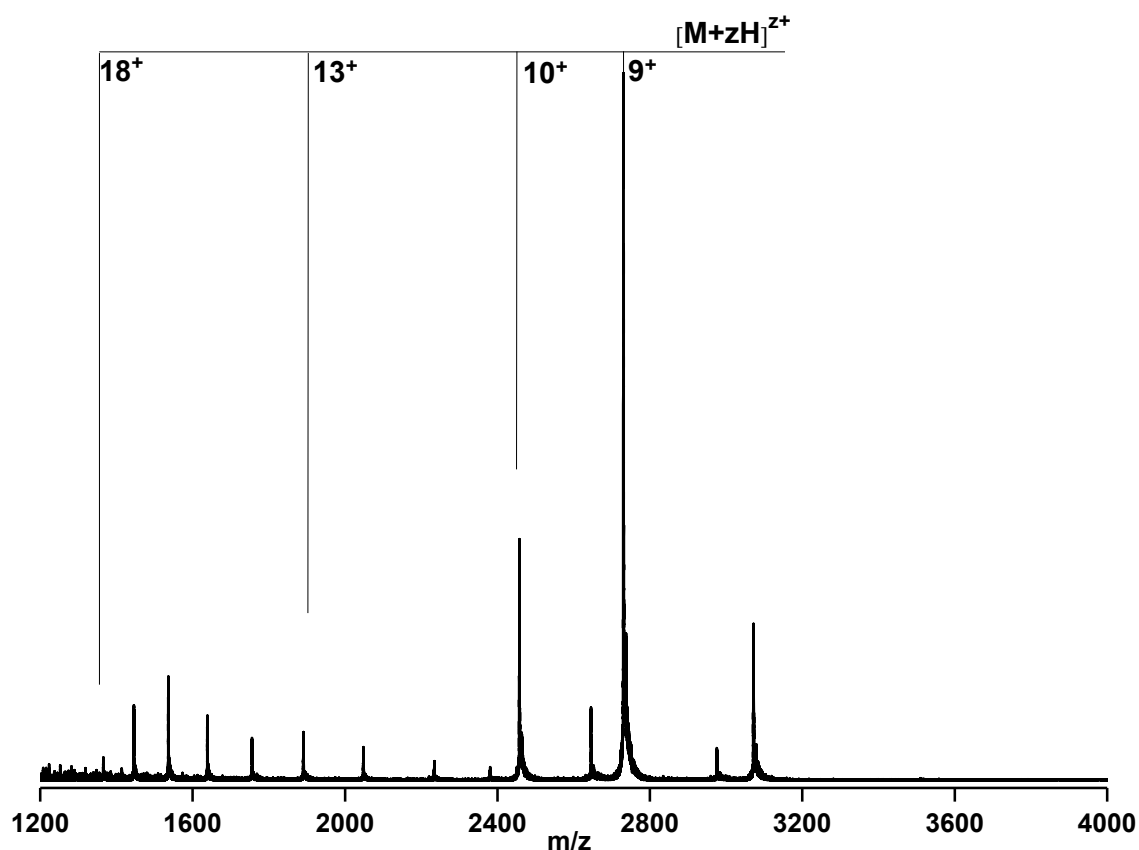
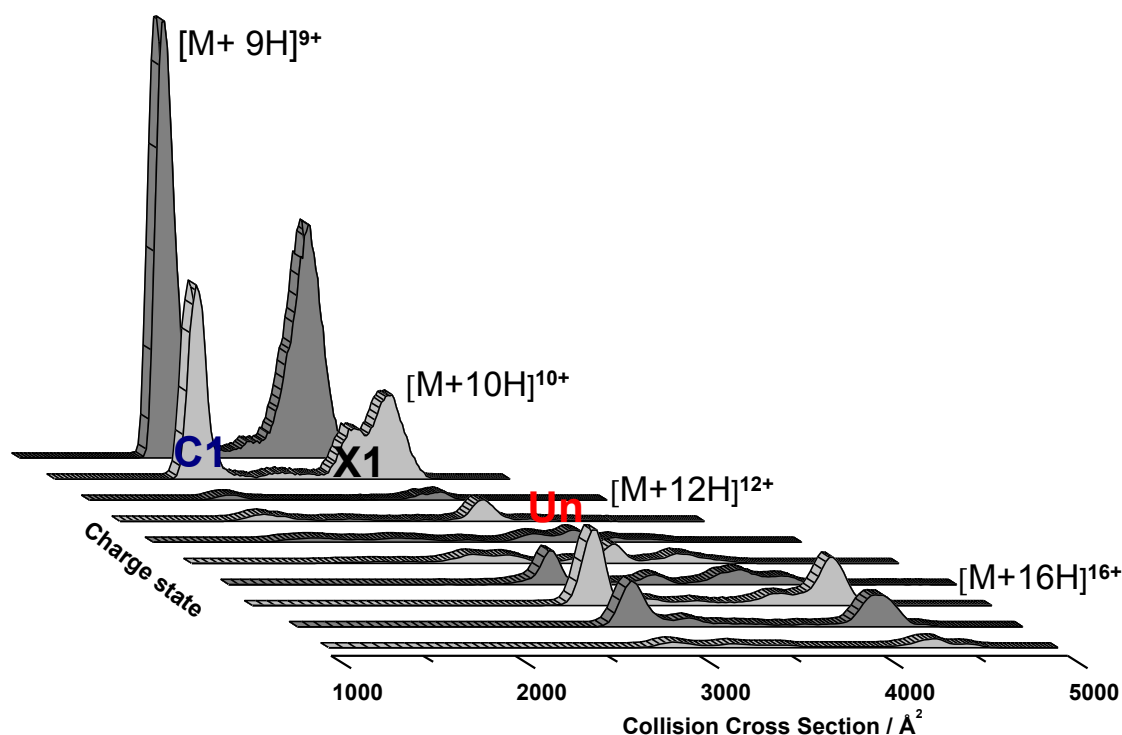


Figure 4.20: n-ESI spectrum of 99  $\mu\text{M}$  R273H/H115N mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

#### 4.6.6. Double Mutant R248Q/H115N

Data for the R248Q mutation in isolation was not obtained but the conformational landscape occupied by a double mutant R248Q/H115N was probed. Arg-248 is a hot spot residue in loop L3 for a contact mutation which introduces extensive structural changes in the loop L3 and L2 regions, affecting the DNA-binding interface Merabet *et al.* (13) reported that H115N mutation is capable of rescuing the DNA-binding in R248Q mutant, as shown by binding of R248Q/H115N to *gadd35* (the promoter of a gene involved in cell-cycle arrest).



**Figure 4.21:** Waterfall plot representing the *in vacuo* ‘conformational phenotype’ of R248Q/H115N double mutant sprayed from buffered solution conditions (83  $\mu$ M solution in 50 mM ammonium acetate, containing 10% by volume propan-2-ol), obtained via the TW IM-MS.

In addition the variable temperature far-UV CD spectra revealed the conformational signature of this double mutant appears to be similar to the WT p53 (13). Here the mobility profile is very dynamic, for every charge state the whole spectrum of conformations is observed, from a very compact (C1) to unfolded species (Un) (Figure 4.21 and Figure 4.22 for the mass spectrum). The compact population exists when  $9 \leq z \leq 13$ , although the abundance is decreased with each additional proton. Extraordinarily, the more extended conformers (X1) are retained over a wide range of charge states ( $9 \leq z \leq 18$ ) and the signal intensity is significantly higher than that of all previously shown mutants or WT p53 proteins. For  $12 \leq z \leq 18$ , multiple unfoldomers (Un) are observed in each charge state, and the number of resolvable Un conformers is at least four. Overall the mutant exhibits the most unfolded structures (the highest values of CCS) in comparison with all other mutants and WT p53-DBD. Here speculatively an increase in the more extended populations could originate from the addition of second site suppressor H115N.



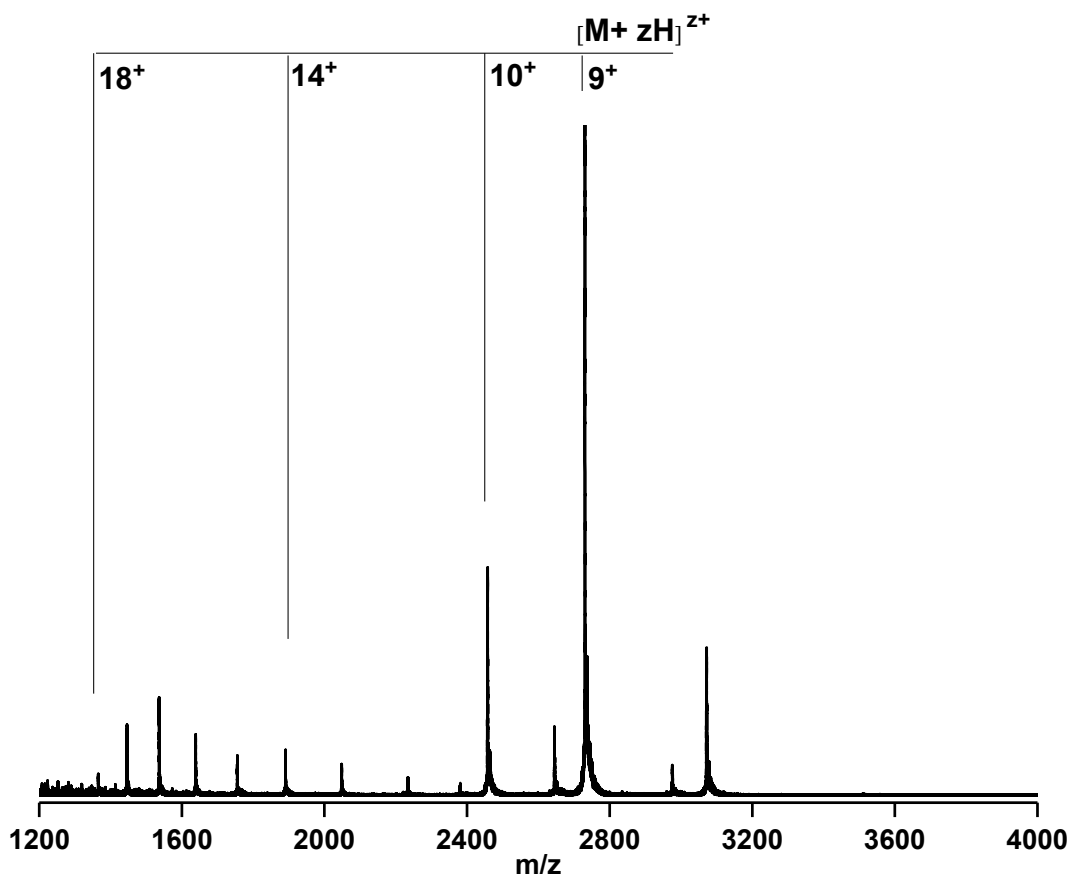


Figure 4.22: n-ESI spectrum of 83  $\mu$ M R273H/H115N mutation obtained via the TW IM-MS from buffered solution with 50 mM ammonium acetate, containing 10% by volume propan-2-ol.

## 4.7. Summary

The application of IM-MS and its capabilities for probing the conformational flexibility of p53 carcinogenic mutants has been demonstrated. The data paints an intriguingly diverse picture of the effects of common single oncogenic mutations on p53-DBD, in particular on the fraction of folded and unfolded protein. The designed cold spot mutation H115N was found to inflict a drastic change to the observed conformational fingerprint, revealing a very plastic and disordered nature of this mutation. The aftermath of introducing H115N mutant as a second-site suppressor on the hot spot mutations are demonstrated in the conformational spread afforded by the double mutants.

A technique that is able to rapidly distinguish p53 mutants at low concentrations could certainly have marked benefits for cancer screening assays and also for drug

discovery. For example, a small-molecule inhibitor that might mimic the function of the H115N mutation could be screened by using this IM-MS approach. Furthermore IM-MS could be exploited in elucidation of the relationship between p53 antineoplastic activities and its structure and functions, as well as in other medically important proteins in the human genome. The ability of this method to delineate both stoichiometry as well as the conformational spread of this tumour suppressor protein in the presence of binding partners will provide new insights to its structure-function relationships.

## 4.8. References

1. Olivier, M., Eeles, R., Hollstein, M., Khan, M.A., Harris, C.C. and Hainaut, P. (2002) The IARC TP53 database: new online mutation analysis and recommendations to users. *Human Mutation*, **19**, 607-614.
2. <http://www-p53.iarc.fr/>.
3. Cho, Y., Gorina, S., Jeffrey, P.D. and Pavletich, N.P. (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, **265**, 346-355.
4. Dearth, L.R., Qian, H., Wang, T., Baroni, T.E., Zeng, J., Chen, S.W., Yi, S.Y. and Brachmann, R.K. (2007) Inactive full-length p53 mutants lacking dominant wild-type p53 inhibition highlight loss of heterozygosity as an important aspect of p53 status in human cancers. *Carcinogenesis*, **28**, 289-298.
5. Blagosklonny, M.V. (2000) p53 from complexity to simplicity: mutant p53 stabilization, gain-of-function, and dominant-negative effect. *Faseb Journal*, **14**, 1901-1907.
6. Bargonetti, J., Reynisdottir, I., Friedman, P.N. and Prives, C. (1992) Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. *Genes & Development* **6**, 1886-1898.
7. Halazonetis, T.D. and Kandil, A.N. (1993) Conformational shifts propagate from the oligomerization domain of p53 to its tetrameric DNA binding domain and restore DNA binding to select p53 mutants. *Embo Journal* **12**, 5057-5064.
8. Nicholls, C.D., McLure, K.G., Shields, M.A. and Lee, P.W. (2002) Biogenesis of p53 involves cotranslational dimerization of monomers and posttranslational dimerization of dimers. Implications on the dominant negative effect. *Journal of Biological Chemistry*, **277**, 12937-12945.
9. Weinberg, R.L., Veprintsev, D.B. and Fersht, A.R. (2004) Cooperative binding of tetrameric p53 to DNA. *Journal of Molecular Biology*, **341**, 1145-1159.
10. Chan, W.M., Siu, W.Y., Lau, A. and Poon, R.Y.C. (2004) How many mutant p53 molecules are needed to inactivate a tetramer? *Molecular & Cellular Biology*, **24**, 3536-3551.
11. Bullock, A.N., Henckel, J., DeDecker, B.S., Johnson, C.M., Nikolova, P.V., Proctor, M.R., Lane, D.P. and Fersht, A.R. (1997) Thermodynamic stability of wild-type and mutant p53 core domain. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 14338-14342.
12. Butler, J.S. and Loh, S.N. (2006) Folding and misfolding mechanisms of the p53 DNA binding domain at physiological temperatures *Protein Science*, **15**, 2457-2465.
13. Merabet, A., Houilleberghs, H., MacLagan, K., Akanho, E., Bui, T.T., Pagano, B., Drake, A.F., Fraternali, F. and Nikolova, P.V. (2010) Mutants of the tumour suppressor p53 L1 loop as second-site suppressors for restoring DNA binding to oncogenic p53 mutations: structural and biochemical insights. *Biochemical Journal*, **427**, 225-236.

14. Pfeifer, G.P., Denissenko, M.F., Olivier, M., Tretyakova, N., Hecht, S.S. and Hainaut, P. (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*, **21**, 7435-7451.
15. Joerger, A.C., Ang, H.C., Veprintsev, D.B., Blair, C.M. and Fersht, A.R. (2005) Structures of p53 cancer mutants and mechanism of rescue by second-site suppressor mutations. *Journal of Biological Chemistry*, **280**, 16030-16037.
16. Joerger, A.C., Ang, H.C. and Fersht, A.R. (2006) Structural basis for understanding oncogenic p53 mutations and designing rescue drugs. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 15056-15061.
17. Ang, H.C., Joerger, A.C., Mayer, S. and Fersht, A.R. (2006) Effects of common cancer mutations on stability and DNA binding of full-length p53 compared with isolated core domains. *Journal of Biological Chemistry*, **281**, 21934-21941.
18. Ryan, K.M. and Vousden, K.H. (1998) Characterization of structural p53 mutants which show selective defects in apoptosis but not cell cycle arrest. *Molecular Cell Biology*, **18**, 3692-3698.
19. Bullock, A.N., Henckel, J. and Fersht, A.R. (2000) Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene*, **19**, 1245-1256.
20. Nikolova, P.V., Wong, K.-B., DeDecker, B., Henckel, J. and Fersht, A.R. (2000) Mechanism of rescue of common p53 cancer mutations by second-site suppressor mutations. *Embo Journal*, **19**, 370-378.
21. Bykov, V.J., Issaeva, N., Zache, N., Shilov, A., Hultcrantz, M., Bergman, J., Selivanova, G. and Wiman, K.G. (2005) Reactivation of mutant p53 and induction of apoptosis in human tumor cells by maleimide analogs. *Journal of Biological Chemistry*, **280**, 30384-30391.
22. Lane, D.P. and Hupp, T.R. (2003) Drug discovery and p53. *Drug Discovery Today*, **8**, 347-355.
23. Foster, B.A., Coffey, H.A., Morin, M.J. and Rastinejad, F. (1999) Pharmacological rescue of mutant p53 conformation and function. *Science*, **286**, 2507-2510.
24. Bykov, V.J.N., Selivanova, G. and Wiman, K.G. (2003) Small molecules that reactivate mutant p53. *European Journal of Cancer*, **39**, 1828-1834.
25. Wiman, K.G. (2006) Strategies for therapeutic targeting of the p53 pathway in cancer. *Cell Death & Differentiation*, **13**, 921-926.
26. Samuels-Lev, Y., O'Connor, D.J., Bergamaschi, D., Trigiante, G., Hsieh, J.K., Zhong, S., Campargue, I., Naumovski, L., Crook, T. and Lu, X. (2001) ASPP proteins specifically stimulate the apoptotic function of p53. *Molecular Cell*, **8**, 781-794.
27. Friedler, A., Hansson, L.O., Veprintsev, D.B., Freund, S.M.V., Rippin, T.M., Nikolova, P.V., Proctor, M.R., Rudiger, S. and Fersht, A.R. (2002) A peptide that binds and stabilizes p53 core domain: Chaperone strategy for rescue of oncogenic mutants. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 937-942.

28. Issaeva, N., Friedler, A., Bozko, P., Wiman, K.G., Fersht, A.R. and Selivanova, G. (2003) Rescue of mutants of the tumor suppressor p53 in cancer cells by a designed peptide. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 13303-13307.
29. Rippin, T.M., Bykov, V.J., Freund, S.M., Selivanova, G., Wiman, K.G. and Fersht, A.R. (2002) Characterization of the p53-rescue drug CP-31398 in vitro and in living cells. *Oncogene*, **21**, 2119-2129.
30. Bykov, V.J., Issaeva, N., Shilov, A., Hultcrantz, M., Pugacheva, E., Chumakov, P., Bergman, J., Wiman, K.G. and Selivanova, G. (2002) Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nature Medicine*, **8**, 282-288.
31. Cañadillas, J.M.P., Tidow, H., Freund, S.M.V., Rutherford, T.J., Ang, H.C. and Fersht, A.R. (2006) Solution structure of p53 core domain: Structural basis for its instability. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 2109-2114.
32. Patel, S., George, R., Autore, F., Fraternali, F., Ladbury, J.E. and Nikolova, P.V. (2008) Molecular interactions of ASPP1 and ASPP2 with the p53 protein family and the apoptotic promoters PUMA and Bax. *Nucleic Acids Residues*, **36**, 5139-5151.
33. Patel, S., Bui, T.T.T., Drake, A.F., Fraternali, F. and Nikolova, P.V. (2008) The p73 DNA binding domain displays enhanced stability relative to Its nomologue, the tumor suppressor p53, and exhibits cooperative DNA binding. *Biochemistry*, **47**, 3235-3244.
34. Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J.P., Slade, S.E., Thalassinios, K., Bateman, R.H., Bowers, M.T. and 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**, 1-12.
35. McCullough, B.J., Kalapothakis, J., Eastwood, H., Kemper, P., MacMillan, D., Taylor, K., Dorin, J. and Barran, P.E. (2008) Development of an ion mobility quadrupole time of flight mass spectrometer. *Analytical Chemistry*, **80**, 6336-6344.
36. Wong, K.B., DeDecker, B.S., Freund, S.M., Proctor, M.R., Bycroft, M. and Fersht, A.R. (1999) Hot-spot mutants of p53 core domain evince characteristic local structural changes. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 8438-8442.
37. Kato, S., Han, S.Y., Liu, W., Otsuka, K., Shibata, H., Kanamaru, R. and Ishioka, C. (2003) Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 8424-8429.
38. Lee, E.-W., Lee, M.-S., Camus, S., Ghim, J., Yang, M.-R., Oh, W., Ha, N.-C., Lane, D.P. and Song, J. (2009) Differential regulation of p53 and p21 by MKRN1 E3 ligase controls cell cycle arrest and apoptosis. *Embo Journal*, **28**, 2100-2113.

39. Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E. and Shakked, Z. (2006) Structural basis of DNA recognition by p53 tetramers. *Molecular Cell*, **22**, 741-753.
40. Kubbutat, M.H.G., Jones, S.N. and Vousden, K.H. (1997) Regulation of p53 stability by Mdm2. *Nature* **387**, 299 - 303.
41. Tidow, H., Veprintsev, D.B., Freund, S.M. and Fersht, A.R. (2006) Effects of oncogenic mutations and DNA response elements on the binding of p53 to p53-binding protein 2 (53BP2). *Journal of Biological Chemistry*, **281**, 32526-32533.
42. Joerger, A.C. and Fersht, A.R. (2007) Structure-function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene*, **26**, 2226-2242.
43. Brachmann, R.K., Yu, K.X., Eby, Y., Pavletich, N.P. and Boeke, J.D. (1998) Genetic selection of intragenic suppressor mutations that reverse the effect of common p53 cancer mutations. *Embo Journal*, **17**, 1847-1859.
44. Baroni, T.E., Wang, T., Qian, H., Dearth, L.R., Truong, L.N., Zeng, J., Denes, A.E., Chen, S.W. and Brachmann, R.K. (2004) A global suppressor motif for p53 cancer mutants. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 4930-4935.
45. Wiczorek, A.M., Waterman, J.L.F., Waterman, M.J.F. and Halazonetis, T.D. (1996) Structure-based rescue of common tumor-derived p53 mutants. *Nat. Med.*, **2**, 1143-1146.
46. Nikolova, P.V., Henckel, J., Lane, D.P. and Fersht, A.R. (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14675-14680.
47. Pan, Y., Ma, B., Venkataraghavan, R.B., Levine, A.J. and Nussinov, R. (2005) In the quest for stable rescuing mutants of p53: computational mutagenesis of flexible loop L1. *Biochemistry*, **44**, 1423-1432.
48. Ahn, J., Poyurovsky, M.V., Baptiste, N., Beckerman, R., Cain, C., Mattia, M., McKinney, K., Zhou, J., Zupnick, A., Gottifredi, V. *et al.* (2009) Dissection of the sequence-specific DNA binding and exonuclease activities reveals a superactive yet apoptotically impaired mutant p53 protein. *Cell Cycle*, **8**, 1603-1615.
49. Inga, A. and Resnick, M.A. (2001) Novel human p53 mutations that are toxic to yeast can enhance transactivation of specific promoters and reactivate tumor p53 mutants. *Oncogene*, **20**, 3409-3419.
50. Zupnick, A. and Prives, C. (2006) Mutational analysis of the p53 core domain L1 loop. *Journal of Biological Chemistry*, **281**, 20464-20473.
51. Joerger, A.C., Allen, M.D. and Fersht, A.R. (2004) Crystal structure of a superstable mutant of human p53 core domain. Insights into the mechanism of rescuing oncogenic mutations. *Journal of Biological Chemistry*, **279**, 1291-1296.
52. Baker, E.S., Hong, J.W., Gaylord, B.S., Bazan, G.C. and Bowers, M.T. (2006) PNA/dsDNA complexes: Site specific binding and dsDNA biosensor applications. *Journal of the American Chemical Society*, **128**, 8484-8492.

53. Pan, Y., Ma, B., Levine, A.J. and Nussinov, R. (2006) Comparison of the human and worm p53 structures suggests a way for enhancing stability. *Biochemistry*, **45**, 3925-3933.
54. Inga, A., Monti, P., Fronza, G., Darden, T. and Resnick, M.A. (2001) p53 mutants exhibiting enhanced transcriptional activation and altered promoter selectivity are revealed using a sensitive, yeast-based functional assay. *Oncogene*, **20**, 501-513.
55. Saller, E., Tom, E., Brunori, M., Otter, M., Estreicher, A., Mack, D.H. and Iggo, R. (1999) Increased apoptosis induction by 121F mutant p53. *Embo Journal*, **18**, 4424-4437.
56. Klein, C., Planker, E., Diercks, T., Kessler, H., Kunkele, K.P., Lang, K., Hansen, S. and Schwaiger, M. (2001) NMR spectroscopy reveals the solution dimerization interface of p53 core domains bound to their consensus DNA. *Journal of Biological Chemistry*, **276**, 49020-49027.
57. James, L.C. and Tawfik, D.S. (2003) Conformational diversity and protein evolution – a 60-year-old hypothesis revisited. *Trends in Biochemical Sciences*, **28**, 361-368.
58. Frauenfelder, H., Sligar, S.G. and Wolynes, P.G. (1991) The energy landscapes and motions of proteins. *Science*, **254**, 1598-1603.
59. Huang, F., Rajagopalan, S., Settanni, G., Marsh, R.J., Armoogum, D.A., Nicolaou, N., Bain, A.J., Lerner, E., Haas, E., Ying, L. *et al.* (2009) Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 20758-20763.

# 5

## Promiscuous Interactions of p53

*Many structural facets of the function of p53 remain to be unravelled. Of particular interest is the mechanism behind the dynamics of disorder-to-order transitions in intrinsically disordered regions to explain the structural basis of binding promiscuity. Structural studies on p53 bound to domains of signaling proteins and partner proteins in the cell cycle and transcriptional pathways are critical to obtain a detailed picture of the structural organization of higher-order complexes. Herein the conformational diversity of three functional regions of p53 was interrogated: natively unfolded N-terminal region, superstable quadruple mutant of DNA-binding domain and p53-N-DBD construct, consisting of N-terminus and the core domain. The molecular interaction of p53 N-terminus with the oncoprotein murine double minute 2, as well as with the antiapoptotic factor B-cell lymphoma-extra-large is probed.*



## 5. Promiscuous Interactions of p53

The powerful tumour suppressor protein p53 induces or suppresses the expression of an array of target genes engaged in cell cycle control, senescence, and apoptosis in response to oncogenic or other cellular stress signals. It exercises its function as guardian of genome through a convoluted interplay of independently folded and intrinsically disordered functional domains. Furthermore the cellular protein-protein interactions play a major role in fine tuning p53 activity at different levels. All steps of this regulatory mechanism are controlled by interactions with positive or negative modulators, promoting p53 covalent modifications, governing its stability, subcellular localization, determining its specificity for selected promoters, or modulating its transactivation potential (1). Accordingly, p53 is a highly connected protein and can form physical complexes with many cellular proteins. Understanding of the greater picture of p53 biology will not be complete until these interacting proteins are fully delineated. Moreover, some of these interactions are potential therapeutic targets and detailing the basis of interaction will aid in the design of molecules as effective drugs.

In this chapter, insights into the miscellaneous structure of p53 are provided by segmenting p53 into individual & multi-domain constructs (shown in Figure 5.1),



**Figure 5.1: Domain structure of full-length p53 consisting of an N-terminal domain (NTD), followed by the central DNA-binding domain (DBD), the oligomerization domain (residues 318 - 360) and the extreme C-terminus (residues 360 - 393); Recombinant constructs used herein: NTD (residues 1 - 102), DBD (residues 102 - 312) and NT-DBD (residues 1 - 292).**

as well as interrogating interactions with the most important p53 regulator the E3 ubiquitin ligase Mdm2 and the Bcl-xL which is involved in transcription-independent apoptotic role of p53.

## 5.1. Methodology

### 5.1.1. Protein Expression and Purification

Protein domain constructs p53-N, superstable mutant of p53-DBD, p53-N-DBD, and Bcl-xL $\Delta$ LAC were expressed and purified by Arielle Viacava Follis (St Jude Children's Research Hospital, Memphis, USA) according to the method described here briefly. Constructs containing p53-DBD (superstable p53-DBD, p53-N-DBD) were modified at four residues - M133L/V203A/N239Y/N268D. Each cloned plasmid was transferred into *Escherichia coli* BL21 (DE3) expression strain. Cells were grown at 37°C until their optical density at 600 nm ( $OD_{600}$ ) reached a value of 0.6, then induced with 0.5 mM isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG) for 5 hours and harvested. For expression of superstable mutant of p53-DBD and p53-N-DBD, the temperature of the media was reduced to 20°C prior to IPTG induction. Bcl-xL $\Delta$ LAC construct was lysed using osmotic shock in a buffer containing 25 mM Tris, 500 mM NaCl, 5 mM imidazole, 20% w/v sucrose, 1 mg/mL lysozyme, pH 8. Purification was carried out in Tris buffer with either imidazole or NaCl gradient and pH 8.0 or 7.0 for the Ni-affinity or anion exchange step respectively. The N-terminal domain of p53 was lysed with a microfluidizer in 25 mM Tris, 500 mM NaCl, 5 mM imidazole, pH 8.0. Cell pellets containing p53-DBD were lysed sonication in 50 mM potassium phosphate, 300 mM NaCl, 2 mM TCEP, 1 mg/mL lysozyme, pH 8.0, followed by Ni-affinity purification in 25 mM Hepes, 0.5 M NaCl, 2 mM TCEP, pH 7.0 with an imidazole gradient. Protein sequences are shown below:

p53-N (residues 1-105)

```
GSHMEEPQSD PSVEPPLSQE TFSDLWKLIP ENNVLSPLPS QAMDDLMLSP
DDIEQWFTED PGPDEAPRMP EAAPPVAPAP AAPTAAAPAP APSWPLSSSV
PSQKT
```

**Superstable p53-DBD (residues 102-312)**

GSHMKTYQGS YGFR LGFLHS GTAKSVTCTY SPALNKLFCQ LAKTCPVQLW  
 VDSTPPPGTR VRAMAIYKQS QHMTFVVRRC PHHERCSDSD GLAPPQHLIR  
 VEGNLRAEYL DDRNTFRHSV VVPYEPPEVG SDCTTIHYN MCNSSCMGGM  
 NRRPILTIIT LEDSSGNLLG RDSFEVRVCA CPGRDRRTEE ENLRKKGEPH  
 HELPPGSTKR ALPNNT

**p53-N-DBD (residues 1-292)**

GSHMEEPQSD PSVEPPLSQE TFS DLWKLLP ENNVLSPLPS QAMDDLMLSP  
 DDIEQWFTED PGPDEAPRMP EAAPPVAPAP AAPTAAAPAP APSWPLSSSV  
 PSQKTYQGSY GFR LGFLHSG TAKSVTCTYS PALNKLFCQL AKTCPVQLWV  
 DSTPPPGTRV RAMAIYKQSQ HMTFVVRRC PHHERCSDSDG LAPPQHLIRV  
 EGNLRAEYLD DRNTFRHSV VVPYEPPEVGS DCTTIHYNM CNSSCMGGMN  
 RRPILTIITL EDSSGNLLGR DSFEVRVCAC PGRDRRTEEE NLRKK

**Bcl-xL $\Delta$ LAC (residues 1-209, lacking the loop region residues 44-85 and twenty two residues at the C-terminus)**

MSMAMSQSNR ELVVDFLSYK LSQKGYWSWQ FSDVEENRTE APEGTESEAV  
 KQALREAGDE FELRYRAFS DLTSQLHITP GTAYQSFEQV VNELFRDGVN  
 WGRIVAFFSF GGALCVESVD KEMQVLVSRI AAWMATYLND HLEPWIQENG  
 GWDTFVELYG NNAAAESRKG QERLEHHHHH H

Human wild-type Mdm2-N was expressed and purified by Dr Erin Worrall (University of Edinburgh, UK). The amino acid sequence of the N-terminal Mdm2 is presented below:

**Mdm2-N (residues 1-126)**

MCNTNMSVPT DGAVTTSQIP ASEQETLVRP KPLLLKLLKS VGAQKDTYTM  
 KEVLFYLGQY IMTKRLYDEK QQHIVYCSND LLGDLFGVPS FSVKEHRKIY  
 TMIYRNLVVV NQQESSDSGT SVSENR

**5.1.2. Sample Preparation**

Protein samples were flash frozen in liquid nitrogen and stored in a - 80°C freezer. Prior to mass spectrometry and ion mobility mass spectrometry measurements, samples were thawed and dialysed using BIO-RAD micro bio-spin chromatography columns (Bio-Rad Laboratories, Inc). Buffer exchange was carried out to 99.99 % effectiveness by bio-spin chromatography column washing with four 500  $\mu$ L aliquots of 50 mM ammonium acetate, followed each time by centrifugation for 2 minutes. Each of the protein samples was added to the buffer-exchanged chromatography

column and centrifuged for further 4 minutes. Protein concentrations were confirmed post dialysis spectrophotometrically (NanoDrop Spectrophotometer ND 1000 Thermo Scientific, USA).

### **5.1.3. Mass Spectrometry and Ion Mobility Mass Spectrometry**

Drift tube ion mobility mass spectrometry (DT IM-MS) experiments were performed on an in-house modified quadrupole time-of-flight mass spectrometer, the MoQToF (2). More detailed description of this instrument can be found in Chapter 2, Section 2.4.1. Ions were produced by positive n-ESI using Z-spray source, within a spray voltage range 1.0 to 1.8 kV and a source temperature of 80°C. The drift cell was filled with helium buffer gas at the average pressure of 3.5 - 4.0 Torr (4.7 - 5.3 mBar) and a temperature between 298 - 302 K. The electric potential across the cell was varied from 60 to 15 V and ATDs were recorded at eight drift voltages. The mobility of the ion of interest was calculated from a linear plot of average arrival time versus pressure/drift voltage ( $P/V$ ) and from this the rotationally-averaged collision cross section (CCS) for each resolvable species at a given charge state were obtained.

## **5.2. Highly Disordered p53-N-Terminal Region**

The N-terminus of p53 consists of an acidic transcription activation domain (TAD) and a proline-rich domain. The TAD (amino acid residues 1-63) is often further subdivided into ill-defined subdomains TAD1 (residues 1-42) and TAD2 (residues 43-63) (3,4). The adjacent proline-rich domain (residues 64-92) contains five repeats of the sequence PXXP (SH3- domain binding motifs) which contributes to the apoptotic function of p53 (3-5). The polyproline regions have conformationally-constrained backbones indicating a certain degree of structural rigidity. It is possible that the prevalence of prolines may facilitate its folding. Such role has not yet been experimentally validated for p53, but folding of other proteins has been shown to be chaperoned by the proline-rich regions (6,7). The TAD is a promiscuous binding site for a multitude of interacting proteins, such as modules of the transcriptional

machinery (via BOX-I region) (8-10), negative regulators Mdm2/Mdm4 which trigger ubiquitination and ultimately degradation of p53 (11-13), and the transcriptional coactivators p300/CBP (14,15).

The N-terminal region of p53 is intrinsically disordered (16,17), although residual secondary structure is observed in functionally essential, conserved hydrophobic residues (18). Upon binding of the p53 TAD to partner proteins, regions of nascent helical structure that are present in the native state, rigidify and become entirely folded. A fragment consisting of residues 15-29 (TAD1), assumes an  $\alpha$ -helical structure on binding to the N-terminal domain of Mdm2 and Mdm4. While natively unfolded segment of TAD2 folds into amphipathic  $\alpha$ -helices on binding to replication protein A and the Tfb1 subunit of yeast TFIIH. So far, the only known atomic high resolution structures of N-terminal p53 have been solved in complex with their binding partners, when the p53 fragment forms a helical conformation (10,11,19-23). Intrinsic disorder provides a platform for multiple binding of diverse target proteins with overlapping binding sites on TAD and, often proteins with antagonistic biological effects. The classic sequence-function model is irrelevant here. It is estimated that approximately 25% of mammalian proteins are intrinsically disordered and the presence of natively unfolded regions is a recurring motif in proteins that are at the hub of a plethora of signalling pathways (24,25).

The N-terminus of p53 is also heavily phosphorylated; it has nine phosphorylation sites, seven of which are located within the binding segment of regulatory proteins (Ser-15, Thr-18, Ser-20, Ser-33, Ser-37, Ser-46, and Thr-55) (26). Phosphorylation at position threonine 18 considerably lessens the affinity of TAD for Mdm2 because of electrostatic repulsion between the negatively charged area on the Mdm2 surface and the phosphate group (12,27,28). In contrast, there is evidence suggesting that phosphorylation at Thr-18 and other sites strengthens p53/p300 binding (14,29,30). Multiple phosphorylation of the TAD is able to alter the relative binding affinity of p53 for competing binding proteins (for example Mdm2 versus Taz2 domain of p300), highlighting how a phosphorylation cascade can provide modulation of p53 activity in response to stress signals (14).

### 5.2.1. MS of p53-N-Terminus in buffered solution conditions

From the buffered solution, a charge state distribution of p53-N monomer (calculated  $M_R = 11162.3$  Da, observed  $M_R = 11164.2$  Da) is observed with  $6 \leq z \leq 11$  (Figure 5.2). The most abundant species correspond to three charge states  $[M+6H]^{6+}$ ,  $[M+7H]^{7+}$  and  $[M+8H]^{8+}$  indicative of a compact form of the protein with a limited number of charge-accessible residues available for protonation.

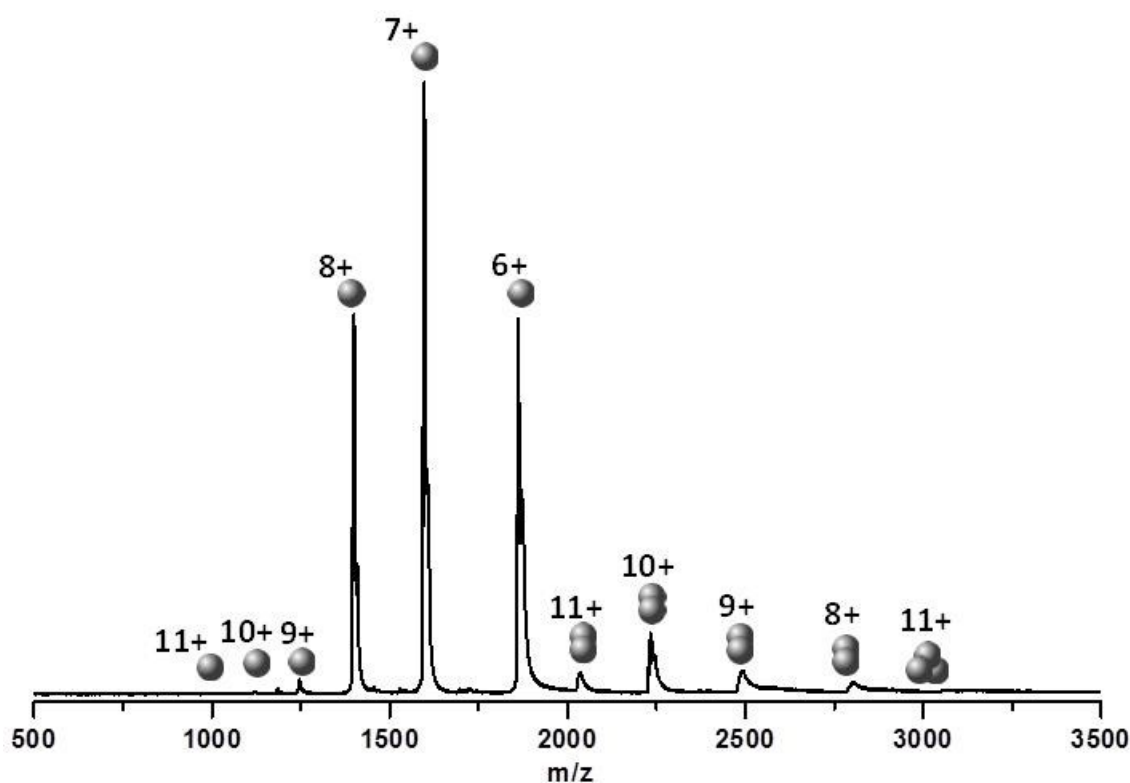


Figure 5.2: n-ESI spectrum of 200  $\mu$ M p53-N obtained via the DT IM-MS from buffered solution with 50 mM ammonium acetate.

The spectrum also displays a less intense populations for dimeric species  $[2M+zH]^{z+}$  and one unique trimer  $[3M+11H]^{11+}$ , representing higher order multimeric association in solution.

### 5.2.2. Conformational Flexibility of p53-N-Terminal Region

Along with mass spectrometry, ion mobility measurements have been performed on p53-N-terminus sprayed from buffered solution conditions. Collision cross sections obtained are tabulated in Table 5.1 and 5.2 and plotted as a function of charge state in Figure 5.3. The data shows a smooth, near linear increase in cross-section with respect to charge to circa  $1500 \text{ \AA}^2$  at  $z = 8$ , before the CCS values reach a plateau. At a given charge state two stable conformational isomers were resolved where their CCS varied by 10%, evidence of conformational diversity of the N-terminal region of p53. ATDs for the most intense charge states  $[M+6H]^{6+}$ ,  $[M+7H]^{7+}$  and  $[M+8H]^{8+}$ , displaying the abundance of the conformation with the protein are shown in Figure 5.4. For  $z = 7-8$  the larger conformer is highly abundant, whereas for the lowest observable charge state  $z = 6$  the more compact conformation is more intense. The lack of baseline resolution between the dynamic states of p53 can be attributed to the resolution of the MoQToF instrument, or rather interconverting conformers over the timescale of the experiment.

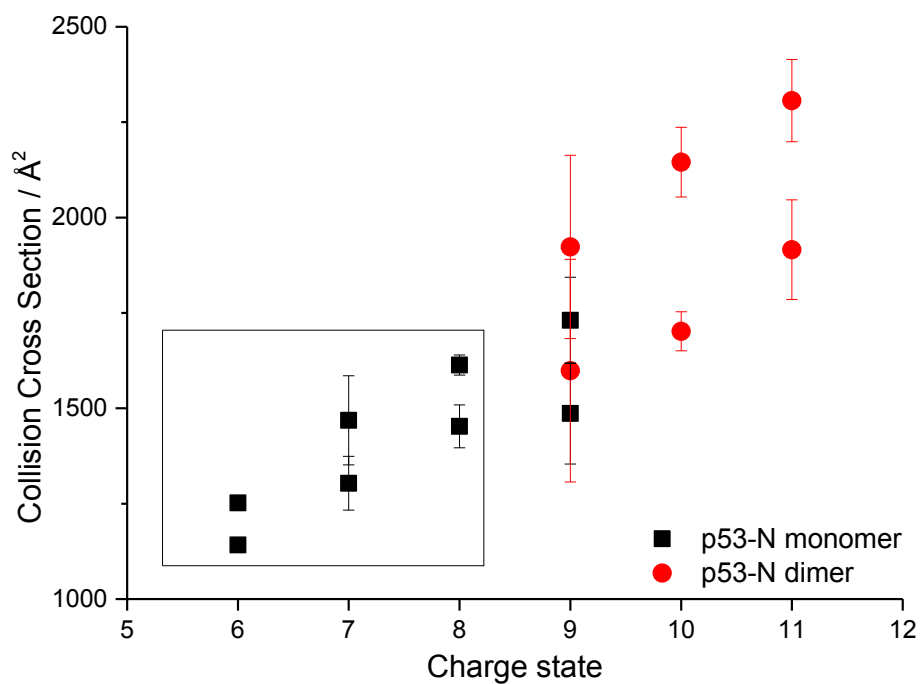


Figure 5.3: Collision cross sections versus charge obtained via the DT IM-MS of p53-N monomer and dimer from buffered solution conditions (200  $\mu$ M p53-N in 50 mM ammonium acetate). The black box highlights the most abundant charge states.

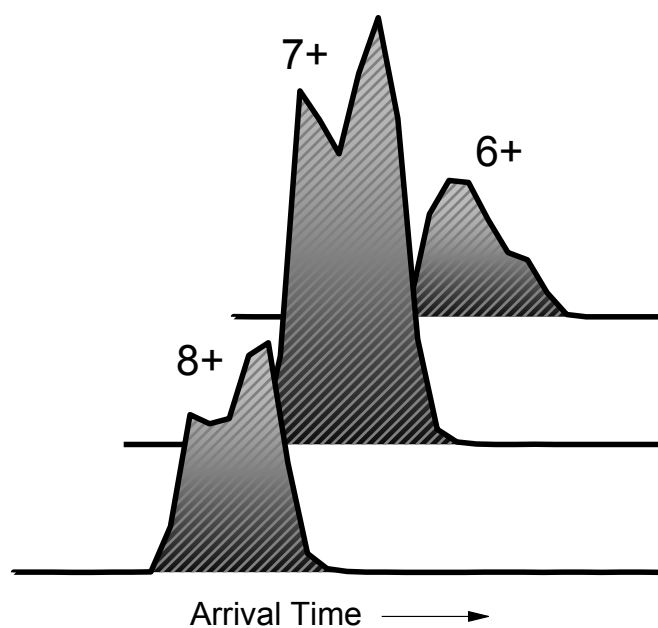


Figure 5.4: Experimental ATDs (taken at 40 V) of three most intense charge states  $[M+6H]^{6+}$ ,  $[M+7H]^{7+}$ ,  $[M+8H]^{8+}$  showing at least two species can be resolved for each charge state.



The prevalence of intrinsic structural flexibility is signature behaviour of transactivation domains of transcription factors, proposed by Liu *et al.* (31). His PONDR analysis revealed that 85% of transcription factors possess extended regions of intrinsic disorder. The conformational flexibility of the p53 transactivation domain may have been evolutionary carefully chosen to allow p53 to interact with a wide array of proteins, since each p53-protein complex may stabilize a specific p53 conformational isomer suitable for the particular interaction. Further, because of the conformational variability and adaptability, the p53 transactivation domain residues can adjust to the active sites of various enzymes (kinases, acetylases, etc.), which explains why many post-translational modifications of p53 map to the transactivation domain.

Charge state	Collision Cross Section / Å <sup>2</sup>	
6	1142±9,	1252±12,
7	1304±70,	1469±117
8	1453±56,	1613±26
9	1486±132,	1730±113

**Table 5.1:** Collision cross sections of monomer form of p53-N obtained via the DT IM-MS from buffered solution condition (200 µM p53-N in 50 mM ammonium acetate). Values are an average of three replicates and stated with the standard deviation of the mean.

Charge state	Collision Cross Section / Å <sup>2</sup>	
9	1599±300,	1923±240
10	1702±52,	2145±92
11	1916±131,	2306±108

**Table 5.2:** Collision cross sections of dimer form of p53-N obtained via the DT IM-MS from buffered solution condition (200 µM p53-N in 50 mM ammonium acetate). Values are an average of three replicates and stated with the standard deviation of the mean.

### 5.3. Molecular Architecture of p53-N-DBD

To map possible interactions of the p53 N-terminus with the DBD, mass spectrometry and ion mobility mass spectrometry measurements were conducted, using a construct comprising the superstable quadruple mutant of p53-DBD (residues

102-312) with full N-terminus (residues 1-102). Natan *et al.* (32) reported that the region of the N-terminus interacting with the p53-DBD could be pinpointed to residues 86-93, the hinge section between the folded DBD and the proline-rich region. These interactions increase the melting temperature of the tetrameric p53 by up to 2 °C indicating that p53-DBD tetramer is stabilised through additional interactions by its N-terminal region. Light scattering measurements have also demonstrated that a short segment from the TAD (residues 89-94) inhibits accelerated aggregation of the full-length p53 tetramer. In general this highlights the potential repercussions for studies of multi-domain proteins where weak ordered-disordered domain interactions can regulate the properties of proteins of complex structure.

### 5.3.1. MS of p53-N-DBD in buffered solution conditions

The observed spectrum of p53-N-DBD displays a wide charge state envelope over the acquired mass range (1000-5500  $m/z$ ). Under these conditions the most dominant species of p53 corresponds to a monomer (calculated  $M_R = 32613.7$  Da, observed  $M_R = 32625.9$  Da), with a wide-ranging charge state distribution of  $10 \leq z \leq 20$ . The spectrum displays a tetramodal charge distribution with centres at  $[M+14H]^{14+}$ ,  $[M+11H]^{11+}$ ,  $[2M+17H]^{17+}$  and  $[3M+21H]^{21+}$ , suggesting that at least four conformational families are present in the solution and transferred into the gas phase environment (MS shown in Figure 5.5). The overall charge state distribution of p53-N-DBD appears dissimilar to the spectrum of isolated superstable p53-DBD (Figure 5.10). Closer examination of spectra reveals a shift towards higher charge states of most intense peaks  $[M+11H]^{11+}$  and  $[M+14H]^{14+}$ , perhaps due to the more flexible and disordered nature of p53 in the presence of the N-terminus. The intrinsically disordered N-terminus also induces an increase in formation of more stable higher order multimeric species such as dimers and trimers, in accord with the findings of Natan *et al.* The ion envelope for  $[2M+zH]^{z+}$  is  $15 \leq z \leq 19$  along with the less intense population for  $[3M+zH]^{z+}$ , where  $20 \leq z \leq 22$ . Even under ‘near native’ conditions p53-N-DBD exhibits a wide charge state distribution for monomeric

species ( $10 \leq z \leq 20$ ), which is a distinctive behaviour of IDPs *in vacuo* as reported earlier by mass spectrometry (33).

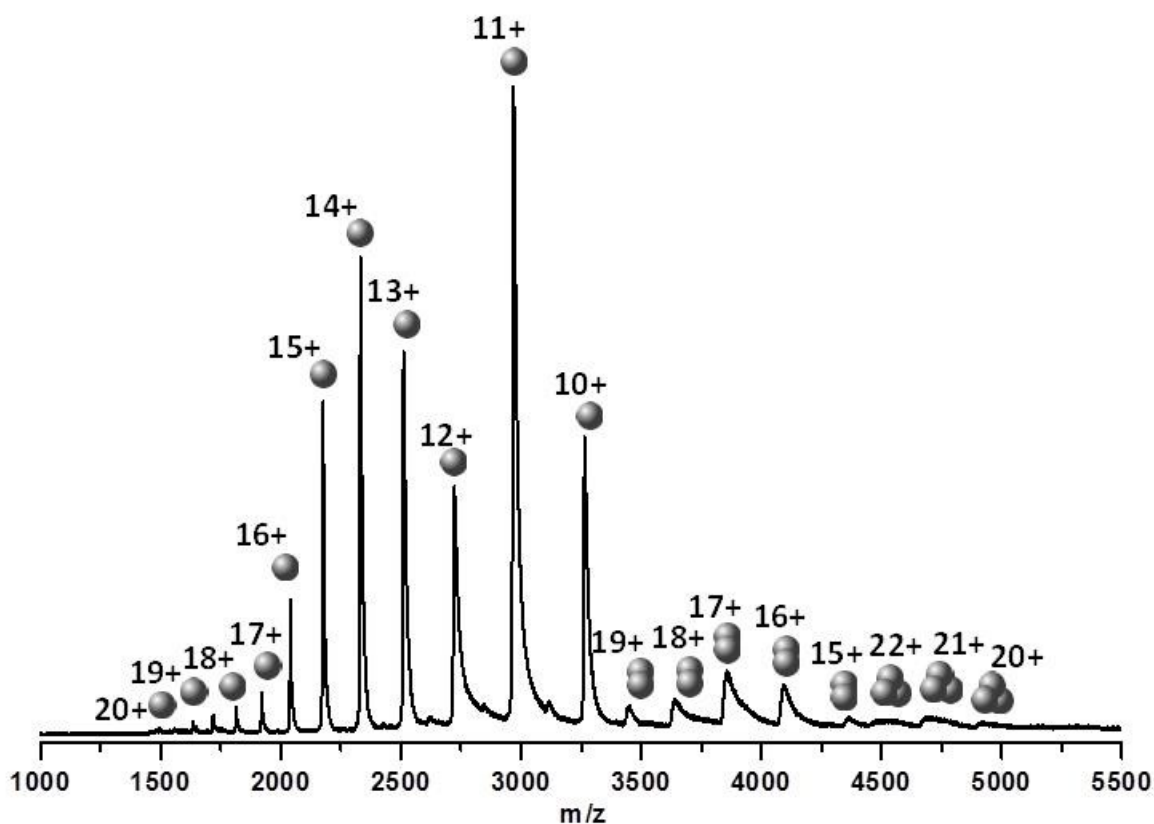
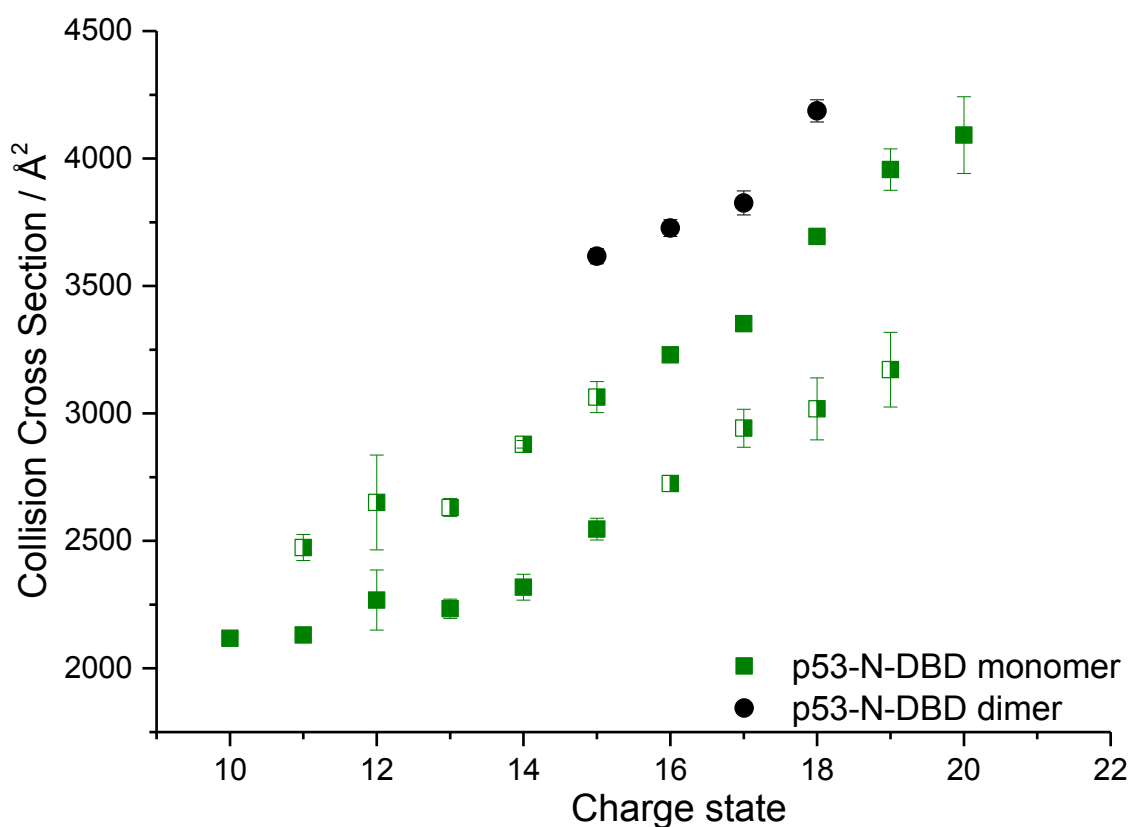


Figure 5.5: n-ESI spectrum of 90  $\mu$ M p53-N-DBD obtained via the DT IM-MS from buffered solution with 50 mM ammonium acetate.

### 5.3.2. Conformational Landscapes adopted by p53-N-DBD

Collision-cross sections measured for p53-N-DBD from aqueous solution buffered to pH 6.8 with 50 mM ammonium acetate versus its charge state are reported in Figure 5.6. The mean cross-section value of each ion increases with increasing charge, which can be attributed to expansion of the protein structure to overcome Coulomb repulsion as an increasing number of protons are accepted, a well-established phenomenon of gas-phase ions (34,35). The lowest charge state,  $z = 10$  provides a single resolvable conformer, indicative of a compact structure, whereas the most intense charge state  $z = 11$  adopts two conformations, with the most populated being

2131  $\text{\AA}^2$  and a lower intensity elongated form of 2474  $\text{\AA}^2$  (ATDs is shown in Figure 5.7) With increasing charge, the unfolding pathways of the protein are clearly evident and a minimum of two conformational isomers can be resolved for  $11 \leq z \leq 19$ . The majority of the protein for  $z = 10$ -15 is present in a compact structure, which also exists for  $z = 16$ -19, although here the more extended conformational family is the more prominent. At the highest charge state  $z = 20$ , only the more elongated conformer is observed (4092  $\text{\AA}^2$ ). Remarkably for the dimeric species only a single conformation is detected, suggesting that this is the only stable gas phase dimeric form of p53, nonetheless still in a collapsed state (CCS only larger by 13% than CCS for the equivalent monomeric forms).



**Figure 5.6:** Collision cross sections versus charge obtained via the DT IM-MS of p53-N-DBD monomer and dimer from buffered solution conditions (90  $\mu\text{M}$  p53-N-DBD in 50 mM ammonium acetate). The half-filled symbol represents less abundant conformer at a given charge state.

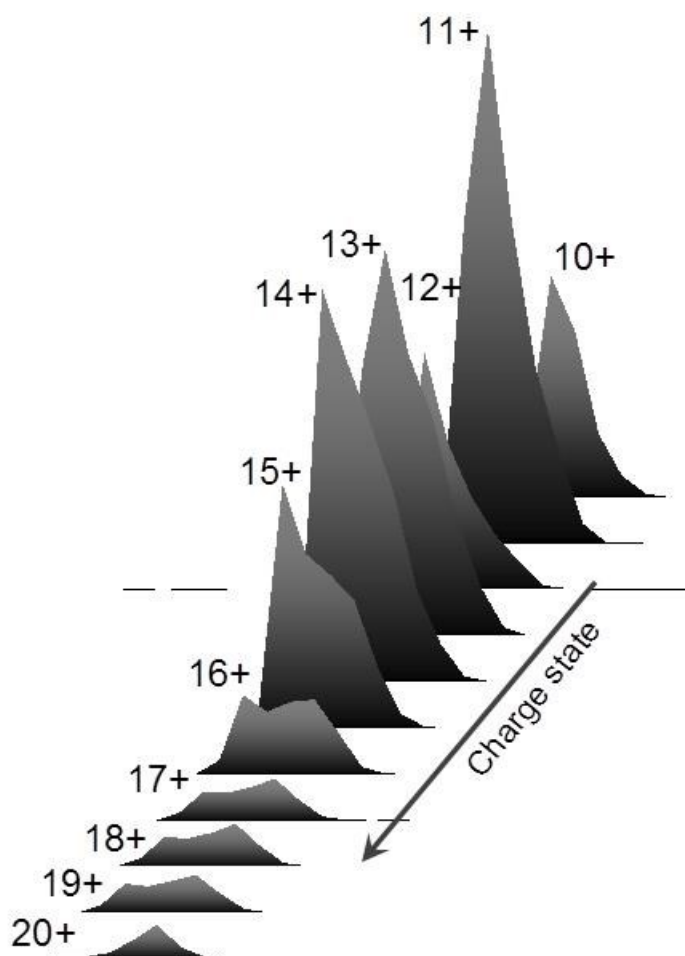


Figure 5.7: Experimenta ATDs taken at 35 V of p53-N-DBD for  $z = 10 - 20$ .

Charge state	Collision Cross Section / $\text{\AA}^2$	
10	2118 $\pm$ 9	
11	2131 $\pm$ 3,	2474 $\pm$ 51
12	2268 $\pm$ 118,	2651 $\pm$ 186
13	2234 $\pm$ 38,	2631 $\pm$ 35
14	2318 $\pm$ 50,	2879 $\pm$ 15
15	2546 $\pm$ 43,	3064 $\pm$ 61
16	2725,	3230
17	2942 $\pm$ 74,	3352
18	3018 $\pm$ 122,	3694 $\pm$ 28
19	3172 $\pm$ 146,	3957 $\pm$ 81
20		4092 $\pm$ 150

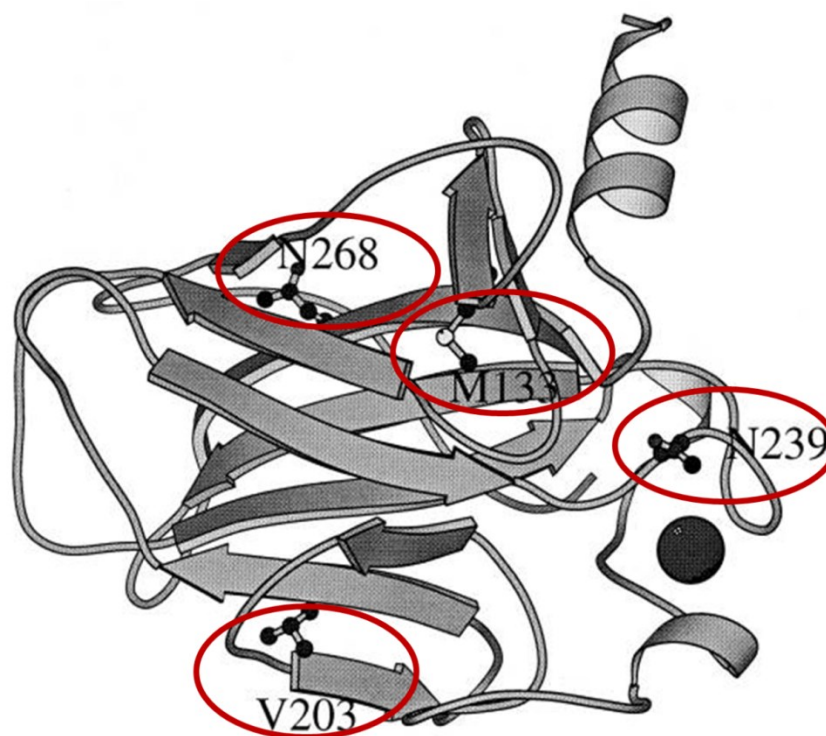
Table 5.3: Collision cross sections of monomeric p53-N-DBD obtained via the DT IM-MS from buffered solution condition (90  $\mu\text{M}$  p53-N-DBD in 50 mM ammonium acetate). Values are an average of two replicates and stated with the standard deviation of the mean.

Charge state	Collision Cross Section / $\text{\AA}^2$
15	3617 $\pm$ 30
16	3727 $\pm$ 33
17	3826 $\pm$ 47
18	4187 $\pm$ 44

**Table 5.4: Collision cross sections of p53-N-DBD dimer obtained via the DT IM-MS from buffered solution condition (90  $\mu\text{M}$  p53-N-DBD in 50 mM ammonium acetate). Values are an average of two replicates and stated with the standard deviation of the mean.**

## **5.4. The Quest for Stable p53 - Quadruple Mutant of p53-DBD: M113/V203/N239Y/N268D**

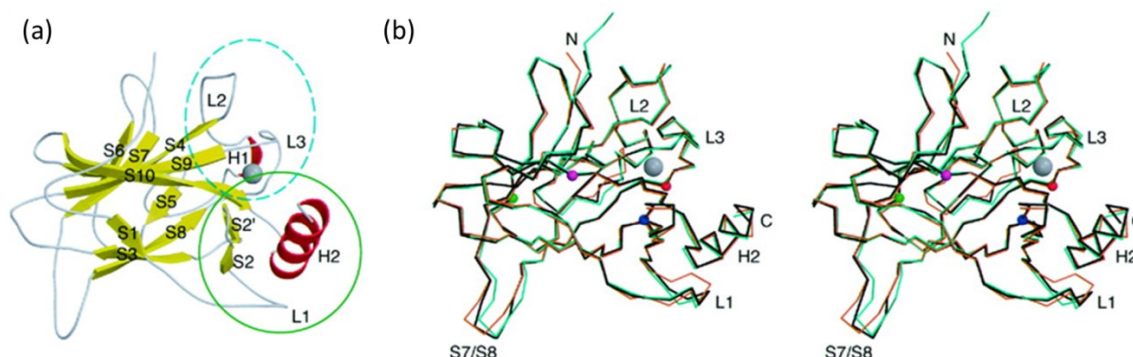
The low thermodynamic stability and associated aggregation propensity is a known Achilles' heel of human p53 protein that has hampered many biophysical and structural studies. By adopting a semi-rational design based on the molecular evolution of p53, Nikolova *et al.* (36) designed a superstable (SS) quadruple mutant of the human p53-DNA binding domain. In this approach naturally occurring amino acid substitutions were identified by assessing the sequences of p53 homologues from more than 20 species and subsequently introduced into p53 core domain to measure their effect on the thermodynamic stability of the protein. In addition, a N239Y substitution was also introduced although does not occur naturally (37). Four most stable point mutations M133L, V203A, N239Y and N268D were combined that increased overall rigidity and thermostability of p53. The mutations are positioned in different regions of the core domain shown in the Figure 5.8.



**Figure 5.8:** Diagram of the DNA-binding core domain of p53 produced by MOL-SCRIPT showing the location of four mutations used to design the superstable quadruple mutant. Image and caption taken from Nikolova *et al.* (36).

The effect of these point mutations is nearly additive, and the main contribution to increased stability arises from N239Y and N268D mutations, both known to act as second-site suppressors for various cancer-associated mutations (37). Nikolova *et al.* (36) also measured the DNA binding affinity of the wild-type and quadrupule mutant using surface plasmon resonance and reported the apparent binding constants to be nearly identical. The crystal structure of the engineered mutant (residues 96-290) shows small local structural changes in the loop regions, whereas the overall structure of  $\beta$ -sandwich scaffold and the DNA binding surface is maintained (Figure 5.9) (38). Mutation M133L located in strand S2 of the loop-sheet-helix motif causes only minor deviations in the neighbouring residues. Similarly mutation V203A in turn between  $\beta$ -strands S5 and S6 induces a local re-packing of the side chains around the mutation site. The crystal structure also explains the larger effect of the N268D and N239Y substitutions on the stability increase of p53 protein. These mutations act on different regions of the core domain. Mutation N268D alters

hydrogen bonding network in  $\beta$ -strand S10, which is part of both the  $\beta$ -sandwich and the loop-sheet-helix motif. Contrary to the other three mutations, the N239Y has so far not been identified in any known p53 sequence.



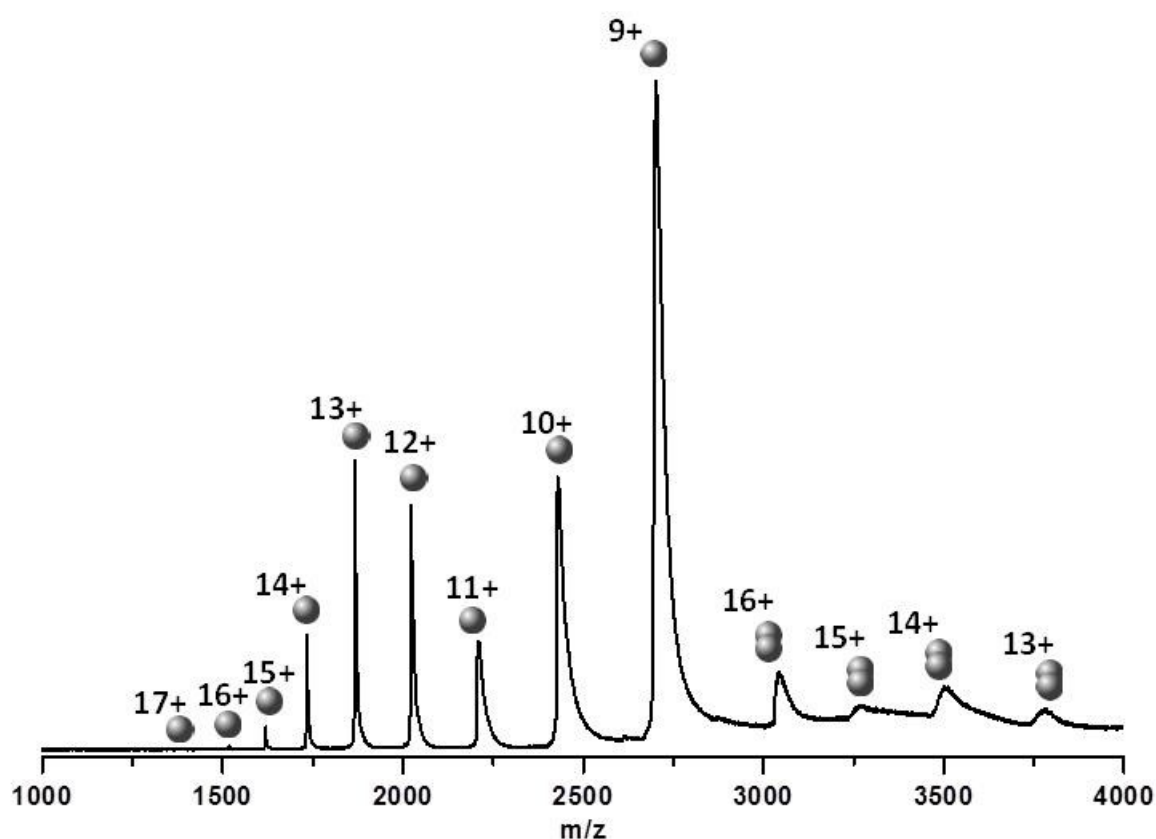
**Figure 5.9:** The X-Ray structure of p53 core DNA-binding domain quadruple mutant M113L/V203A/N239Y/N268D. (a) Ribbon image showing the  $\beta$ -sandwich (in yellow), loop 1 and 2 regions (encircled in blue dashed line) connected via a zinc ion (grey sphere) and a loop-sheet-helix motif (encircled in green solid line). (b) Stereo view of the  $Ca$  trace of the p53 core DNA-binding domain quadruple mutant (shown in black) superimposed on DNA-free wild type (PDB 1TSR; chain A, blue) and DNA-bound wild type (PDB 1TSR; chain B, orange). The mutation sites in the quadruple mutant are marked by small coloured spheres: M113L (blue), V203A (green), N239A (red), and N239Y (magenta). Image and caption taken from Joerger *et al.* (38).

#### 5.4.1. MS of SS p53-DBD in buffered solution conditions

From the buffered solution, the superstable quadruple mutant p53 monomer is observed in a charge state distribution with  $9 \leq z \leq 18$  along with few additional multimeric species. Each of the monomer charge states can be assigned to the zinc-bound form of the protein (calculated  $M_R = 24245.4$  Da, observed  $M_R = 24263.4$  Da) which suggests that all conformations present in solution retain the zinc binding site, and that it is not perturbed by the n-ESI desolvation process or by any subsequent coulombically driven unfolding. The most dominant species correspond to monomeric  $[M+9H]^{9+}$  and  $[M+10H]^{10+}$  indicative of compact conformations prevalent in solution as there are few solvent-accessible sites available for protonation. Lower intensity monomeric species between charge states  $11 \leq z \leq 17$



can be attributed to more unfolded states of p53. The wide charge state distribution even under buffered solution conditions is a distinctive signature for IDPs as reported previously (33). A less abundant ion envelope for the dimeric form of p53 is also observed ( $z = 13-16$ ), representing higher order multimeric association in solution, again zinc is bound stoichiometrically. Overall the charge state distribution is relatively similar to the *de facto* WT p53 DBD (refer to Chapter 3, Section 3.4.1, Figure 3.3 for the mass spectrum) but the higher charge species appear to be more populated manifesting a more disordered nature for this superstable quadruple mutant.



**Figure 5.10:** n-ESI spectrum of superstable 73  $\mu$ M p53-DBD obtained via the DT IM-MS from buffered solution with 50 mM ammonium acetate.

### 5.4.2. Conformational Populations occupied by SS p53-DBD

Rotationally averaged CCS obtained from aqueous solution buffered with 50 mM ammonium acetate are tabulated in Table 5.5 and plotted as a function of charge state in Figure 5.11. Low charge states generally adopt compact conformations indicative of native like structure of the protein. The dominant charge states  $z = 9$  and 10 assume multiple conformations, where populations with the lowest CCS are the most abundant (Figure 5.12 illustrates the abundance of the conformation within the protein). Remarkably, this most compact form of the superstable mutant (represented as black filled squares) is retained over a wide array of charge states ( $9 \leq z \leq 17$ ), whereas for the *de facto* WT p53-DBD this compact ensemble was found to span only charge state  $z = 9 - 13$  (refer to Chapter 3, Section. 3.4.4, Figure 3.7 for the CCS plot *versus* charge state).

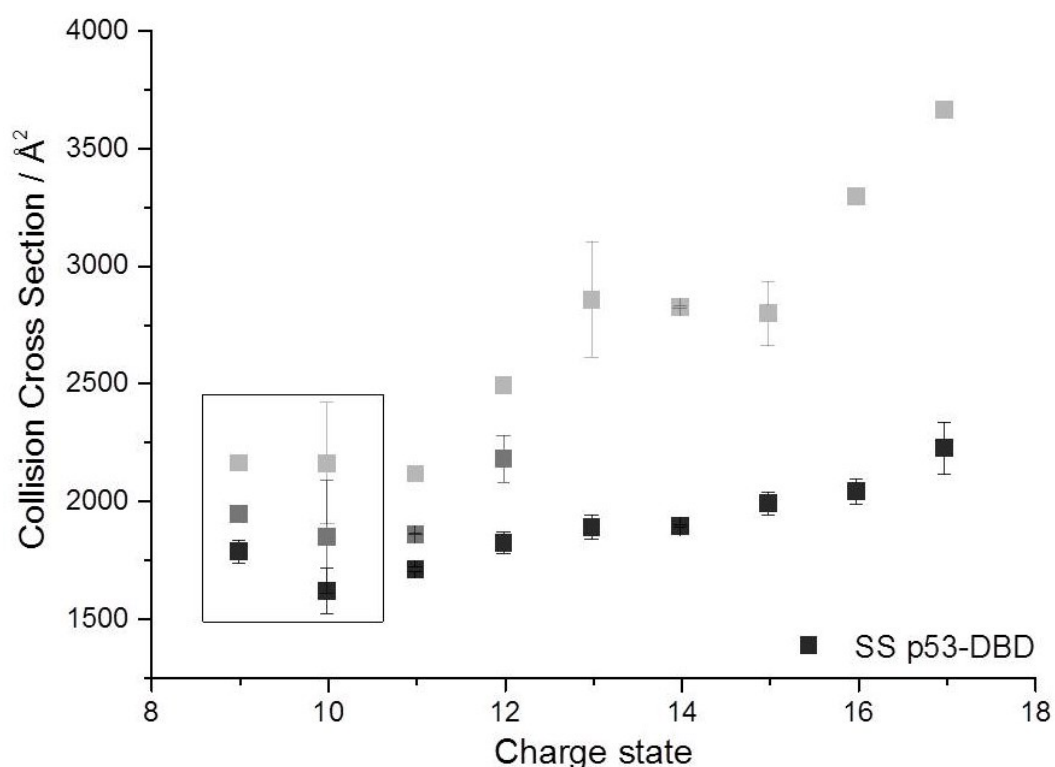
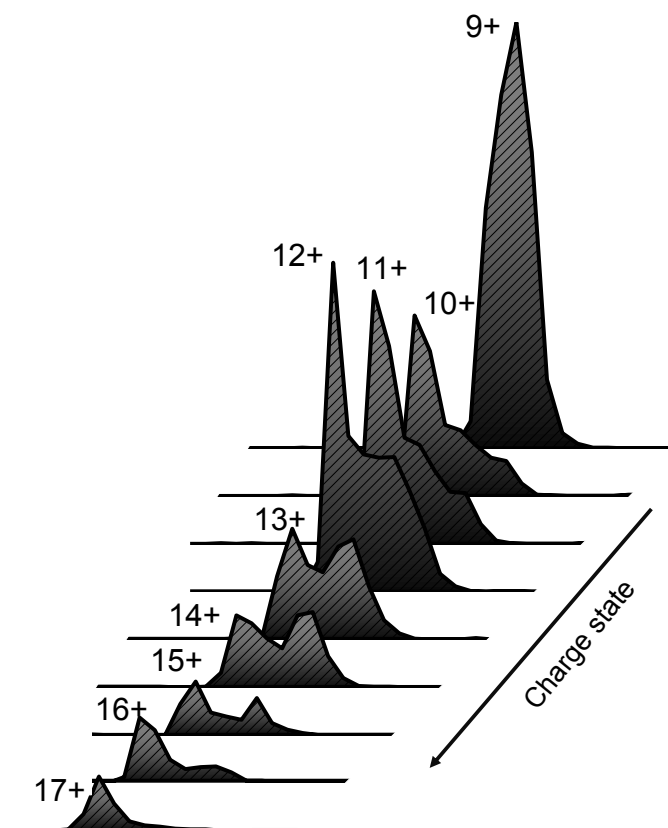


Figure 5.11: Collision cross sections versus charge obtained via the DT IM-MS of superstable quadruple mutant of p53-DBD monomer from buffered solution conditions (73  $\mu$ M SS p53-DBD in 50 mM ammonium acetate). The black box highlights the most abundant charge states.



**Figure 5.12:** Experimental ATDs taken at 35 V of monomer form of SS p53-DBD for  $z = 9 - 17$ .

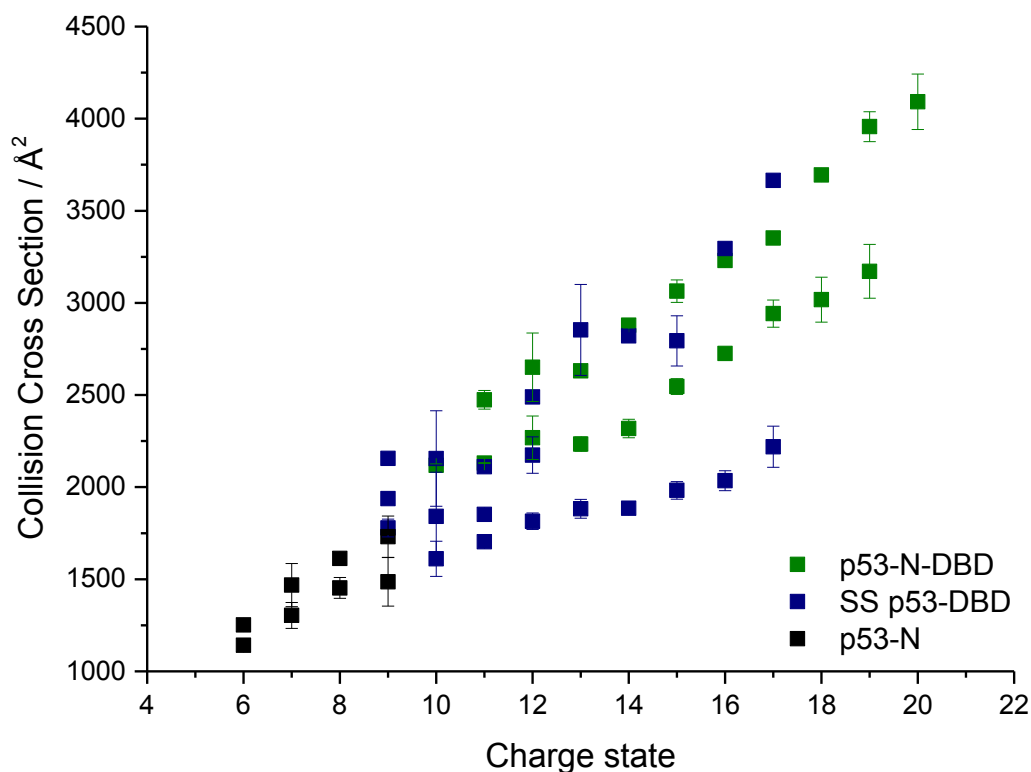
This idiosyncratic rigidity and clear resistance to Coulombically-driven unfolding of the ‘native like’ structure of superstable p53-DBD perhaps might be an aftermath of introduction of four stabilising mutations M133L, V203A, N239Y and N268D. Overall the unfolding trajectory of the superstable mutant displays an extraordinary extent of stability regions seen as ‘shelving’, cumulating around lower charge states  $z = 9 - 12$ . The most compact and rigid conformer unfolds minimally, for  $z = 10$ , CCS of  $1611 \text{ \AA}^2$  is measured, whereas for  $z = 17$  the CCS of  $2219 \text{ \AA}^2$  is observed (which is only larger by 17%). The intermediate conformer (shown as dark grey symbol) is only detected at the lower charge states and is entirely lost at  $z = 13$ . The third conformational family (shown as light grey filled squares) with a CCS of  $1937 \text{ \AA}^2$  at  $z = 9$  demonstrates two sharp jumps in CCS during the unfolding pathway, occurring at  $z = 12$  and  $z = 16$ .

Charge state	Collision Cross Section / $\text{\AA}^2$		
9	1778 $\pm$ 48,	1937	
10	1611 $\pm$ 96,	1841 $\pm$ 240,	2155 $\pm$ 259
11	1703 $\pm$ 9,	1852 $\pm$ 2,	2111
12	1815 $\pm$ 45,	2174 $\pm$ 99,	2489
13	1882 $\pm$ 51,		2854 $\pm$ 247
14	1885 $\pm$ 5,		2821 $\pm$ 5
15	1983 $\pm$ 48,		2794 $\pm$ 137
16	2035 $\pm$ 54,		3295
17	2219 $\pm$ 111,		3665

**Table 5.5: Collision cross sections of monomeric SS p53-DBD obtained via the DT IM-MS from buffered solution condition (79  $\mu\text{M}$  SS p53-DBD in 50 mM ammonium acetate). Values are an average of two replicates and stated with the standard deviation of the mean.**

Overall the IM-MS data paints an intriguing picture of the engineered mutant of p53-DBD, conformationally diverse with multiple isomers detected, yet exhibiting some exceptional intrinsic rigidity of the structure. The design of the superstable mutant, which maintains the overall structural features of the WT protein, unquestionably supports biophysical and biochemical studies, particularly in tackling the quaternary structure of full length p53. Nonetheless, the results should be interpreted carefully and when possible the *de facto* WT p53 construct should be used, especially whilst only the isolated p53-DBD is investigated.

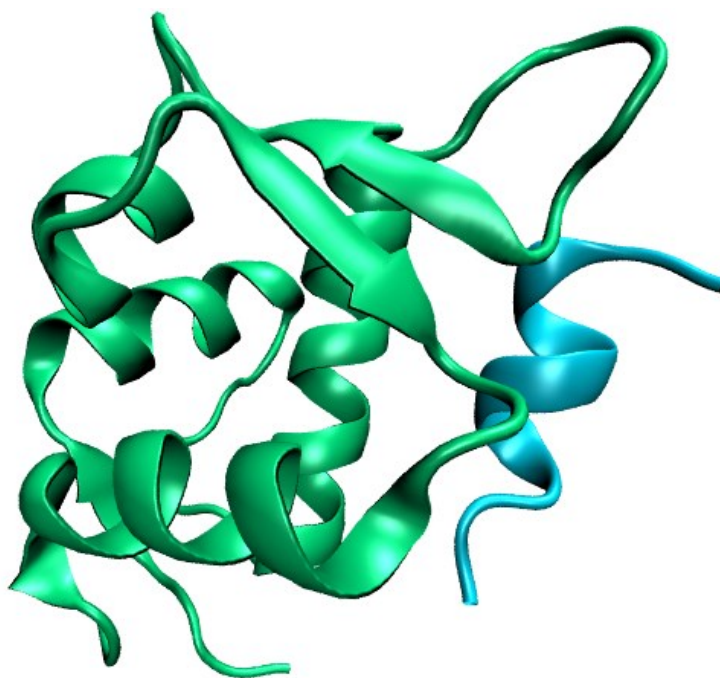
Collision cross sections ranges measured for individual & multi-domain p53 constructs are plotted in Figure 5.13.



**Figure 5.13:** Comparison of collision cross sections versus charge state measured for monomeric p53-N, superstable quadruple mutant of p53-DBD and p53-N-DBD.

## 5.5. Architecture of p53:Mdm2 Complex

Critical to the inhibitory action of the oncogene product, Mdm2, on the tumour suppressor p53, is association of the N-terminal domain of Mdm2 with the transactivation domain of p53. Much light has been shed on the interaction between p53 and Mdm2 by determination of the crystal structure of the N-terminal domain of Mdm2 (residues 25-109) bound to 15-residue transactivation domain peptide of p53 (residues 15-29) by Kussie *et al.* (11).



**Figure 5.14: X-Ray crystal structure of human p53-N (residues 15-29, shown in blue colour) in complex with Mdm2-N (shown in green); (PDB ID 1YCQ) (11).**

The region of Mdm2 that interacts with p53 consists of an independently-folding domain comprising of a four helix bundle flanked on either side by  $\beta$ -sheets. A fragment containing residues 15 - 29 of p53 folds as an  $\alpha$ -helix that interacts with a hydrophobic cleft on the Mdm2 surface (structure shown in Figure 5.14). The interaction is interceded by highly conserved amino acids Phy-19, Trp-23 and Leu-26 of p53 that interact with multiple Mdm2 hydrophobic residues, including Leu-54, Met-62, Tyr-67 and Val-93, whose side chains form the cleft on the Mdm2 surface. The amphipathic nature of the p53 helix (hydrophobic on the side that interacts with Mdm2 and hydrophilic on the side exposed to solvent) is crucial for the binding.

### **5.5.1. Mdm2 N-terminal - MS and IM-MS perspective**

The NMR experiment performed by Uhrinova *et al.* (39) revealed an unstable nature and multiple conformations of ligand-free Mdm2-N (residues 2-118). This solution structure enabled the hypothesis that upon binding to p53, an N-terminal segment of

apo-Mdm2 that occludes contact with the end of the p53-binding cleft is displaced. Further, Uhrinova argues that Mdm2-N undergoes expansion upon binding, which is attained through conformational rearrangement of its two sub-domains resulting in a more ordered, and open conformation of the binding cleft.

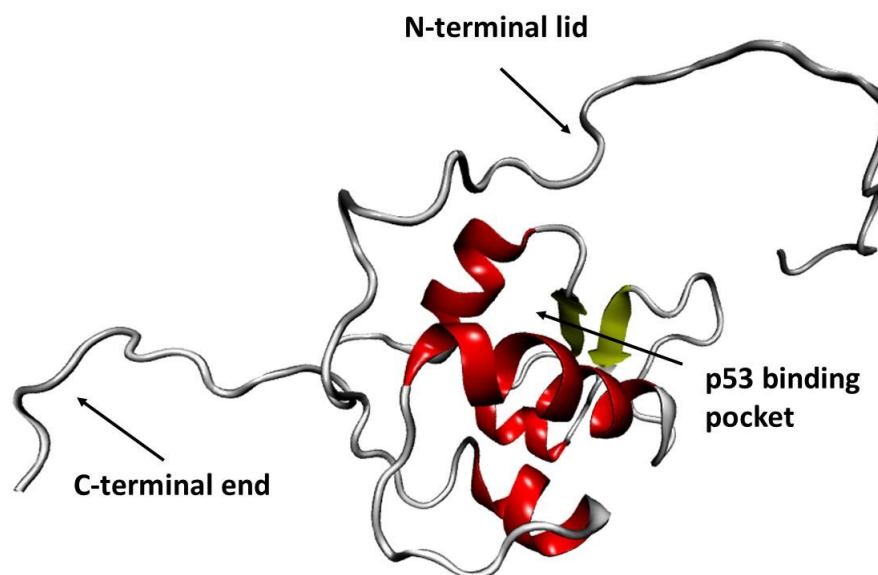


Figure 5.15: Solution NMR structure of Mdm2-N (PDB ID 1Z1M) (39) highlighting the C-terminal end, N-terminal lid region and the hydrophobic p53 binding pocket.

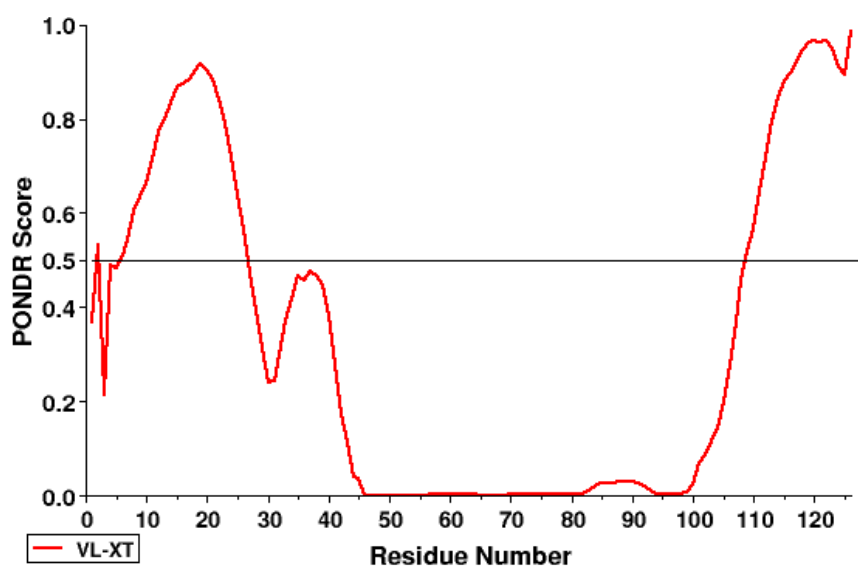
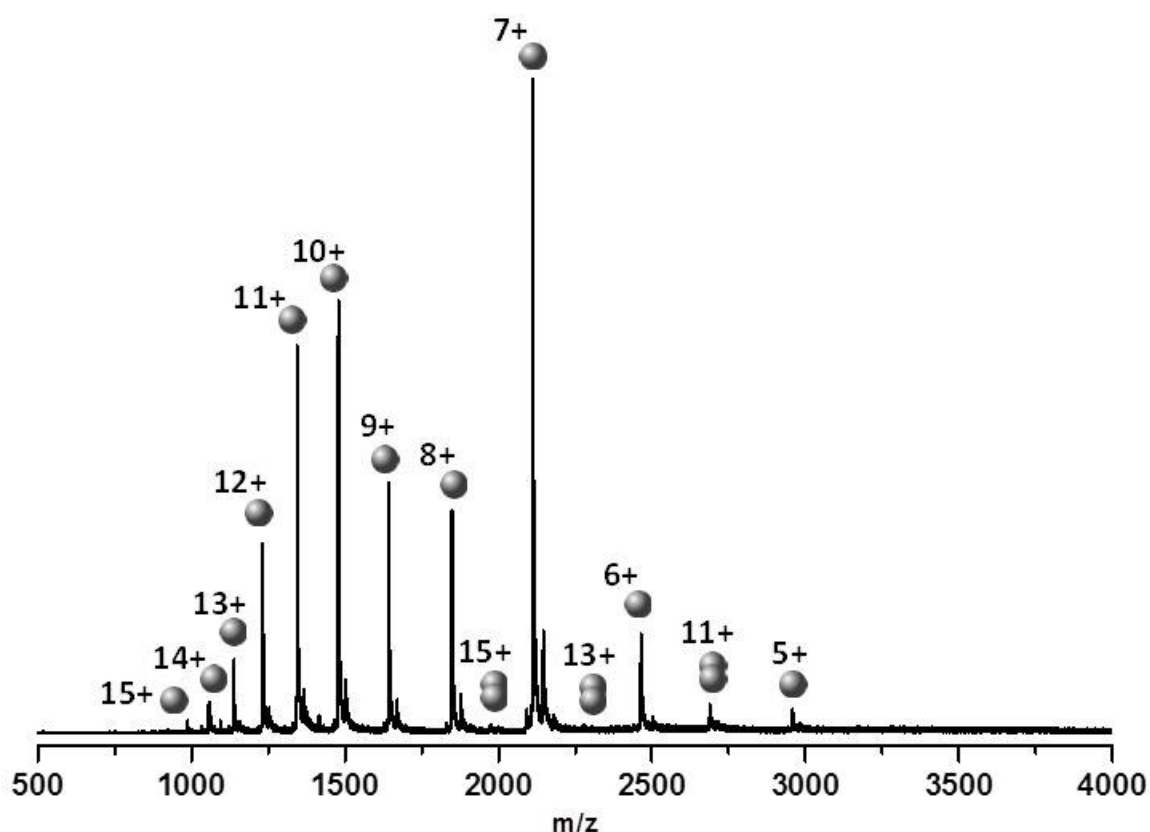


Figure 5.16: PONDR score of disorder for Mdm2-N, showing the disordered lid region at residues 4-27 (40).

Schon *et al.* (12) also reported extensive changes in chemical shifts when Mdm2-N (residues 2-125) is in complex with peptides from the TAD of p53 (residues 15-29 and 17-26). These perturbations were not only restricted to the binding cleft but could be mapped to the other parts of the protein, suggestive of extensive conformational rearrangement. McCoy *et al.* (41) also suggested that the N-terminal residues of Mdm2-N form a flexible lid over the p53-binding groove (highlighted in the structure of Mdm2-N in Figure 5.15 as well as in the PONDR score in Figure 5.16), it was therefore important to use a fragment of Mdm2 that contains the sequence prior to residues 18, in contrast to some previous studies that employed truncated constructs (11,41).

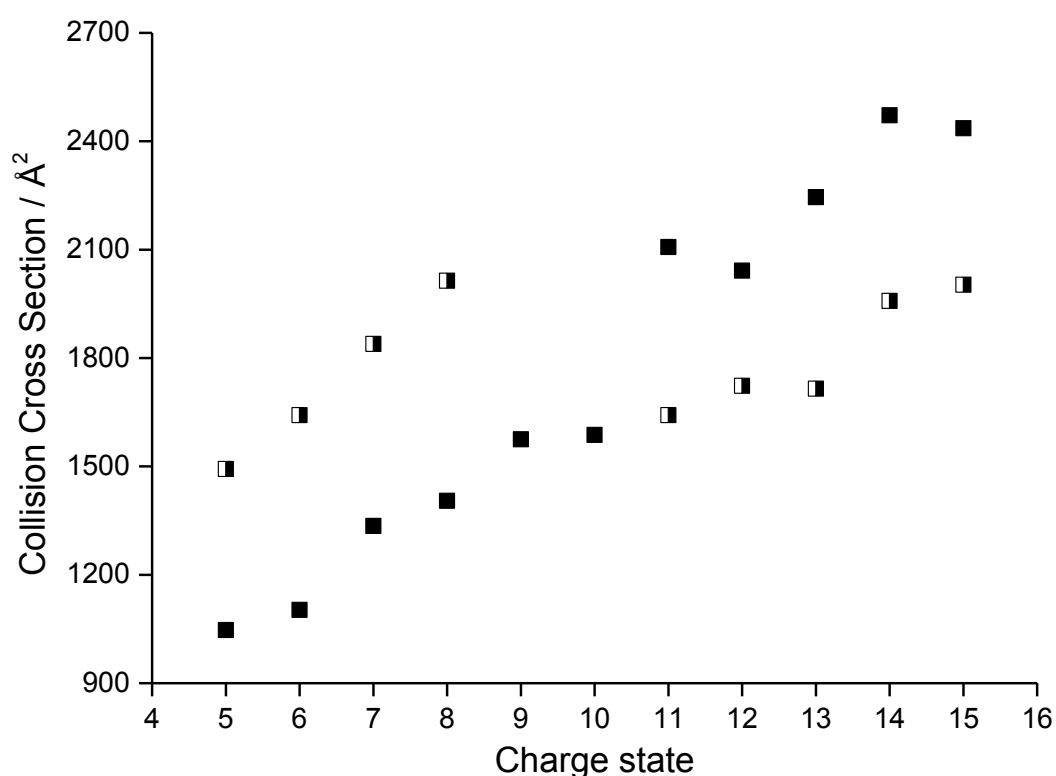
The mass spectrum obtained for WT N-terminal Mdm2 sprayed from near native conditions (50 mM, ammonium acetate) is illustrated in Figure 5.17.



**Figure 5.17:** n-ESI spectrum of 115  $\mu$ M Mdm2-N obtained via the DT IM-MS from buffered solution with 50 mM ammonium acetate.



It displays a wide charge state distribution for monomer ( $5 \leq z \leq 15$ ); characteristic behaviour of IDPs. The central spectral peak can be assigned to  $z = 7$  evidence of a predominant compact structure to the protein. CCS obtained are tabulated in Table 5.6 and plotted as a function contrasting experimental CCS values *versus* charge state in Figure 5.18. Multiple conformations are detected; specifically up to two conformations can be resolved for each charge state. The data overall shows a smooth increase in cross-section with respect to charge. However the unfolding trajectory demonstrates some sharp ‘jumps’ in the CCS values, one at  $z = 6$  (from 1103 to 1335 Å<sup>2</sup>) and  $z = 12-13$  (2042 to 2245 Å<sup>2</sup>, and 1715 to 1958 Å<sup>2</sup>).



**Figure 5.18:** Collision cross sections versus charge obtained via the DT IM-MS of monomer form of Mdm2-N from buffered solution condition (59 µM Mdm2-N in 50 mM ammonium acetate). The half-filled square symbol represents less abundant conformer at a given charge state.

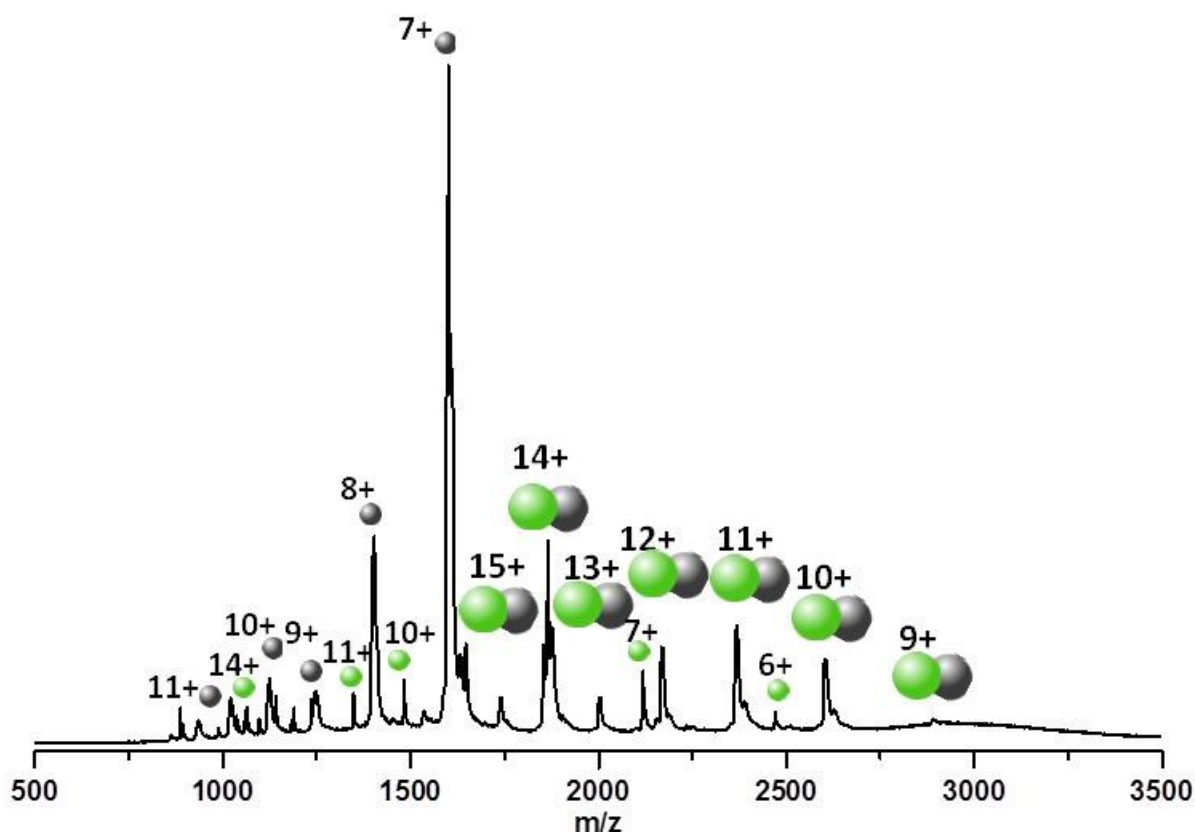
Charge state	Collision Cross Section / Å <sup>2</sup>	
5	1047,	1493
6	1103,	1642
7	1335,	1839
8	1405,	2014

Charge state	Collision Cross Section / Å <sup>2</sup>
9	1575,
10	1587,
11	1642, 2107
12	1723, 2042
13	1715, 2245
14	1958, 2472
15	2003, 2436

**Table 5.6: Collision cross sections of monomer form of Mdm2-N sprayed from buffered solution condition (59 µM Mdm2-N in 50 mM ammonium acetate) obtained via the DT IM-MS.**

### 5.5.2. Interactions of p53-Mdm2 probed by IM-MS

Mass spectrometry and ion mobility mass spectrometry was used to study complex formation between p53-N (residues 1-105,  $M_R = 11162.3$  Da) and Mdm2-N (residues 1-126,  $M_R = 14926.1$  Da) (Figure 5.19 shows the resultant mass spectrum). Proteins were incubated up to three hours; an increase in the incubation time did not alter the binding strength indicating an instant binding mechanism.

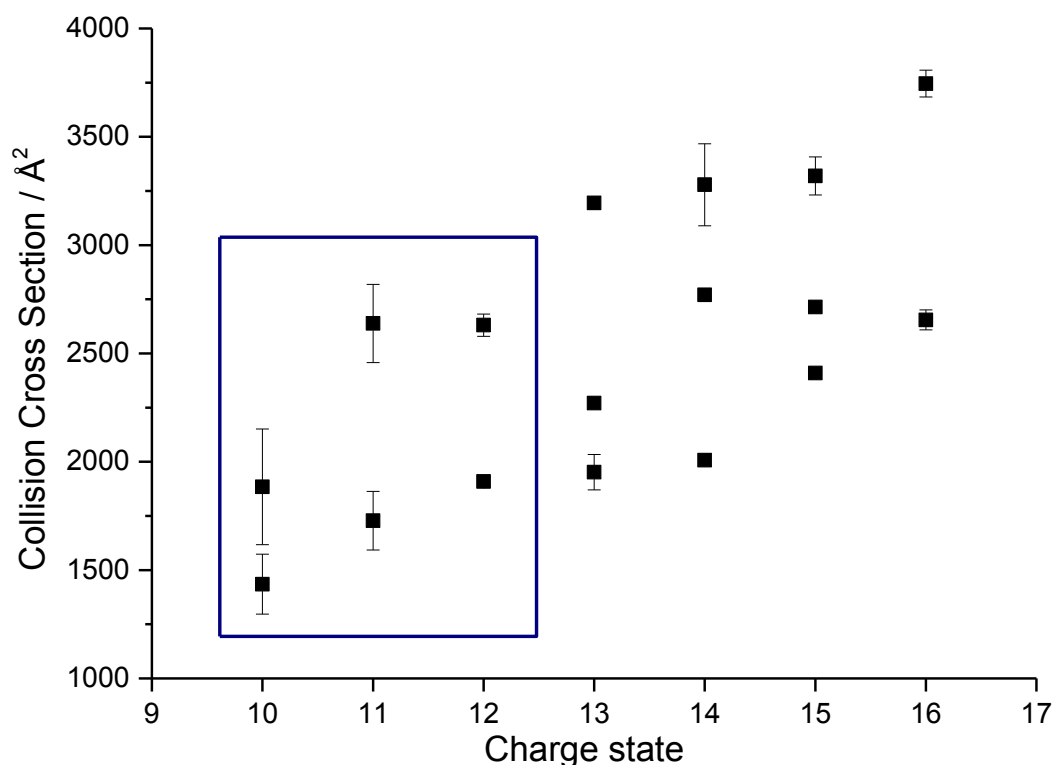


**Figure 5.19:** n-ESI spectrum of 49  $\mu\text{M}$  p53-N in complex with 16  $\mu\text{M}$  Mdm2-N. Mass spectrum was obtained from solution buffered with 50 mM ammonium acetate, incubated for five minutes and recorded via the DT IM-MS. The single grey spheres represent p53-N monomer, whereas the single green spheres represent Mdm2-N monomer. The mixed dimers spheres (green and grey) represent the p53-N – Mdm2-N complex.

The p53-N protein was found to bind Mdm2-N as 1:1 complex only. A charge state distribution of  $9 \leq z \leq 15$  for the complex was observed, the dominant peak of the complex is assigned to  $[\text{M}+11\text{H}]^{11+}$ . In the spectrum, ion envelopes of unbound monomeric ensembles for p53-N and Mdm2-N were also detected, which may perhaps be a consequence of transient and dynamic character of the interaction between the two proteins. There is also a notable depletion of all charge states corresponding to N-terminal Mdm2 in the presence of p53-N binding partner. The formation of full-length p53-Mdm2 complex was assessed by Funari et al. (42) using single molecule atomic force spectroscopy. The binding strength and dissociation rate constant was found to fall within a range of values characteristic of dynamical protein-protein interactions similar to several other ligand-receptor complexes, such as cadherins and selectins (43,44). Indeed, the cellular activity of the p53-Mdm2

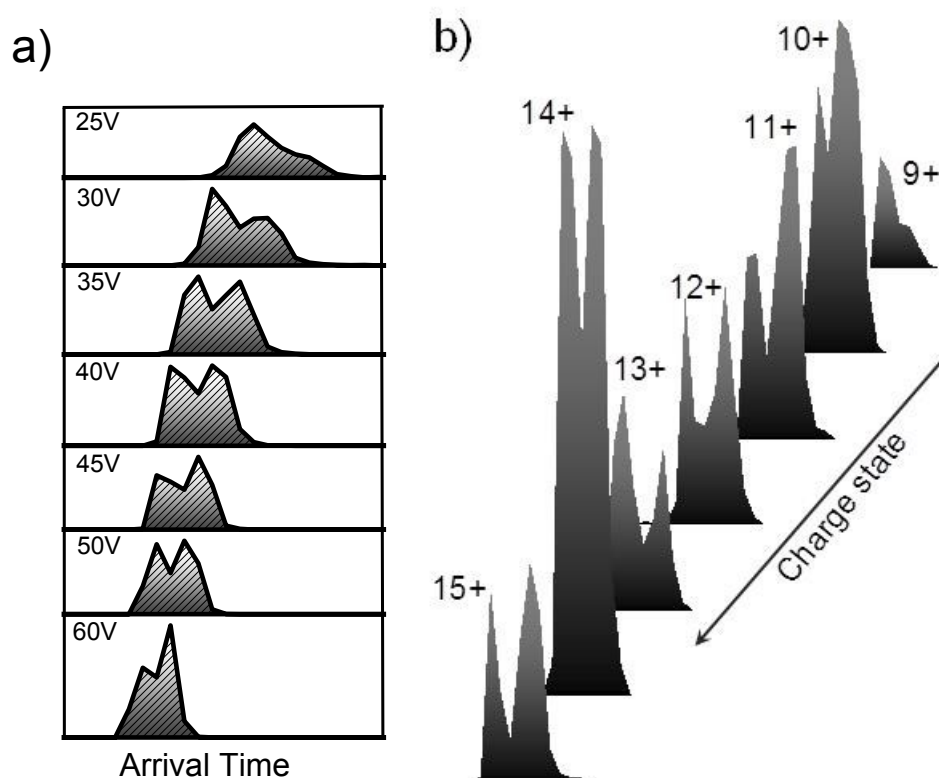
complex is more likely to require a dissociation rate constant typical of transient, dynamic interactions to permit, for instance an ubiquitination and the subsequent dissociation of p53 to undergo degradation (45). Moreover, Funari's findings are in close agreement with a stopped-flow fluorimetric study by Schon et al. (12), which utilised the two N-terminal domains of both p53 and Mdm2 proteins (rather than full-length proteins). It can be inferred that the two cognate binding sites of the complex experience a bio-recognition process without any steric hindrance effect from the presence of the entire milieu of the proteins (42).

Rotationally averaged CCS obtained from aqueous solution buffered with 50 mM ammonium acetate are reported in Table 5.7 and plotted as a function of charge state in Figure 5.20. The dominant charge states ( $z = 10 - 12$ ) provide at least two resolvable conformers, for  $z = 13 - 15$  three conformational isomers were resolved. ATDs of the complex for  $z = 9-15$ , displaying the abundance of the conformation within the protein are shown in Figure 5.21. Here for  $z = 10 - 12$  and  $z = 15$  the larger conformer is highly abundant (arriving later), whereas for  $z = 9$  and  $13$  the more compact conformation is more intense. For  $z = 14$  the intensity of conformers appears to be shared equally.



**Figure 5.20:** Collision cross sections versus charge obtained via the DT IM-MS of p53-Mdm2 complex from buffered solution conditions (50 mM ammonium acetate). The blue box highlights the most abundant charge states.

Additionally the intensity of each conformer switches during different drift voltages, the electric potential difference across the drift cell (shown in Figure 5.21a), perhaps again suggesting the extraordinary dynamic and transient nature of the complex. To simplify, structured, folded proteins with little conformational flexibility in solution, would give rise to narrower ATDs (peak width would be expected to be smaller), typically presenting one conformation at different drift voltages applied to the cell. In turn, less structured or unfolded proteins would give rise to wider ATDs often with multiple conformations (as seen here). Analysis of the width of ATDs could potentially act as an indicator of intrinsic disorder of proteins, semi quantifying the conformational flexibility and spread of structures in a family of the systems.



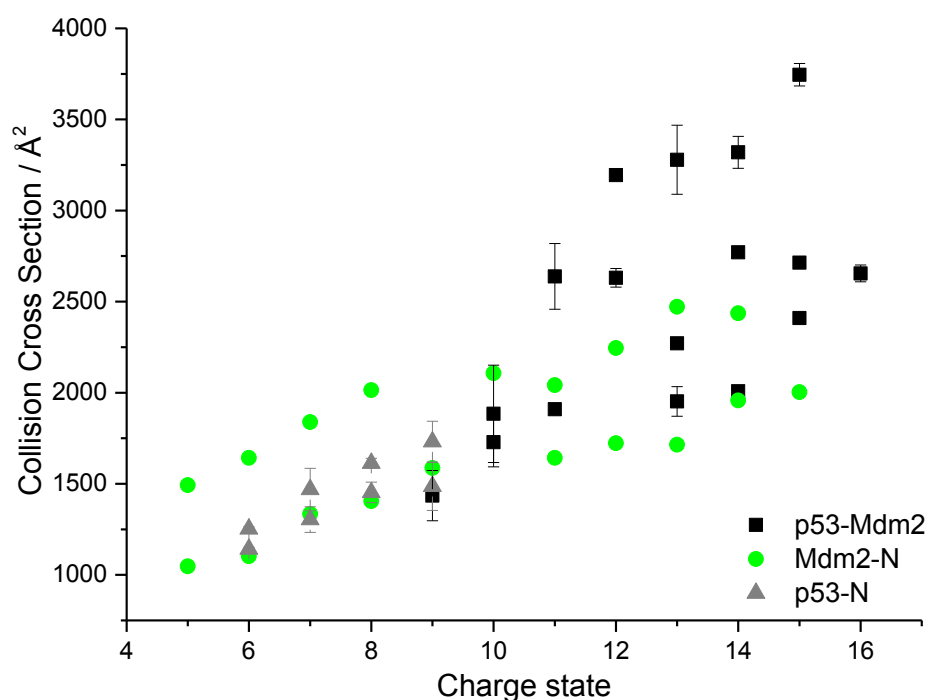
**Figure 5.21:** a) Experimental ATD of p53-N-Mdm2-N for a single charge state  $[M + 14H]^{14+}$  taken at seven different drift voltages (60, 50, 45, 40, 35, 30, 25 V) showing the switching intensity of conformers. At 60 and 45 V the conformer arriving later appears to be more intense whereas at 35, 30 and 25 V first conformation is more abundant; at 50 and 35 V intensity of conformers seems to be shared equally. b) Experimental ATDs taken at 40 V of p53-Mdm2 complex for  $z = 9 - 15$ .

Overall the cross-section values are observed to increase upon addition of charge attributable to unfolding of the protein structure to minimise Coulomb repulsion (34). Interestingly the unfolding pathway of the complex appears to be more heterogeneous than that of individual domains of apo-Mdm2 or isolated p53. The explicit-solvent molecular dynamics simulations performed for bound and apo-Mdm2 also showed that their unfolding pathways are different and the transition state of bound Mdm2 is more heterogeneous than that of apo-Mdm2 (46). The unfolding trajectory lacks any major stability regions seen as ‘shelving’ of several charge states, providing once more evidence of ‘the ephemeral’ and conformationally dynamic character of the complex.

Charge state	Collision Cross Section / $\text{\AA}^2$		
10	1435±138,		1884±267
11	1728±135,		2638±180
12	1909±23,		2631±51
13	1952±81,	2271,	3194
14	2007,	2770,	3279±189
15	2410,	2714,	3319±88
16	2655±46,		3746±62

**Table 5.7: Collision cross sections of monomeric p53-Mdm2 complex obtained via the DT IM-MS from buffered solution (50 mM ammonium acetate). Values are an average of two replicates and stated with the standard deviation of the mean.**

Collision cross-section ranges measured for the complex and individual components are plotted in Figure 5.22. Cross-section values for the complex range approximately between 1450 to 3750  $\text{\AA}^2$ , for p53-N between 1150 - 1730  $\text{\AA}^2$  and for Mdm2-N between 1050 - 2450  $\text{\AA}^2$ , respectively. Remarkably it appears that ‘shape’ of the complex is configured by the Mdm2 protein, which has a molecular weight only twenty five percent larger than of p53-N.



**Figure 5.22: Comparison of collision cross sections versus charge state measured for monomeric p53-N (grey triangles), Mdm2-N (green circles) and p53-Mdm2 complex (black squares).**

The Mdm2-p53 interaction is an example of a protein-protein interaction that is a good candidate for inhibition by a small, drug-like ligand (47). Precise structural understanding of this compact protein-protein interface will hopefully lead to a successful identification of a more specific and potent small molecule drug-like antagonist for use in anticancer therapy.

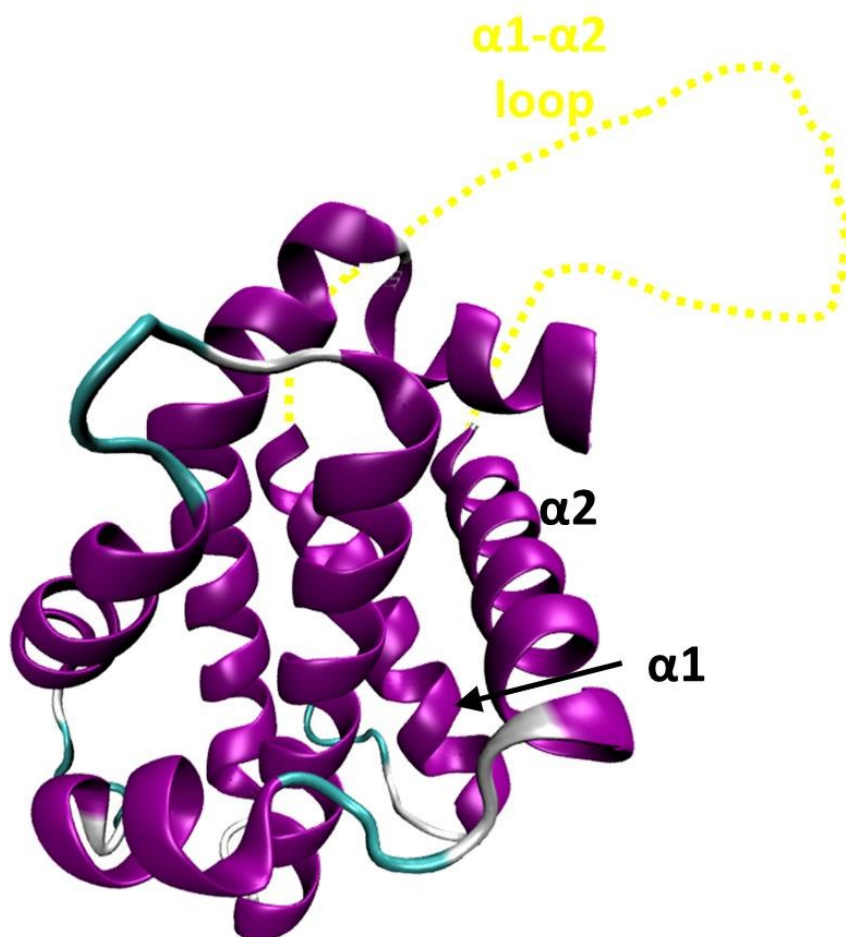
## **5.6. BcL-xL-p53 Connection: Alternative Route to Cell Death**

The basis for p53's outstanding apoptotic and tumour suppressive effectiveness lies in its pleiotropism that comprises transcription-dependant and -independent functions. This p53 transcription-independent apoptosis precedes p53 target gene activation and is mediated by the direct interaction of p53 with Bcl-2 family members, the anti-apoptotic members BcL-2 and BcL-xL, and the pro-apoptotic Bax, Bak, and Bad. p53 is both a transcription factor and a membrane permeability regulator. From genetic and biochemical evidence it has been proposed that p53 engages apoptotic caspase activation mainly through the intrinsic, mitochondrial pathway (48). This pathway is regulated by Bcl-2 family proteins, which govern the permeability of the outer mitochondrial membrane and the release of death factors such as cytochrome c. Chipuk *et al.* (49) demonstrated that PUMA, a proapoptotic BH3-only protein, can release p53 from the p53-BcL-xL complex to bind and activate Bax or Bak, in consequence allowing Bax and Bak to prompt mitochondrial permeabilization. Another pro-apoptotic protein Bad was shown to be upregulated by p53 and form a complex with p53 at the mitochondria inducing apoptosis (50). This example describes the function of cytoplasmic p53, which integrates the functions of p53 with the Bcl-2 family proteins, and supports the hypothesis that the complex formation of p53 and BcL-xL is of high significance in p53 transcription-independent apoptotic activity.

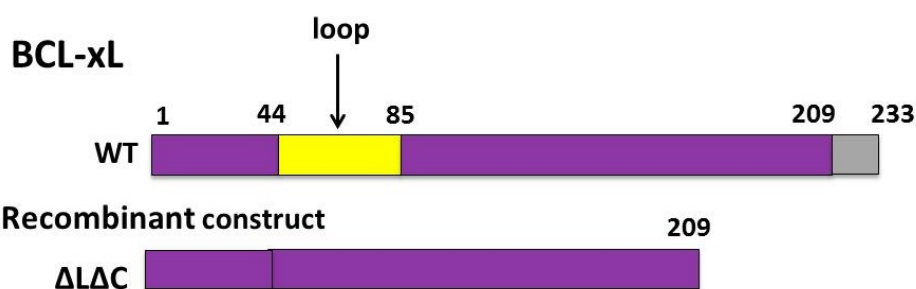
The critical interaction between p53 and Bcl-2 family members in response to apoptotic stimuli has aroused much interest and detailed information on the mitochondrial p53 apoptotic mechanism is still accumulating (50-55). Here, in the



mass spectrometry study for the molecular interaction between BcL-xL (Figure 5.23 shows the overall structure) and N-terminal p53 (residues 1-102), a BcL-xL $\Delta$ L $\Delta$ C construct was used (graphically presented in domain structure in Figure 5.24), in which the C-terminal transmembrane twenty two residues as well as the unstructured loop region connecting  $\alpha$ -helix 1 ( $\alpha$ 1) to  $\alpha$ 2 (from residues 44-55) was truncated.



**Figure 5.23:** X-Ray crystal structure of BcL-xL (PDB 1R2D) (56) adapted to display the long disordered loop region (in yellow) connecting  $\alpha$ -helix 1 to  $\alpha$ -helix 2.



**Figure 5.24: Domain structure of wild-type (WT) Bcl-xL highlighting the disordered loop region connecting  $\alpha 1$  to  $\alpha 2$ ; Recombinant construct Bcl-xL $\Delta L\Delta C$  lacking the loop region as well as the twenty two amino acid residues at the C-terminus.**

### **5.6.1. MS of Bcl-xL $\Delta L\Delta C$ in buffered solution conditions**

The observed spectrum of Bcl-xL $\Delta L\Delta C$  (Figure 5.25) displays a wide charge state distribution for a monomer,  $7 \leq z \leq 17$  (calculated  $M_R = 20780.9$  Da, observed  $M_R = 20745.1$  Da). Under buffered solution conditions the dominant spectral peak can be assigned as  $[M+8H]^{8+}$ , indicative of a highly compact conformation. In addition a less intense ion envelope for  $[2M+zH]^z$ , is also observed (where  $10 \leq z \leq 13$ ).

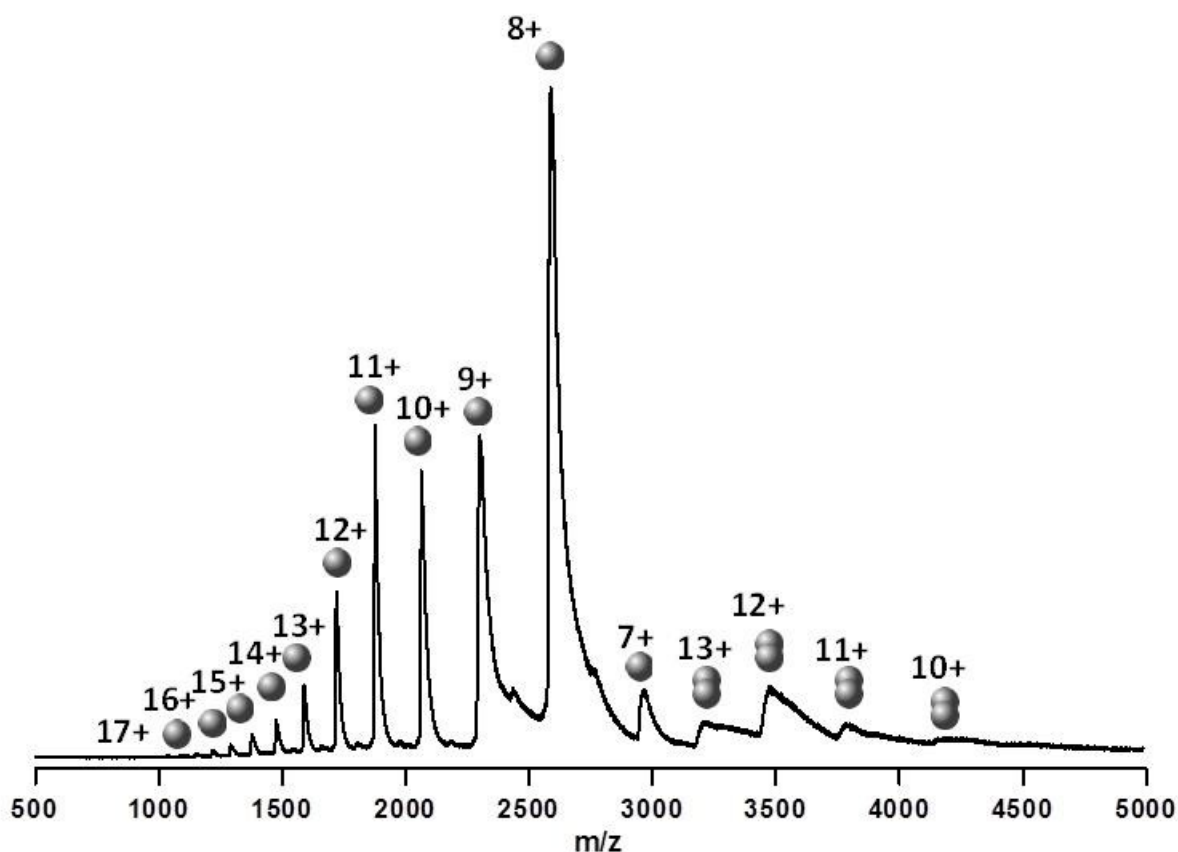
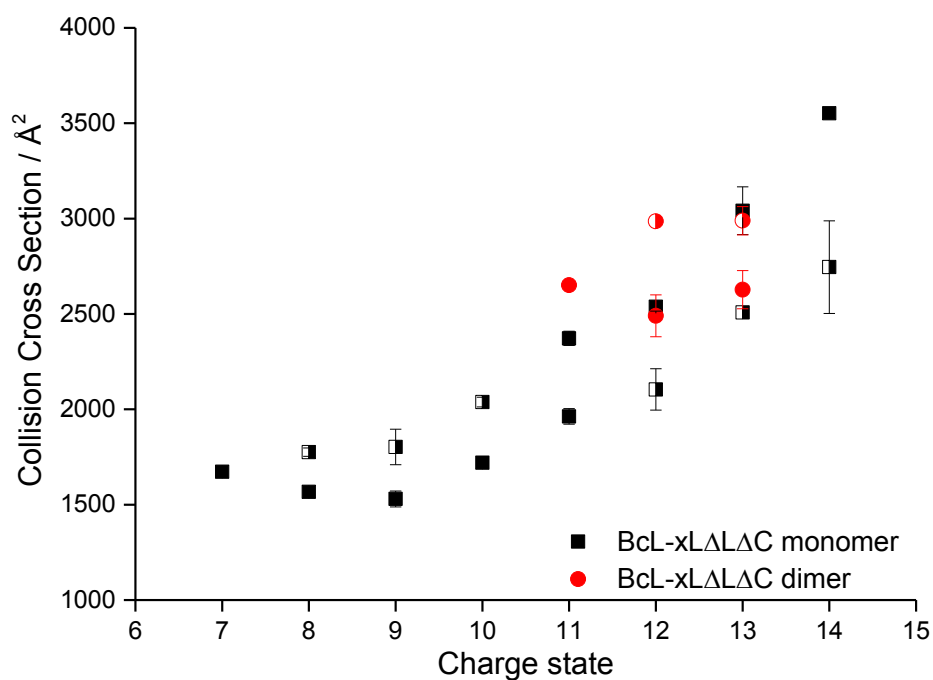


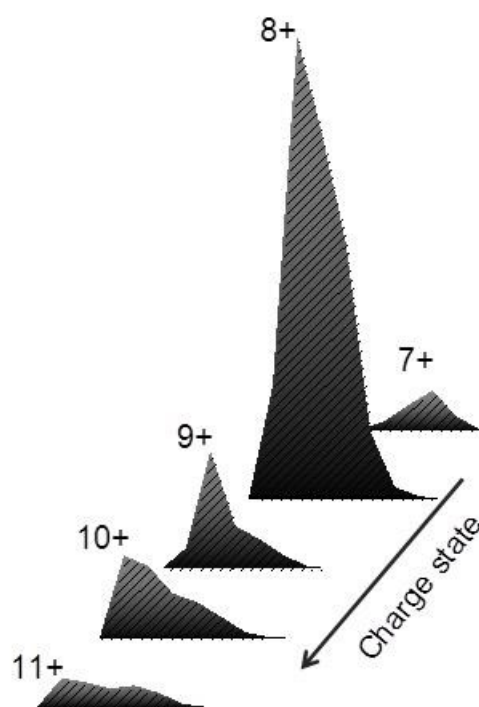
Figure 5.25: n-ESI spectrum of 135  $\mu\text{M}$  BcL-xL $\Delta$ L $\Delta$ C obtained via the DT IM-MS from buffered solution with 50 mM ammonium acetate.

### 5.6.2. Collision cross sections of BcL-xL $\Delta$ L $\Delta$ C

CCS values obtained for the monomeric and dimeric populations are tabulated in Table 5.8 and 5.9 and displayed graphically as a function of charge state in Figure 5.26. The cross-section is observed to increase upon addition of charge and is attributed to unfolding of the protein structure to minimise Coulomb repulsion (34). However the data also displays some regions of stability seen as ‘shelving’ in the CCS values for the lowest charge states ( $7 \leq z \leq 9$ ), providing evidence for several stable populations on the unfolding pathway.



**Figure 5.26:** Collision cross sections versus charge obtained on the DT IM-MS of monomer and dimer form of 135  $\mu$ M BcL-xL $\Delta\Delta\Delta$ C from buffered solution with 50 mM ammonium acetate. The half-filled symbol represents less abundant conformer at a given charge state.



**Figure 5.27:** Experimental ATDs taken at 35 V of the most intense charge states ( $z = 7 - 11$ ) for monomer form of BcL-xL $\Delta\Delta\Delta$ C.

The lowest charge state,  $z = 7$  provides a single resolvable conformer, with a CCS of  $1673 \text{ \AA}^2$ , whereas for  $z = 10$  two species are resolved, with the most populated being  $1567 \text{ \AA}^2$  and a lower intensity elongated form of  $1775 \text{ \AA}^2$  (shown in the ATDs in Figure 5.27) With increasing charge two conformational isomers can be elucidated for each charge state which indicates that multiple coexisting conformations are produced. At the highest charge,  $z = 14$ , the observed collision cross-section continues to rise in a near linear manner, CCS  $2746 \text{ \AA}^2$  and  $3553 \text{ \AA}^2$  are obtained. The majority of the protein for  $z = 7$ -10 is found in a compact conformational isomer, which also exists for  $z = 11$ -14, although here the more extended conformational family is the more pronounced. CCS obtained for dimers are on average only 14% larger than CCS of corresponding monomeric conformers (even less for  $z = 13$  where the difference is 4%), suggestive of significant collapse of the dimeric structure in the gas phase.

Charge state	Collision Cross Section / $\text{\AA}^2$	
7		1673 $\pm$ 1
8	1567 $\pm$ 10,	1775 $\pm$ 22
9	1530 $\pm$ 42,	1803 $\pm$ 93
10	1720 $\pm$ 2,	2038 $\pm$ 25
11	1963 $\pm$ 41,	2372 $\pm$ 37
12	2104 $\pm$ 108,	2538 $\pm$ 23
13	2508,	3041 $\pm$ 125
14	2746 $\pm$ 243,	3553

**Table 5.8: Collision cross sections of monomeric BcL-xL $\Delta$ L $\Delta$ C obtained via the DT IM-MS from 135  $\mu$ M solution buffered with 50 mM ammonium acetate. Values are an average of two replicates and stated with the standard deviation of the mean.**

Charge state	Collision Cross Section / $\text{\AA}^2$	
11		2651 $\pm$ 12
12	2490 $\pm$ 110,	2986
13	2627 $\pm$ 101,	2989 $\pm$ 74

**Table 5.9: Collision cross sections of dimeric BcL-xL $\Delta$ L $\Delta$ C obtained via the DT IM-MS from 135  $\mu$ M solution buffered with 50 mM ammonium acetate. Values are an average of two replicates and stated with the standard deviation of the mean.**

### 5.6.3. Molecular Interaction of N-terminal Domain of p53 with anti-apoptotic BcL-xL

A number of researchers have postulated that the anti-apoptotic protein BcL-xL interacts only with the DNA-binding region of p53 not with its N- and C-terminal regions, and subsequently the proposed molecular interaction has been defined by NMR spectroscopy (51,57-59). Controversially, Chipuk *et al.* (60) reported that the amino terminus of p53 including its proline-rich regulatory domain is necessary and sufficient to induce transcription-independent apoptosis. Further to that Xu *et al.* (61) characterised the binding interface between p53-N and BcL-xL by employing NMR-based binding study and a FRET assay in cells.

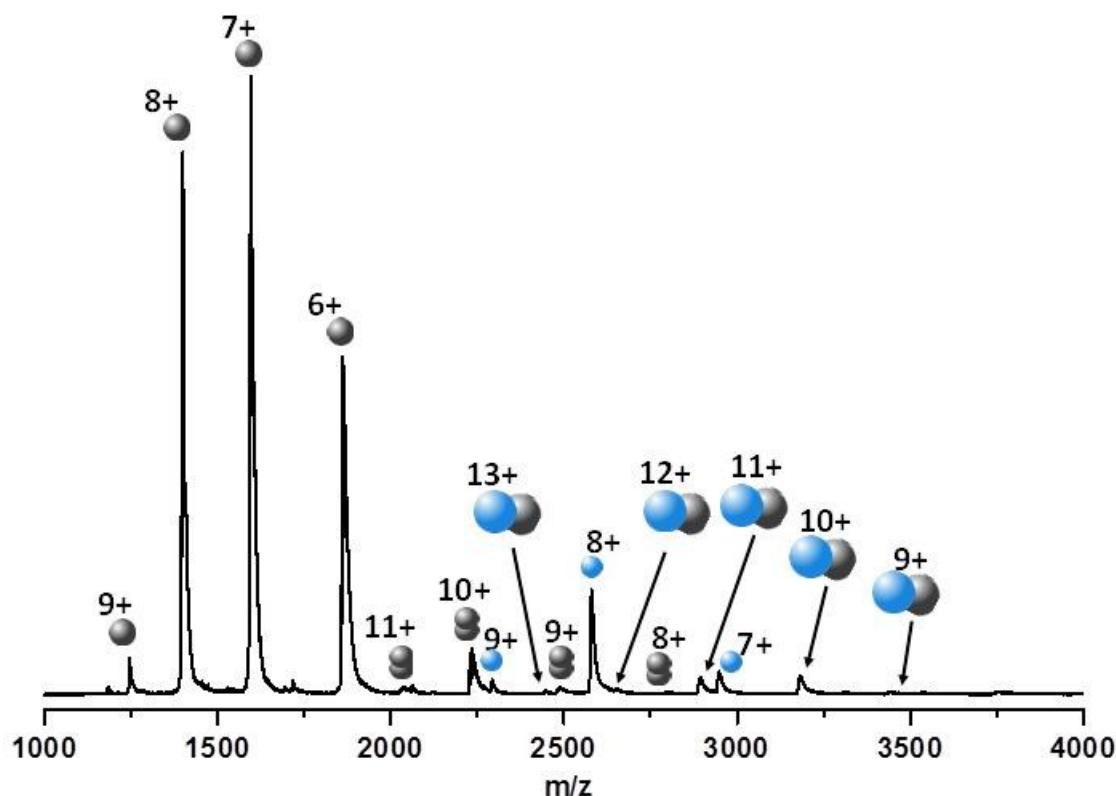


Figure 5.28: n-ESI spectrum of 180  $\mu$ M p53-N with 135  $\mu$ M BcL-xL $\Delta\Delta\Delta$ C. Mass spectrum was obtained from solution buffered with 50 mM ammonium acetate, incubated for five minutes and recorded via the DT IM-MS. The single grey spheres represent p53-N monomer, whereas the single blue spheres represent BcL-xL $\Delta\Delta\Delta$ C monomer. The mixed dimers spheres (blue and grey) represent the p53-N-BcL-xL $\Delta\Delta\Delta$ C complex.

Mass spectrometry was used to study the complexation between p53-N ( $M_R = 11162.3$  Da) and truncated construct BcL-xL $\Delta\Delta\Delta$ C ( $M_R = 20745.1$  Da) (Figure 5.28 shows the produced mass spectrum). The preliminary data showed that p53-N binds BcL-xL $\Delta\Delta\Delta$ C to form 1:1, 31.9 kDa complexes only, where  $z = 9-13$ . Notable comparable charge states were found for the p53-Mdm2 complex (spectrum shown in Figure 5.19, Section 5.5.2) which is not surprising as the Mdm2-binding region in the transactivation domain of p53 has also been identified as BcL-xL-binding motif (62). In the spectrum shown in Figure 5.21, ion envelopes of unbound ensembles for p53-N and BcL-xL $\Delta\Delta\Delta$ C were also detected. There is also a notable depletion of all charge states corresponding to BcL-xL $\Delta\Delta\Delta$ C in the presence of p53-N binding partner. These initial results on p53-N-BcL-xL interaction could perhaps be improved by lengthening of the incubation time or by employing different recombinant BcL-xL constructs, in particular wild-type BcL-xL. A longer incubation time would potentially results in strengthened binding (and a decreased amount of unbound proteins).

In conclusion, this study provides introductory information for the elucidation of the interaction between p53 and BcL-xL, further work is necessary to determine the mechanism and the roles they play in the p53-mediated transcription-independent apoptotic functions.

## 5.7. Summary

The past years have brought notable advancements in the understanding of the structure and the dynamic behaviour of p53 protein. The challenge now is to visualise the structure of multi-domain and full-length p53 in complex with its interacting proteins, which will provide further insights into the regulation of p53 activity via its intrinsically disordered amino- terminus and the crosstalk between different domains. The characterisation and quantification of p53 protein-protein interactions is essential for understanding the nature and relevance of these interactions under physiological and pathological conditions. An important and much needed contribution comes from elucidating how the function of a higher-order

complex diverges from that of its individual modules, and accurately determining the role of each subunit in terms of its identity, stoichiometry, structure, and binding constant, and how these might be affected by post-translational modifications. Mass spectrometry in tandem with ion mobility mass spectrometry captures the multiple populations of protein assemblies and is capable of providing snapshots when monitoring the overall activity snapshots in the biogenesis pathway of the protein complex. Combining such data from *in vivo* with that of cellular protein levels will provide a much clearer picture of p53 interactome.



## 5.8. References

1. Collavin, L., Lunardi, A. and Del Sal, G. (2010) p53-family proteins and their regulators: hubs and spokes in tumor suppression. *Cell Death & Differentiation*, **17**, 901-911.
2. McCullough, B.J., Kalapothakis, J., Eastwood, H., Kemper, P., MacMillan, D., Taylor, K., Dorin, J. and Barran, P.E. (2008) Development of an ion mobility quadrupole time of flight mass spectrometer. *Analytical Chemistry*, **80**, 6336-6344.
3. Chang, J., Kim, D.H., Lee, S.W., Choi, K.Y. and Sung, Y.C. (1995) Transactivation ability of p53 transcriptional activation domain is directly related to the binding-affinity to tata-binding protein. *The Journal of Biological Chemistry*, **270**, 25014-25019.
4. Walker, K.K. and Levine, A.J. (1996) Identification of a novel p53 functional domain that is necessary for efficient growth suppression. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 15335-15340.
5. Sakamuro, D., Sabbatini, P., White, E. and Prendergast, G.C. (1997) The polyproline region of p53 is required to activate apoptosis but not growth arrest. *Oncogene*, **15**, 887-898.
6. Kusano, K., Sakaguchi, M., Kagawa, N., Waterman, M.R. and Omura, T. (2001) Microsomal p450s use specific proline-rich sequences for efficient folding, but not for maintenance of the folded structure. *Journal of Biochemistry*, **129**, 259-269.
7. Wang, J., Tan, N.S., Ho, B. and Ding, J.L. (2002) Modular arrangement and secretion of a multidomain serine protease. Evidence for involvement of proline-rich region and N-glycans in the secretion pathway. *The Journal of Biological Chemistry*, **277**, 36363-36372.
8. Lu, H. and Levine, A.J. (1995) Human taf(ii)31 protein is a transcriptional coactivator of the p53 protein. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 5154-5158.
9. Thut, C.J., Chen, J.L., Klemm, R. and Tjian, R. (1995) p53 transcriptional activation mediated by coactivators taf(ii)40 and taf(ii)60. *Science*, **267**, 100-104.
10. Di Lello, P., Jenkins, L.M., Jones, T.N., Nguyen, B.D., Hara, T., Yamaguchi, H., Dikeakos, J.D., Appella, E., Legault, P. and Omichinski, J.G. (2006) Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. *Molecular Cell*, **22**, 731-740.
11. Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A.J. and Pavletich, N.P. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science*, **274**, 948-953.
12. Schon, O., Friedler, A., Bycroft, M., Freund, S.M.V. and Fersht, A.R. (2002) Molecular mechanism of the interaction between MDM2 and p53. *Journal of Molecular Biology*, **323**, 491-501.

13. Marine, J.C. and Jochemsen, A.G. (2005) Mdmx as an essential regulator of p53 activity. *Journal of Biochemical and Biophysical Research Communications*, **331**, 750-760.
14. Teufel, D.P., Bycroft, M. and Fersht, A.R. (2009) Regulation by phosphorylation of the relative affinities of the N-terminal transactivation domains of p53 for p300 domains and Mdm2. *Oncogene*, **28**, 2112-2118.
15. Gu, W., Shi, X.L. and Roeder, R.G. (1997) Synergistic activation of transcription by CBP and p53. *Nature*, **387**, 819-823.
16. Bell, S., Klein, C., Muller, L., Hansen, S. and Buchner, J. (2002) p53 contains large unstructured regions in its native state. *Journal of Molecular Biology*, **322**, 917-927.
17. Dawson, R., Muller, L., Dehner, A., Klein, C., Kessler, H. and Buchner, J. (2003) The N-terminal domain of p53 is natively unfolded. *Journal of Molecular Biology*, **332**, 1131-1141.
18. Lee, H., Mok, K.H., Muhandiram, R., Park, K.H., Suk, J.E., Kim, D.H., Chang, J., Sung, Y.C., Choi, K.Y. and Han, K.H. (2000) Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *The Journal of Biological Chemistry*, **275**, 29426-29432.
19. Bochkareva, E., Kaustov, L., Ayed, A., Yi, G.S., Lu, Y., Pineda-Lucena, A., Liao, J.C., Okorokov, A.L., Milner, J., Arrowsmith, C.H. *et al.* (2005) Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15412-15417.
20. Rosal, R., Pincus, M.R., Brandt-Rauf, P.W., Fine, R.L., Michl, J. and Wang, H. (2004) NMR solution structure of a peptide from the mdm-2 binding domain of the p53 protein that is selectively cytotoxic to cancer cells. *Biochemistry*, **43**, 1854-1861.
21. Grasberger, B.L., Lu, T., Schubert, C., Parks, D.J., Carver, T.E., Koblish, H.K., Cummings, M.D., LaFrance, L.V., Milkiewicz, K.L., Calvo, R.R. *et al.* (2005) Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. *Journal of Medicinal Chemistry*, **48**, 909-912.
22. Sakurai, K., Schubert, C. and Kahne, D. (2006) Crystallographic analysis of an 8-mer p53 peptide analogue complexed with MDM2. *Journal of American Chemical Society*, **128**, 11000-11001.
23. Popowicz, G.M., Czarna, A., Rothweiler, U., Szwagierczak, A., Krajewski, M., Weber, L. and Holak, T.A. (2007) Molecular basis for the inhibition of p53 by Mdmx. *Cell Cycle*, **6**, 2386-2392.
24. Dunker, A.K., Silman, I., Uversky, V.N. and Sussman, J.L. (2008) Function and structure of inherently disordered proteins. *Journal of Current Opinion in Structural Biology*, **18**, 756-764.
25. Gsponer, J. and Babu, M.M. (2009) The rules of disorder or why disorder rules. *Prog. Biophys. Mol. Biol.*, **99**, 94-103.
26. Joerger, A.C. and Fersht, A.R. (2010) The tumor suppressor p53: from structures to drug discovery. *Cold Spring Harbor Perspectives in Biology*, **2**, a000919.

27. Sakaguchi, K., Saito, S.i., Higashimoto, Y., Roy, S., Anderson, C.W. and Appella, E. (2000) Damage-mediated phosphorylation of human p53 threonine 18 through a cascade mediated by a casein 1-like kinase: Effect on Mdm2 binding. *The Journal of Biological Chemistry*, **275**, 9278-9283.
28. Brown, C.J., Srinivasan, D., Jun, L.H., Coomber, D., Verma, C.S. and Lane, D.P. (2008) The electrostatic surface of MDM2 modulates the specificity of its interaction with phosphorylated and unphosphorylated p53 peptides. *Cell Cycle*, **7**, 608-610.
29. Ferreón, J.C., Lee, C.W., Arai, M., Martinez-Yamout, M.A., Dyson, H.J. and Wright, P.E. (2009) Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 6591-6596.
30. Jenkins, L.M.M., Yamaguchi, H., Hayashi, R., Cherry, S., Tropea, J.E., Miller, M., Wlodawer, A., Appella, E. and Mazur, S.J. (2009) Two distinct motifs within the p53 transactivation domain bind to the Taz2 domain of p300 and are differentially affected by phosphorylation. *Biochemistry*, **48**, 1244-1255.
31. Liu, J.G., Perumal, N.B., Oldfield, C.J., Su, E.W., Uversky, V.N. and Dunker, A.K. (2006) Intrinsic disorder in transcription factors. *Biochemistry*, **45**, 6873-6888.
32. Natan, E., Baloglu, C., Pagel, K., Freund, S.M., Morgner, N., Robinson, C.V., Fersht, A.R. and Joerger, A.C. (2011) Interaction of the p53 DNA-binding domain with its n-terminal extension modulates the stability of the p53 tetramer. *Journal of Molecular Biology*, **409**, 358-368.
33. Jurneczko, E., Cruickshank, F., Porrini, M., Nikolova, P., Campuzano, I.D., Morris, M. and Barran, P.E. (2012) Intrinsic disorder in proteins: a challenge for (un)structural biology met by ion mobility-mass spectrometry. *Biochemical Society Transactions*, **40**, 1021-1026.
34. Valentine, S.J., Counterman, A.E. and Clemmer, D.E. (1997) Conformer-dependent proton-transfer reactions of ubiquitin ions. *Journal of the American Society for Mass Spectrometry*, **8**, 954-961.
35. Clemmer, D.E., Hudgins, R.R. and Jarrold, M.F. (1995) Naked protein conformations: cytochrome c in the gas phase. *Journal of the American Chemical Society*, **117**, 10141-10142.
36. Nikolova, P.V., Henckel, J., Lane, D.P. and Fersht, A.R. (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14675-14680.
37. Brachmann, R.K., Yu, K., Eby, Y., Pavletich, N.P. and Boeke, J.D. (1998) Genetic selection of intragenic suppressor mutations that reverse the effect of common p53 cancer mutations. *Embo Journal*, **17**, 1847-1859.
38. Joerger, A.C., Allen, M.D. and Fersht, A.R. (2004) Crystal structure of a superstable mutant of human p53 core domain - Insights into the mechanism of rescuing oncogenic mutations. *The Journal of Biological Chemistry*, **279**, 1291-1296.
39. Uhrinova, S., Uhrin, D., Powers, H., Watt, K., Zheleva, D., Fischer, P., McInnes, C. and Barlow, P.N. (2005) Structure of free MDM2 N-terminal

- domain reveals conformational adjustments that accompany p53-binding. *Journal of Molecular Biology*, **350**, 587-598.
40. <http://www.disprot.org/pondr-fit.php>.
  41. McCoy, M.A., Gesell, J.J., Senior, M.M. and Wyss, D.F. (2003) Flexible lid to the p53-binding domain of human Mdm2: implications for p53 regulation. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 1645-1648.
  42. Funari, G., Domenici, F., Nardinocchi, L., Puca, R., D'Orazi, G., Bizzarri, A.R. and Cannistraro, S. (2010) Interaction of p53 with Mdm2 and azurin as studied by atomic force spectroscopy. *Journal of Molecular Recognition*, **23**, 343-351.
  43. Baumgartner, W., Hinterdorfer, P., Ness, W., Raab, A., Vestweber, D., Schindler, H. and Drenckhahn, D. (2000) Cadherin interaction probed by atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 4005-4010.
  44. Hanley, W., McCarty, O., Jadhav, S., Tseng, Y., Wirtz, D. and Konstantopoulos, K. (2003) Single molecule characterization of P-selectin/ligand binding. *The Journal of Biological Chemistry*, **278**, 10556-11061.
  45. Haupt, Y., Maya, R., Kazaz, A. and Oren, M. (1997) Mdm2 promotes the rapid degradation of p53. *Nature*, **387**, 296-299.
  46. Chen, H.-F. and Luo, R. (2007) Binding Induced Folding in p53-Mdm2 Complex. *Journal of the American Chemical Society*, **129**, 2930-2937.
  47. Dickens, M.P., Fitzgerald, R. and Fischer, P.M. (2010) Small-molecule inhibitors of MDM2 as new anticancer therapeutics. *Seminars in Cancer Biology*, **20**, 10-18.
  48. Schuler, M. and Green, D.R. (2001) Mechanisms of p53-dependent apoptosis. *Biochemical Society Transactions*, **29**, 684-688.
  49. Chipuk, J.E., Bouchier-Hayes, L., Kuwana, T., Newmeyer, D.D. and Green, D.R. (2005) PUMA couples the nuclear and cytoplasmic proapoptotic function of p53. *Science*, **309**, 1732-1735.
  50. Jiang, P., Du, W., Heese, K. and Wu, M. (2006) The Bad guy cooperates with good cop p53: Bad is transcriptionally up-regulated by p53 and forms a Bad/p53 complex at the mitochondria to induce apoptosis. *Molecular Cell Biology*, **26**, 9071-9082.
  51. Mihara, M., Erster, S., Zaika, A., Petrenko, O., Chittenden, T., Pancoska, P. and Moll, U.M. (2003) p53 has a direct apoptogenic role at the mitochondria. *Molecular Cell*, **11**, 577-590.
  52. Leu, J.I., Dumont, P., Hafey, M., Murphy, M.E. and George, D.L. (2004) Mitochondrial p53 activates Bak and causes disruption of a Bak-Mcl1 complex. *Nature Cell Biology*, **6**, 443-450.
  53. Chipuk, J.E., Kuwana, T., Bouchier-Hayes, L., Droin, N.M., Newmeyer, D.D., Schuler, M. and Green, D.R. (2004) Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis. *Science*, **303**, 1010-1014.
  54. Vaseva, A.V. and Moll, U.M. (2009) The mitochondrial p53 pathway. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, **1787**, 414-420.

55. Han, J., Goldstein, L.A., Hou, W., Gastman, B.R. and Rabinowich, H. (2010) Regulation of Mitochondrial Apoptotic Events by p53-mediated Disruption of Complexes between Antiapoptotic Bcl-2 Members and Bim. *The Journal of Biological Chemistry*, **285**, 22473-22483.
56. Manion, M.K., O'Neill, J.W., Giedt, C.D., Kim, K.M., Zhang, K.Y. and Hockenbery, D.M. (2004) Bcl-XL mutations suppress cellular sensitivity to antimycin A. *The Journal of Biological Chemistry*, **279**, 2159-2165.
57. Sot, B., Freund, S.M. and Fersht, A.R. (2007) Comparative biophysical characterization of p53 with the pro-apoptotic BAK and the anti-apoptotic BCL-xL. *The Journal of Biological Chemistry*, **282**, 29193-29200.
58. Petros, A.M., Gunasekera, A., Xu, N., Olejniczak, E.T. and Fesik, S.W. (2004) Defining the p53 DNA-binding domain/Bcl-x(L)-binding interface using NMR. *Febs Letters*, **559**, 171-174.
59. Hagn, F., Klein, C., Demmer, O., Marchenko, N., Vaseva, A., Moll, U.M. and Kessler, H. (2010) BclxL changes conformation upon binding to wild-type but not mutant p53 DNA binding domain. *The Journal of Biological Chemistry*, **285**, 3439-3450.
60. Chipuk, J.E., Maurer, U., Green, D.R. and Schuler, M. (2003) Pharmacologic activation of p53 elicits Bax-dependent apoptosis in the absence of transcription. *Cancer Cell*, **4**, 371-381.
61. Xu, H., Tai, J., Ye, H., Kang, C.B. and Yoon, H.S. (2006) The N-terminal domain of tumor suppressor p53 is involved in the molecular interaction with the anti-apoptotic protein Bcl-XL. *Journal of Biochemical and Biophysical Research Communications*, **341**, 938-944.
62. Xu, H., Ye, H., Osman, N.E., Sadler, K., Won, E.-Y., Chi, S.-W. and Yoon, H.S. (2009) The Mdm2-binding region in the transactivation domain of p53 also acts as a Bcl-xL-binding motif. *Biochemistry*, **48**, 12159-12168.

# 6

## Conclusions

*Mass spectrometry and ion mobility mass spectrometry have provided valuable insights into the conformational diversity of tumour suppressor, multi-domain transcription factor p53 protein. In this chapter the conclusions of this thesis are summarised.*

## 6. Conclusions

In the last ten years mass spectrometry in tandem with ion mobility mass spectrometry has blossomed as a powerful biophysical technique capable of providing unique insights into the structure and dynamics of proteins. It is an attractive tool in the elucidation of IDP structure, capturing detail on the conformational range of a polypeptide at a single conformer level, rather than averaged data on multiple conformations adopted by the protein, which is of central importance in assigning structure-function relationship to proteins that are conformationally mobile. This method can test the (many) biophysical assertions and predictions made about why they differ from structured proteins. For example, from the analysis of the relative intensities of ions in the charge state distribution presented by a protein following nano-electrospray ionization, it is possible to infer the range of conformational ensembles present in solution and hence the extent of disorder. Although it does not provide atomic resolution compared to NMR and X-ray crystallography the combination of mass spectrometry with ion mobility delivers rotationally averaged collision cross-sections of molecular ions which can be correlated with dynamic states of IDPs.

In this thesis, the structural features and disorder/function relationship for the multi-talented ‘guardian of the genome’ p53 protein were explored.

In *Chapter Three*, the central core of p53, the DNA-binding domain was put under the ‘magnifying glass’ of mass spectrometry. The plasticity of this domain was reflected in the existence of multiple conformational isomers, independent of any binding stimuli. Under ‘near native’ conditions a wide charge state distribution was detected, manifesting the signature behaviour of IDPs *in vacuo*. In the presence of DNA, p53-DBD was found to possess a remarkable self-organizing ability, binding to double-stranded 12mer DNA as a dimer and as ‘dimer of dimers’ to double-stranded 20mer DNA, to form 58 kDa and 110 kDa complexes. The importance of conformational diversity in binding was also highlighted.

In *Chapter Four*, the *in vacuo* conformational phenotypes exhibited by most frequent single-point carcinogenic mutations of p53 DNA-binding domain: R249A, R273H, K292I, A276Y, and R175H were defined. The study on the second-site suppressor from loop L1, H115N, showed that this designed mutation inflicts a drastic change to the observed conformational fingerprint, revealing a very plastic and disordered nature of this mutant. In the double mutant approach, the conformational changes triggered by the H115N in the hotspot cancer mutations were also examined, detailed in the conformational spread afforded by the double mutants.

In *Chapter Five*, insights into the complex structure of p53 were provided by segmenting p53 into individual and multi-domain constructs, as well as interrogating the molecular interactions of p53 N-terminus with the most important p53 regulator, the E3 ubiquitin ligase Mdm2 and with BcL-xL involved in transcription-independent apoptotic role of p53. The prevalence of intrinsic structural flexibility of the N-terminal region of p53 was highlighted via the presence of two persistently stable conformers in the unfolding pathway. The intrinsically disordered N-terminus induced an increase in formation of more stable higher order multimeric species such as dimers and trimers, in the resultant spectrum of p53-N-DBD. The superstable quadruple mutant of the p53 DNA-binding domain was found to inhabit a diverse conformational landscape, and showed remarkable rigidity and clear resistance to Coulombically driven unfolding of the most compact form which was interpreted as a possible aftermath of the introduction of four stabilising mutations. Further, the investigation of p53-Mdm2 interaction revealed its extraordinary transient and conformationally dynamic character. Multiple conformations were resolved for the complex. The unfolding trajectory was found to lack any major stability regions, seen as ‘shelving’ of several charge states, and appeared to be more heterogeneous than that of individual domains of apo-Mdm2 or isolated p53.

The p53 system presented here, undoubtedly exhibits a number of mechanistic features common to all IDPs: promiscuous interactions, persistent flexibility, partial folding in isolation, folding upon binding, conformational diversity and extensive regulation by post-translational modifications. Notably, this study has provided insights into disorder/function relationships - the functional purposes as to why



disorder is utilised. IDPs are understood only to some extent on a structural level and it is expected that the 'jargon' of disorder/function relationship is much more extensive than is currently known. Further research of the physical properties of IDPs and revelation of their links with function are necessary to fully understand the mechanistic language of IDPs.

# Appendices

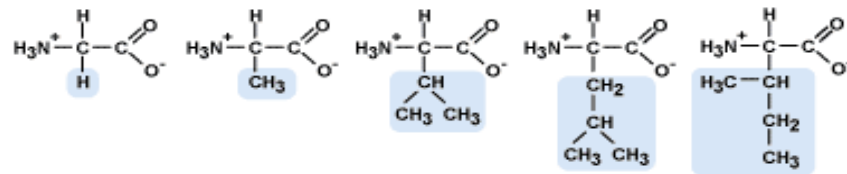
## Appendix 1: Amino Acid Abbreviations

Amino Acid	Three Letter Code	One Letter Code	Molecular Formula	Monoisotopic Mass	Average Mass
<b>Glycine</b>	Gly	G	C <sub>2</sub> H <sub>5</sub> NO <sub>2</sub>	54.021464	57.05
<b>Alanine</b>	Ala	A	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub>	71.037114	71.08
<b>Serine</b>	Ser	S	C <sub>3</sub> H <sub>7</sub> NO <sub>3</sub>	87.032029	87.08
<b>Proline</b>	Pro	P	C <sub>5</sub> H <sub>9</sub> NO <sub>2</sub>	97.052764	97.12
<b>Valine</b>	Val	V	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub>	99.068414	99.07
<b>Threonine</b>	Thr	T	C <sub>4</sub> H <sub>9</sub> NO <sub>3</sub>	101.04768	101.1
<b>Cysteine</b>	Cys	C	C <sub>3</sub> H <sub>7</sub> NO <sub>2</sub> S	103.00919	103.1
<b>Leucine</b>	Leu	L	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	113.08406	113.2
<b>Isoleucine</b>	Ile	I	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	113.08406	113.2
<b>Asparagine</b>	Asn	N	C <sub>4</sub> H <sub>8</sub> N <sub>2</sub> O <sub>3</sub>	114.04293	114.1
<b>Aspartic Acid</b>	Asp	D	C <sub>4</sub> H <sub>7</sub> NO <sub>4</sub>	115.02694	115.1
<b>Glutamic Acid</b>	Gln	Q	C <sub>5</sub> H <sub>9</sub> NO <sub>4</sub>	128.05858	128.1
<b>Lysine</b>	Lys	K	C <sub>6</sub> H <sub>14</sub> N <sub>2</sub> O <sub>2</sub>	128.09496	128.2
<b>Glutamine</b>	Glu	E	C <sub>5</sub> H <sub>9</sub> NO <sub>4</sub>	129.0496	129.1
<b>Methionine</b>	Met	M	C <sub>5</sub> H <sub>11</sub> NO <sub>2</sub> S	131.04048	131.2
<b>Histidine</b>	His	H	C <sub>6</sub> H <sub>9</sub> N <sub>3</sub> O <sub>2</sub>	137.05891	137.1
<b>Phenylalanine</b>	Phe	F	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	147.06841	147.2
<b>Arginine</b>	Arg	R	C <sub>6</sub> H <sub>14</sub> N <sub>4</sub> O <sub>2</sub>	156.10111	156.2
<b>Tyrosine</b>	Tyr	Y	C <sub>9</sub> H <sub>11</sub> NO <sub>3</sub>	163.06333	163.2
<b>Tryptophan</b>	Trp	W	C <sub>11</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	186.07931	186.2

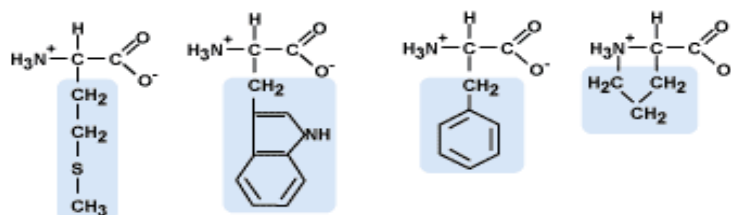
**Table A1:** Table of amino acids showing their three and one letter code, monoisotopic and average masses.

## Appendix 2: Amino Acid R Groups

### Non Polar Amino Acids

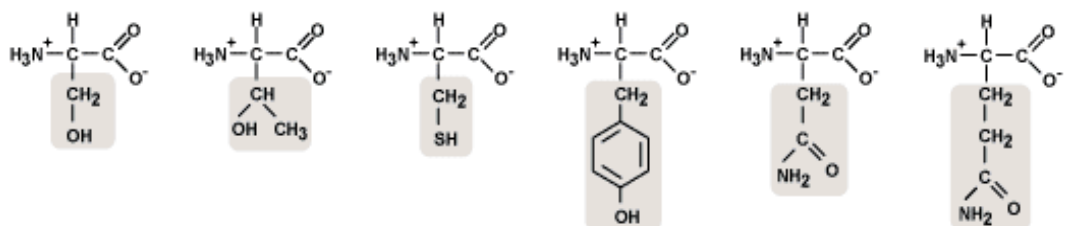


Glycine (Gly)    Alanine (Ala)    Valine (Val)    Leucine (Leu)    Isoleucine (Ile)



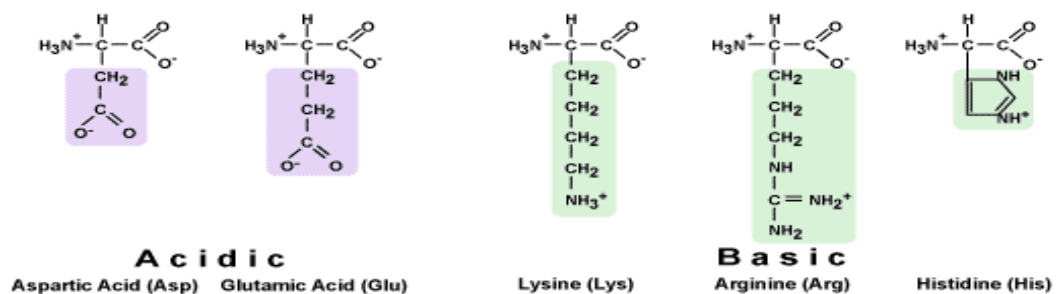
Methionine (Met)    Tryptophan (Trp)    Phenylalanine (Phe)    Proline (Pro)

### Polar Amino Acids



Serine (Ser)    Threonine (Thr)    Cysteine (Cys)    Tyrosine (Tyr)    Asparagine (Asn)    Glutamine (Gln)

### Charged Amino Acids



**Figure A1:** Figures of amino acids from the website of the Department of Biology, Penn State University: <https://wikispaces.psu.edu/display/bio110/Carbon+and+Life>

## Appendix 3: Properties of Drift Gases

Drift Gas	Mass / Da	Polarizability / $10^{-24} \text{ cm}^3$	Effective Radius* / Å
Helium	4.0026	0.205	1.03
Neon	20.1797	0.396	n.d.
Nitrogen	28.0061	1.740	1.73
Argon	39.9624	1.641	1.67

**Table A2:** Molecular mass, polarizability volume and effective radius for the neutral drift gases utilized in this study. Mass of longest lived isotopes. \* Rigid-sphere model. n.d.: not determined. Values for masses of longest lived isotopes were taken from CRC Handbook of Chemistry and Physics, 88<sup>th</sup> edition.

## Appendix 4: Metric System Unit Prefixes

Prefix	Symbol	Factor $10^n$
tera	T	$10^{12}$
giga	G	$10^9$
mega	M	$10^6$
kilo	k	$10^3$
hecto	h	$10^2$
deca	da	$10^1$
deci	d	$10^{-1}$
centi	c	$10^{-2}$
milli	m	$10^{-3}$
micro	$\mu$	$10^{-6}$
nano	n	$10^{-9}$
pico	p	$10^{-12}$
fento	f	$10^{-15}$
atto	a	$10^{-18}$

**Table A3:** Detailed list of metric system units, symbols, and prefixes.

## Appendix 5: Physical Constants, Symbols and Units

Symbol	Name	Unit
<b>e</b>	elementary charge	$1\ e = 1.6021765 \times 10^{-19}\ \text{C}$
<b>E</b>	electric field strength	$\text{V m}^{-1}$
<b>K</b>	mobility	$\text{m}^2 \text{V}^{-1} \text{s}^{-1}$
<b>K<sub>0</sub></b>	reduced mobility	$\text{m}^2 \text{V}^{-1} \text{s}^{-1}$
<b>k<sub>B</sub></b>	boltzman constant	$k_B = 1.3806488(13) \times 10^{-23}\ \text{J}$
<b>L</b>	distance	$\text{m}$
<b>M</b>	mass	$\text{kg}$
<b>MR</b>	relative molecular	$\text{Da}$
<b>N</b>	number density	$\text{m}^{-3}$
<b>P</b>	pressure	$\text{Pa}$ ( $1\ \text{Pa} = 10^{-5}\ \text{bar} = 9.8692 \times 10^{-6}\ \text{atm} = 7.5006 \times 10^{-3}\ \text{Torr} = 145.04 \times 10^{-6}\ \text{psi}$ )
<b>T</b>	temperature	$\text{K}$ ( $[\text{K}] = [^{\circ}\text{C}] + 273.15$ )
<b>t<sub>0</sub></b>	dead time	$\text{S}$
<b>t<sub>a</sub></b>	arrival time	$\text{S}$
<b>t<sub>d</sub></b>	drift time	$\text{S}$
<b>v</b>	velocity	$\text{m s}^{-1}$
<b>V</b>	potential difference	$\text{V}$
<b>v<sub>d</sub></b>	drift velocity	$\text{m s}^{-1}$
<b>μ</b>	reduced mass	$\text{kg}$
<b>ν</b>	RF	$\text{Hz} = \text{s}^{-1}$
<b>Ω</b>	momentum transfer integral	$\text{\AA}^2$ ( $\text{\AA}^2 = 1 \times 10^{-20}\ \text{m}^2$ ( $\text{\AA} = 1 \times 10^{-10}\ \text{m}$ ))

**Table A4:** Detailed list of selected physical constants (along with their symbols and units) that were used in this thesis.

## Appendix 6: Publications Arising from this Thesis

Jurneczko, E., Cruickshank, F., Porrini, M., Clarke, D.J., Campuzano, I.D.G., Morris, M., Nikolova, P.V. and Barran, P.E. (2013) Probing the conformational diversity of cancer-associated mutations in p53 with ion-mobility mass spectrometry. *Angewandte Chemie International Edition*, **52**, 4370-4374.

Jurneczko, E., Cruickshank, F., Porrini, M., Nikolova, P., Campuzano, I.D., Morris, M. and Barran, P.E. (2012) Intrinsic disorder in proteins: a challenge for (un)structural biology met by ion mobility-mass spectrometry. *Biochemical Society Transactions*, **40**, 1021-1026.

Jurneczko, E., Kalapothakis, J., Campuzano, I.D., Morris, M. and Barran, P.E. (2012) Effects of drift gas on collision cross sections of a protein standard in linear drift tube and traveling wave ion mobility mass spectrometry. *Analytical Chemistry*, **84**, 8524-8531.

Jurneczko, E. and Barran, P.E. (2011) How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst*, **136**, 20-28.