Investigating the Interactions and Conformations of Metalloproteins by Means of Mass Spectrometry

Frances D. L. Kondrat BSc (Hons)

A thesis submitted for the degree of Doctor of Philosophy

University of Warwick

School of Life Sciences

September 2012

“Riddles: they either delight or torment. Their delight lies in solutions. Answers provide bright moments of comprehension perfectly suited for children who still inhabit a world where solutions are readily available…..The adult world, however, produces riddles of a different variety. They do not have answers and are often called enigmas or paradoxes. Still the old hint of the riddle’s form corrupts these questions by re-echoing the most fundamental lesson: there must be an answer. From there comes torment.”

House of Leaves, Mark Z. Danielewski

Contents

List of Figures

List of Tables

Acknowledgements

Firstly I would like to thank both my supervisors Prof. Jim Scrivens and Dr. Claudia Blindauer, not only for providing me with the opportunity to carry out this research, but for the constant and invaluable support they provided me over the last few years. Thank you to the EPSRC and RSC who provided financial support for this project.

I recognise how lucky I am, to have been split between two amazing groups. Although from time to time, when walking down Gibbet Hill in the pouring rain or freezing cold I wished they were perhaps a bit closer together. Each and every one of you past and present helped me get this far. Most notably from chemistry Greg and Esther, who I worked closely with whilst conducting my ICP-OES work and were always receptive to my many questions. As for my Biology group, Kostas, Charlie, Sue, Gill, Elle, Nisha, Krisztina, Baharak, Matt, George, Eamonn and Jonathon, you are all great scientists and wonderful people, it was a pleasure to work with and learn from all of you. Many of you became great friends over the years, which made work days enjoyable even through the stressful times when things were not going as planned.

The nature of my research resulted in many instances in which I needed to use new techniques and instrumentation. I would therefore like to thank all the people who put their time and effort into training me. In particular I would like to thank Dr Kevin Purdy and Dr Francis Sweeney for guidance and help whilst using their anaerobic cabinet.

Mum, Simon and Alice I can not thank you enough, you have provided me with invaluable support when I needed it most. In particular, Mum you have been brilliant and have made an extremely stressful time so much easier. Hopefully once I am earning I can start spoiling you.

I would also like to thank Eddie; you have guided me through the last four years with surprising ease. Without your love, support and understanding (not of biochemistry I hasten to add, although you at least now know what a protein is!) I doubt I would have got this far or enjoyed it half as much.

Finally I would like to dedicate this thesis to my Dad. The limited time we got to spend together helped shape me into the person and scientist I am today. You are truly missed.

Declaration

I hereby declare that this thesis, submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy and entitled “Investigating the Interations and Conformations of Metalloproteins by Means of Mass Spectrometry”, represents my own work and has not been previously submitted to this or any other institution for any degree, diploma or other qualification. Work undertaken by my collaborators is explicitly stated where appropriate.

Frances Kondrat

September 2012

Summary

The use of mass spectrometry for the study of proteins and their non-covalent interactions has expanded rapidly since the introduction of ionisation techniques capable of transferring intact complexes into the gas phase. The coupling of the shape selective technique of ion mobility to mass spectrometry has advanced the ustility of these techniques for structural biology studies. The work presented in this thesis showcases the use of mass spectrometry for the study of proteins and their interaction partners including metal ions, proteins and DNA.

SmtB is a homodimeric, DNA binding protein capable of binding zinc ions with high affinity. SmtB functions as a transcriptional repressor in the absence of zinc ions but has the ability to reduce its DNA binding affinity upon a quaternary conformational change induced by the binding of zinc ions. In this fashion SmtB acts as a zinc sensor allowing the zinc chelator SmtA to be produced under zinc stress conditions. Mass spectrometry has been used to probe the zinc binding stoichiometry and monomer/ dimer equilibrium of a recombinant SmtB sample expressed in the presence of added zinc ions. These experiments confirmed that, in the presence of zinc ions, SmtB predominantly existed as a homodimer with four zinc ions bound. The complementary technique of ICP-OES confirmed that the zinc binding stoichiometry of the SmtB dimer was clearly higher than the value which had been presented in the literature.

A variety of different zinc ion removal techniques were implemented to investigate the removal of zinc ions from SmtB together with any associated change in monomer/dimer equilibrium and dimeric conformation. These were characterised by using mass spectrometry and ion mobility mass spectrometry respectively. The removal of zinc ions was found to promote expansion and unfolding of the dimer resulting in the production of monomeric species. The locations of the zinc binding sites were then studied using tandem mass spectrometry.

The interaction of apo-SmtB, produced by dialysis under anaerobic conditions, with a double stranded DNA oligomer containing a SmtB binding site, was successfully transferred into the gas phase. The resulting complex was found to exist in two different stoichiometries, $SmtB_2$:DNA and $2SmtB_2$:DNA, with no indication of a stepwise monomer binding mechanism.

The conformational changes associated with zinc ion removal from the metallothionein SmtA were also investigated utilising ion mobility mass spectrometry. These results revealed an increase in the flexibility of SmtA during zinc ion removal along with an increase of the conformational space occupied.

Abbreviations

A

A	Adenosine
Å	Angstrom
α	Alpha
AAS	Atomic absorption spectroscopy
ADC	Analogue-to-digital converter
Ag	Silver
Ar	Argon
As	Arsenic
ATD	Arrival time distribution
ATP	Adenosine Triphosphate
Au	Gold

B

β	Beta
Bi	Bismuth
bp	Base pair

C

C	Cytosine
CAD	Collisionally activated dissociation
CCS	Collisional cross section
CE	Collision energy
CI	Chemical ionisation
CID	Collisionally induced dissociation
CD	Circular dichroism
Cd	Cadmium
CDF	Cation diffusion facilitator
Co	Cobalt
CsI	Caesium iodide
Cu	Copper
CV	Compensation voltage

D

Da	Dalton
DC	Direct current

DE	Delayed extraction
DNA	Deoxyribonucleic acid
dsDNA	Double stranded deoxyribonucleic acid
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
DTT	Dithiothreitol

E

ECD	Electron capture dissociation
E. coli	Esherishia coli
EDC	Enhanced duty cycle
EDTA	Ethylenediaminetetraacetic acid
EHSS	Exact hard sphere scattering
EI	Electron impact ionisation
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
EXAFS	Extended X-ray absorption fine structure

F

FAB	Fast atom bombardment
FAIMS	Field assisted ion mobility spectroscopy
Fe	Iron
FI	Field ionisation
FPLC	Fast protein liquid chromatography
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width at half maximum

G

G	Guanine
g	Centrifugal acceleration

H

H^+	Proton
HDX	Hydrogen/deuterium exchange
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
Hg	Mercury
HMBD	Heavy metal binding domain
HPLC	High performance liquid chromatography
Hrs	Hours
HTH	Helix-turn-helix

I

ICC	Ion charge control
ICP-MS	Inductively coupled plasma mass spectrometry
ICP-OES	Inductively coupled plasma optical emission spectroscopy
IPTG	Isopropyl-beta-D-thiogalactopyranoside
IM	Ion mobility
IMMS	Ion mobility mass spectrometry
IMS	Ion mobility spectroscopy
IT	Ion trap

K

K_d	Dissociation constant
kDa	Kilodalton
kV	Kilovolt

L

L	Litres
LB	Liquid Broth

M

M	Molar
m	Metres per second
MALDI	Matrix assisted laser desorption ionisation
mbar	Millibar
MCP	Microchannel plate
MD	Molecular dynamics
mg	Milligram
MIC	Minimum inhibitory concentration
min	Minute
mL	Millilitre
mM	Millimolar
Mn	Manganese
m-NBA	m-nitrobenzyl alcohol
Mo	Molybdenum
mRNA	Messanger ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MT	Metallothionein

m/z	Mass over charge
µg	Microgram
µL	Microlitre
µM	Micromolar

N

NaCl	Sodium chloride
Ni	Nickel
nL	Nanolitre
nm	Nanometre
NMR	Nuclear magnetic resonance
ns	Nanoseconds
N-terminal/N-terminus	Amino-terminal/amino terminus

O

OD	Optical density
ORF	Open reading frame

P

PA	Projection approximation
PAR	Pyridyl-azo-resorcinol
Pb	Lead
PBS	Phosphate buffered saline
PDB	Protein data bank
pmol	Picomole
ppm	Parts per million
ps	Picosecond
PSA	Projected superposition approximation
PTR	Proton transfer reaction

Q

Q	Quadrupole
qDNA	Quadplex DNA

R

RF	Radio frequency
RNA	Ribonucleic acid
RND	Resistance nodulation cell
rpm	Revolutions per minute

S

s	Second
Sb	Animony
SEC	Size exclusion chromatography
SEM	Scanning electron microscope
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
S/N	Signal-to-noise
SRIG	Stacked ring ion guide
ssDNA	Single stranded deoxyribonucleic acid

T

T	Thymidine
TDC	Time-to-digital converter
TM	Trajectory method
TOF	Time of flight
Tris	Tris(hydroxymethyl)aminomethane
T-Wave	Travelling wave
TWIMS	Travelling wave ion mobility mass spectrometry

U

U	Uracil
UV	Ultra violet

V

V	Volt

Z

Zn	Zinc

Amino Acid	3 Letter code	1 Letter Code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic Acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Theronine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter One: Introduction

1.1 Biological molecules

Four main macromolecular systems may be recognised in biology, proteins, lipids, nucleic acids and carbohydrates. Each macromolecule is created from specific building blocks, have their own properties and serve a specific function. The structure of proteins and nucleic acids will be focussed on here.

1.1.1 Proteins

Proteins are the most well studied macromolecules due to their diverse function. Their functions include structure, transport, catalysis and regulation. Every protein is made up of a long chain of amino acids joined together via condensation reactions. The order and length of the amino acid chain gives the protein its primary structure. This primary structure encodes the interactions required to form secondary structural elements, such as alpha helices and beta sheets, which can then fold and collapse into the proteins tertiary structure. The specific function of many proteins is dictated by this structure. Some proteins, however, require more than one polypeptide chain to be functional and the interaction between these chains results in a quaternary structure. These folding events allow an extensive repertoire of proteins to be created from just twenty different amino acids.

1.1.2 Nucleic acids

There are two main groups of nucleic acids, ribonucleic acids (RNA) and deoxyribonucleic acids (DNA), both are polymers of nucleotides although the nucleotides differ between them.

Each RNA nucleotide is composed of one of four bases, adenine (A), cytosine (C), guanine (G), and uracil (U) joined to a ribose and a phosphate group. As polymers they are involved in protein synthesis (mRNA and tRNA) and ribosomal structure (rRNA).

DNA nucleotides are also made up of a sugar, in this instance deoxyribose, a phosphate group and one nucleobase either A, C, G or thymine (T) which DNA contains instead of U, in RNA. This chain of nucleotides (sense strand) then forms a

double helical structure with an antisense strand. This structure is stabilised by intrastrand base stacking and the formation of hydrogen bonds between Watson and Crick base pairs. The stability of the resulting structure depends on the base composition of the strands. Three hydrogen bonds are formed between C and G bps, whilst two hydrogen bonds are formed between A and T bps, thus a higher GC content increases the DNA stability.

The base content of DNA also can determine the type of helix if forms under certain conditions. There are three main helices associated with DNA with the most biologically common and recognisable helix being the right handed B-form, Table 1.1.

Table 1.1: Properties of the three forms of DNA helices, A, B and Z (Stryer 1995).

	Helix Type		
	A	B	Z
Helix diameter	25.5Å	23.7Å	18.4Å
Rise per base pair	2.3Å	3.4Å	3.8Å
Screw sense	Right-handed	Right-handed	Left-handed
Major groove	Narrow and very deep	Wide and quite deep	Flat
Minor groove	Very broad and shallow	Narrow and quite deep	Very narrow and deep

Conversions of this form to the A-form, another right handed helix but with different dimensions, Table 1.1, often occur in biology, in sequences with a high C·G content. This conversion can be recreated in vitro by exposing susceptible B- DNA to conditions of high salt/ low humidity < 75 %, however similar forms are observed in double stranded RNA and DNA: RNA hybrids (Stryer 1995). The left handed Z-DNA helix, Table 1.1, is also formed under similar in vitro conditions, although susceptible sequences are usually short with alternating purine and pyrimidine bases (Stryer 1995). To date, no biologically relevant conversion from B-form to Z-form has been demonstrated.

1.2 Metallomics

1.2.1 Metal ions in biology

Initially when thinking about biochemistry, the elements carbon, nitrogen, oxygen, hydrogen, sulphur and phosphorous come to mind. Notably because these are the main constituents of amino acids, nucleic acids, saccharides and lipids; the building blocks of a cell. Cellular systems, however, cannot function without vital trace elements; often metal ions, which perform roles that no other component could (Ma et al. 2009). Recently it was estimated that nearly a third of known proteins require metal ions (Waldron and Robinson 2009). Often these are d-block metal ions, Figure 1.1, and these have been shown to provide an extensive range of structural and functional roles within proteins (Reyes-Caballero et al.), with a smaller incidence being found within RNA (Fedor 2002).

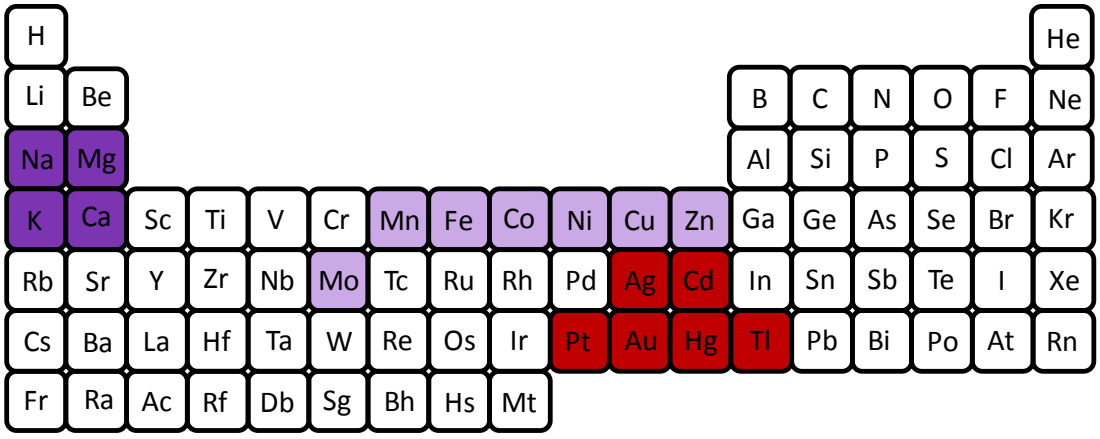


Figure 1.1: A periodic table highlighting the major and essential trace metal ions, dark and light purple respectively, required by biological systems. Metal elements highlighted in red, generally have no biological function and are highly toxic within biological cells. Adapted from (Nies and Silver 2007) and (Summers 2009).

Metal ions involved in structure

The binding of a variety of metal ions to polypeptides has been shown to assist with folding and/or stabilisation. These interactions help stabilise a protein's tertiary structure enabling specific functional motifs to be created and maintained. The extensively studied Cys_2His_2 zinc finger is one of the best examples of this. Without zinc ions the zinc finger peptide is unstructured and therefore biologically inactive. The introduction of zinc ions causes folding and stabilisation of the DNA binding

motif (Cox and McLendon 2000). In this case, although the zinc ion creates a functional motif, its role is purely structural. Metal ions not only stabilise secondary and tertiary structures, they can also play a pivotal role in keeping protein subunits together (quaternary structure), especially when their binding sites lie at protein: protein interfaces (Auld 2001).

Metal ions involved in catalysis

There are many other examples where metal ions function as catalytic cofactors. In these incidences if the metal ion is removed, the protein will still maintain its globular structure but will not be able to perform any enzymatic reaction. The simplest way in which metal ions act as catalysts is their ability to bind two ligands so that they are in close enough proximity to react. The ability of catalytic metal ions to accept and donate electrons means that they can also act as redox centres and Lewis acids. Bioinformatic analysis has identified a total of 13 different metal ions present within enzymes, with all six different classes of enzymes having between 36 and 59% of known structures containing them (Andreini et al. 2008).

1.2.2 Zinc ions in biology

Zinc, unlike many of the other d-block elements, exists in biology in only one oxidation state, Zn^{2+} and is one of the most commonly found biological trace metal ions, after iron. Its environmental availability to cells is imperative due to its requirement in cellular growth and development. Although common throughout prokaryotes and eukaryotes, levels of free zinc are still extremely low within cells. The majority of Zn^{2+} is predicted to be complexed to proteins, especially those involved in transcription and translation (Finney and O'Halloran 2003). The concentration of free zinc ions within an *E. coli* cell has been predicted, from zinc sensor binding constants, to be within the femtomolar range (Eide 2006). This concentration relates to less than one free zinc ion per cell, effectively ruling out the presence of a free Zn^{2+} pool (Finney and O'Halloran 2003). Zinc ions have the unique ability, amongst the d-block metals, to form a catalytic, cocatalytic, interface or structural centre depending on the subset of bound ligands. In the majority of examples zinc ions have a coordination number of four, due to their electronic configuration $[Ar]3d^{10}$, and adopt a tetrahedral geometry utilising sp^3 hybrid orbitals (Blencowe and Morby 2003; Leszczyszyn et al. 2010; Wang et al. 2010). As zinc

ions are borderline hard/soft acids they have the ability to form stable complexes with both hard ligands, such as nitrogen and oxygen atoms from His, Asp and Glu but also with soft ligands, such as sulphur atoms from Cys (Rensing et al. 2007; Leszczyszyn et al. 2010). The preference in catalytic sites is the inclusion of His whereas Cys is often found in structural sites. Often, however, the only difference between structural and catalytic sites is the presence of a ligating water molecule in the latter site instead of a full complement of ligands from protein residues (Auld 2001). The ability of Zn^{2+} to bind to a variety of ligands makes them a highly competitive species. A slight rise in free ions could, therefore, result in the binding of zinc to non-cognate proteins, leading to protein misfolding or the inactivation of enzymes.

1.2.3 Cadmium in biology

Cadmium, like zinc, is found in nature as a divalent metal ion. It shares homologous chemistry to zinc and is also classed as a non-transition element due to its completely filled d shell (Vallee and Ulmer 1972). The measured abundance of cadmium is lower than zinc in both the earths crust and seawater. Unlike zinc, Cd^{2+} is not commonly a biologically required trace element but is usually inherently toxic. Cadmium is a soft acid and therefore has a high affinity to soft ligands (cysteine thiolates) whereas it does not bind preferentially to hard ligands. As described above Zn^{2+} can bind to either soft or hard ligands but tends to have a lower affinity than Cd^{2+} for soft ligands. For this reason cadmium can be found bound in the place of zinc in metalloproteins, leading to either an increase, but more often a decrease in enzyme activity (Vallee and Ulmer 1972). Although predominantly toxic, there are now examples of native cadmium enzymes. The first isolated example was a carbonic anhydrase, CDCA1, from the marine organism *Thalassiosira weissflogii* (Lane et al. 2005). The transcription of *cdca1* was shown to increase upon cadmium exposure and a tetrahedral cadmium binding site was identified, within the purified protein. The appearance of cadmium within marine enzymes has been suggested to be due to selection pressure caused by the low levels of trace metals within sea water (Lane et al. 2005).

1.2.4 Metal ion toxicity

As the metal ion quota required by cells is relatively high, to ensure a sufficient supply for the plethora of proteins they contain, it would be easy to believe that cells would uptake as many metal ions as possible. From Figure 1.1 however, it can be seen that some metal ions are inherently toxic to cells, such as lead, mercury and in most instances cadmium. Furthermore even trace metal ions, such as Zn^{2+}, although vital to cellular systems are also toxic even at relatively low concentrations (Pennella and Giedroc 2005).

Trace metal ions can become toxic to a cell upon binding to the incorrect protein, rendering it either misfolded (Leal et al. 2012) or non-functional. It is therefore important to ensure that the correct metal binds to the correct protein. This challenge becomes even harder when the following is taken into consideration. Some metal ions form more stable complexes with ligands than others and although proteins have predetermined binding sites, they are flexible, often allowing non-cognate metal ions to bind. The Irving-Williams series reveals the complex stability for the first row of the divalent transition elements as, $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$ (Irving and Williams 1953). If metal ion binding was based solely on stability then Cu^{2+} would bind in place of all other divalent metal ions. A number of systems, explained in detail by Waldron and Robinson, have therefore been developed by bacteria to ensure stability is not the governing factor, reducing the likelihood of non-cognate binding. These are centred around metal ion homeostasis and include; folding proteins in different locations and the utilisation of metallochaperones, especially for copper (Waldron and Robinson 2009).

1.2.5 Homeostasis of metal ions in bacteria

The concentration of each metal ion required within a bacterial cell has to be maintained within a certain range. This enables the cell to provide a sufficient level of free metal ions to bind to newly synthesised proteins, but not an excess, which could lead to incorrect binding and eventually toxicity. The process of homeostasis maintains this balance and is dependent on four main areas; metal ion influx, efflux, intracellular trafficking and storage.

Metal ion influx

To reach the cytoplasm, the place in which all bacterial proteins are synthesised, any required metal ions need to cross the inner membrane in gram positive bacterial cells, or the outer and inner membrane in gram negative cells, see Figure 1.2. The latter of which will be focussed on here.

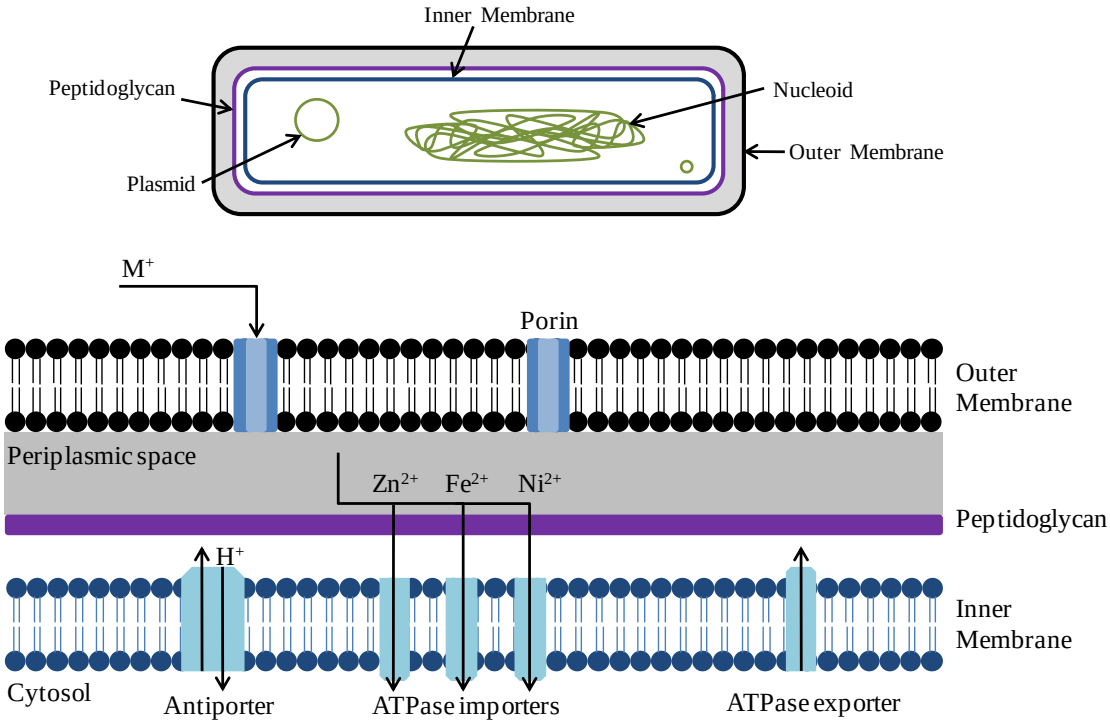


Figure 1.2: The artitecture of the cell wall in gram negative bacteria. Metal ions (M^+) move across the outer membrane through Porins. Movement across the inner membrane is more selective, requiring a range of antiporters or ATP dependant channels to be present.
Adapted from (Ma et al. 2009; Waldron and Robinson 2009).

Due to their high solubility and relatively low molecular weight, the majority of metal ions are thought to simply diffuse across the outer membrane through aqueous channels known as Porins (Waldron and Robinson 2009). Iron is an exception, due to the low solubility of Fe^{3+}, and requires siderophores coupled with active transport to enter the periplasm (Neilands 1995). Evidence has emerged which implies that diffusion is partially controlled by porins (Nies and Silver 2007). This suggests that equilibrium gradients are not solely responsible for metal ion uptake between the

periplasm and the external environment, although another mechanism has not been fully elucidated.

The inner/cytoplasmic membrane is known to exhibit more control over the variety and number of ions which cross it. All metal ions, therefore, require specific metal transporters, such as the ABC-type ATPases, to pass into the cytoplasm. Further details into ABC-type and other uptake mechanisms, such as cation diffusion facilitators, have been reviewed for a range of different metal ions by Ma et al (Ma et al. 2009). Here the focus will be on zinc ion uptake.

Zinc ion uptake

There are two groups of zinc ion importers, high affinity and low affinity. These allow bacteria to uptake zinc ions from environments lacking or rich in zinc, respectively (Hantke 2001).

The high affinity importers, which are almost ubiquitous within bacteria, are the ABC transporters. An example of this is the ZnuABC complex in *E. coli* (Patzer and Hantke 1998*)*. This form of uptake system involves three different protein types; a periplasmic binding protein (ZnuA), a permease (ZnuB) and an ATPase (ZnuC), although the final complex stoichiometry varies between systems (Blencowe and Morby 2003). Import is driven by the hydrolysis of ATP in the cytoplasm and specificity depends upon the ligating residues within the periplasmic protein. For zinc ions these ligating residues are often His, Asp and Glu (Hantke 2005).

The lower affinity uptake systems, such as ZupT, the first ZIP member found in bacteria (Grass et al. 2002), do not require ATP hydrolysis for Zn ion transport. These systems are not as well characterised as the high affinity systems, due to the masking of their uptake contribution by the high affinity uptake systems. More information regarding their mechanism of uptake and metal ion specificity has, however, recently come to light (Taudte and Grass 2010).

Metal ion efflux

To prevent the accumulation of metal ions within the cytoplasm, efflux pumps also reside within the inner membrane. These pumps include a different class of ATPases,

known as P-type ATPases. Although they often act as efflux pumps it has to be recognised that for a few metal ions they are orientated differently and act as influx pumps. Metal ions are also removed from the bacterial cytoplasm by antiporters, belonging to the cation diffuser facilitator (CDF) family (Ma et al. 2009). Metal ions are often pumped from the cytoplasm to the periplasm and then out of the cell. Resistance nodulation cell (RND) efflux systems span the entire cell wall, so metal ions can be pumped directly from the cytoplasm to outside the cell (Nies and Silver 2007).

Zinc ion efflux

The efflux of zinc ions has been relatively well characterised in bacteria, with contributions from CDF members, P-type ATPases and RND type exporters (Hantke 2001). In cyanobacteria however, CDF members have as of yet not been identified and so P-type ATPases are considered to be the main form of exporter (Ma et al. 2009).

These pumps use the energy provided in ATP to transport specific metal ions across the membrane, against their concentration gradient. During this process the pump forms a phosphorylated intermediate, hence the name P-Type. The increased production of P-type ATPase pumps reduces the accumulation of Zn^{2+} in the cytosol and therefore heightens the cells tolerance to increased Zn^{2+} concentrations, as exemplified by ZiaA. The transcription of this zinc specific exporter pump, found in the cyanobacterium *Synechocystis*, is tightly controlled by an ArsR/SmtB family member ZiaR (Thelwell et al. 1998).

A highly characterised example of a P_{1B}-type ATPase is ZntA. This exporter removes divalent zinc, cadmium and lead from *E. coli*. Binding of metal ions occurs on both an exposed cytosolic N-terminal domain and a conserved region within a transmembrane helix (Liu et al. 2005). It is however likely that ATP hydrolysis, and in turn metal ion efflux, is reliant on the coordination geometry formed upon metal ion binding in the transmembrane site. It is this, and not the N-terminal site, which provides the pump with specificity and selectivity (Liu et al. 2005; Ma et al. 2009) . A complete ATPase exporter, CopA, has recently been crystallised. Although CopA is a Cu^{2+} exporter the core structural information obtained is transferable to other P_{IB}-

type ATPases. A platform, composed of two helices, is proposed to be involved with the docking of metallochaperones or the N-terminal heavy metal binding domain (HMBD). In this structure the HMBD was not modelled but the electron density indicates interaction with the cytosolic A-domain suggesting a role in inhibition and/or activation (Gourdon et al. 2011).

The CDF members, such as the zinc ion specific Yiip, are composed of six predicted transmembrane domains and utilise the proton motive force to move metal ions through the inner membrane into the periplasm (Rensing et al. 2007; Ma et al. 2009).

Metal ion storage

Storage proteins, such as ferritins and metallothioneins (MTs), also aid in the prevention of free metal ion accumulation by binding excess free metal ions. These bound metal ions can therefore no longer exhibit a toxic effect upon the cell and may also be transported to necessary sites. The zinc storage MT SmtA will be discussed in further detail in Chapter Five.

Ferritins are proteins expressed in response to a heightened level of Fe. These alpha helical cage-like proteins have the ability to oxidise Fe^{2+} (and subsequently store and concentrate ferric ion within their cavity. The removal of mineralised iron from the centre of ferritin, for use in iron containing co-factors, requires either the use of ferritin pores or degradation of the protein in the lysosome (Liu and Theil 2005).

Metal ion intracellular trafficking

In order to keep the ‘free’ intracellular concentration of some metal ions low, notably copper and zinc, it is predicted from sensor binding affinities that no free cytosolic ion pool exists (Outten et al. 2001; Waldron and Robinson 2009). If metal ion requiring proteins can not receive their required free metal ions from a cytosolic pool, how do they obtain them? Intracellular trafficking is difficult to study due to its dynamic nature. Ubiquitous proteins known as chaperones, however, have been identified and appear to be involved in the movement of both copper and nickel between proteins (Waldron and Robinson 2009). This form of transport system prevents highly active species, such as copper, from forming inappropriate interactions before reaching its target (Harrison et al. 2000). The transfer of the

tightly bound copper from the metallochaperone to its target is highly specific. This specificity is highlighted by the homologous structures CopZ and MNKr2. The chaperone CopZ transfers copper to the repressor CopY, displacing a bound zinc ion, via protein-protein interactions followed by ligand exchange reactions (Cobine et al. 2002). MNKr2 cannot make this transfer and replacement (Cobine et al. 1999) due to a different pattern of surface charged residues (Cobine et al. 2002).

The high capacity of MTs to bind metal ions and the presence of domains containing possible interaction sites makes MTs a very attractive candidate for the role as a zinc chaperone. This proposal, although still speculative, has become more substantial due to the occurrence of MTs in the transport of zinc into mitochondria (Costello et al. 2004).

1.2.6 Metalloregulators

To maintain a beneficial concentration of metal ions; influx, efflux, movement and storage of each metal ion is tightly regulated by sensor proteins, known as metalloregulators. These proteins act as transcriptional regulators of genes or operons which encode for proteins involved in the afore-mentioned homeostatic processes. Regulation at the level of transcription is ideal for biological systems as it prevents the production of proteins when conditions do not warrant them. This saves the cell energy and enables the use of available metabolites for other cellular processes.

Bacterial metalloregulators

An extremely diverse set of metalloregulators has evolved in bacteria, with each protein sensing one or a few metal ions. This means each bacterial cell contains a number of metalloregulators, relating to their metal ion cellular requirements and environment. The currently identified metalloregulators have been split into two different functional groups, Table 1.2 and Table 1.3. They are separated by whether they control genes relating to either metal ion movement into the cell, or the removal and storage of ions from and within the cell respectively. Additionally these regulator proteins can be grouped structurally creating the ten current structural families (Ma et al. 2009), all of which are named after their founding member.

As metalloregulators control the transcription of all the elements involved in maintaining a homeostatic system, it is very important that they bind and respond to the correct metal ions. If this was governed by nothing but absolute metal binding affinities, an increase in high affinity ions, such as Cu^{2+}, would lead to the efflux or storage of all metal ions. Research in this area has revealed that the affinity of metal ions to the sensor is not always the mechanism which determines their specificity, but is also in part down to allostery and access (Waldron and Robinson 2009).

Access dictates metalloregulators selectivity because bacteria inhabit many different environments and have different metal ion requirements. These factors determine to what extent different metal ions are accumulated within the cytosol. If the buffered levels of a metal ion species are significantly lower than the cognate metal ion, then the correct metal will tend to bind to the metalloregulator. In a different bacterial species these buffered levels may vary causing the same metalloregulator to bind other metal ions. The response of the metalloregulator DtxR to exclusively Fe^{2+} in *C. diphtheria*, whereas in *B. subtilis* DtxR responds to both Fe^{2+} and Mn^{2+}, highlights the theory of selectivity by access (Guedon and Helmann 2003). Access to metal ions can also be controlled by intracellular trafficking, ensuring the correct metal ion binding.

The metalloregulator NmtR shows the concept of allostery perfectly. Apo-NmtR has a high affinity for its DNA binding site. Upon the binding of Co or Ni an allosteric change occurs reducing NmtR's DNA affinity and, as a consequence proteins involved with Co or Ni homeostasis are produced. Although Zn^{2+} binds to NmtR at a higher affinity than Co^{2+}, Zn^{2+} binds tetrahedrally, whereas Co^{2+} binds octahedrally. The use of all six ligands is required for the allosteric change, and therefore the sensor is specific to cobalt and not zinc ions. Further structural, and possibly cellular, factors help determine metal ion specificity such as; the complement of ligating residues within a binding site, the location of these binding sites and the possible requirement for metallochaperones (Cavet et al. 2003).

Table 1.2: The structural families of the metalloregulatory proteins which belong to the functional group that causes an increase of metal efflux/sequestration.
The founding member and any additional members associated with metal ion sensing are shown along with the metal ions they sense. Family members sensing oxidative stress have been omitted. Adapted from (Ma et al. 2009).

Family (Founding member)	Additional members	Metals sensed	Mode of action	Structural fold
ArsR/SmtB	CadC, AztR, ZiaR, CmtR, CzrR, NmtR, KmtR, BxmR	As, Sb, Bi, Zn, Cd, Pb, Co, Ni, Cu, Ag	De-repression	αααααββα
CopY	N/A	Cu	De-repression	
CsoR/RcnR	N/A	Cu, Ni, Co	De-repression	
MerR	ZntR, PbrR, CueR, Gols, CoaR	Hg, Zn, Cd, Pb, Cu, Au, Ag, Co	Activation	βααββαααα
NikR	N/A	Ni	Co-repression	βααβαββαβ
TetR	SczA	Zn	De-repression	

Table 1.3: The structural families of the metalloregulatory proteins which belong to the functional group that causes the reduction of metal uptake.
The founding member and any additional members associated with metal ion sensing are shown along with the metal ions they sense. Family members sensing oxidative stress have been omitted. Adapted from (Ma et al. 2009).

Family (Founding member)	Additional members	Metals sensed	Mode of action	Structural fold
DtxR	IdeR, MntR, ScaR, SirR, TroR	Fe, Mn, Cd, Zn	Co-repression	αααββαααα
Fur	Zur, Mur, Nur	Fe, Zn, Mn, Ni	Co-repression / De-repression	αααααββββαβ
LysR	ModE	Mo	Co-repression/ Activation	
MarR (AdcR)	ZitR	Zn	De-repression/ Co-repression	
NikR	N/A	Ni	Co-repression	βααβαββαβ

1.2.7 SmtB/ArsR family of regulators

The SmtB/ArsR family of transcriptional regulators are the most extensively studied structural family out of the ten currently identified. In 2009 the number of genes believed to encode a SmtB/ArsR family member was approximately 550, with the majority of bacteria containing at least one member (Campbell et al. 2007). The family members share a conserved secondary structural fold, αααααββα, and are able to dimerise forming homodimers containing helix-turn-helix (HTH) motifs and metal ion binding sites (Busenlehner et al. 2003). Each family member regulates their operon in a similar fashion; this involves sterically blocking the action of RNA polymerase by binding to specific sites in the operator-promoter region of the DNA. The genes corresponding to efflux and sequestering proteins are then only de-repressed upon binding of metal ions to specific binding sites within the sensor. These metal binding sites vary according to the family member under observation but induce the same effect, a conformational change which reduces the sensors affinity to DNA, allowing RNA polymerase to access the promoter to initiate transcription. The metal ions sensed by this family are diverse, Table 1.4, and include both biologically essential ions, such as zinc, and also inherently toxic ions (Busenlehner et al. 2003) including cadmium, arsenic and lead which often arise from polluted environments.

Table 1.4: Selected members of the SmtB/ArsR regulatory family.
Adapted from (Busenlehner et al. 2003)

Member	Proposed regulatory site	Metals	Regulated proteins
SmtB	α5	Zn	SmtA
ZiaR	α5 and α3N	Zn	ZiaA
CzrA	α5	Co, Zn	CzrB
NmtR	α5C	Ni, Co	NmtA
CadC	α3N	Cd, Pb, Bi, Zn, Co	CadA
ArsR	α3	As, Sb	ArsA, ArsB, ArsC

Sequence alignments and evolutionary analysis of six SmtB/ArsR family members shows the conservation of important residues within the primary structure, particularly in areas which encode certain structural motifs. Further analysis of the sequences reveal that although the binding fold has been conserved between family members, distinct metal binding sites have evolved; explaining how this protein family senses such a broad array of metal ions (Busenlehner et al. 2003).

Alpha 5 sensors

ZiaR and CzrA are proteins homologous to SmtB; both sharing over 45% sequence similarity with SmtB. All three proteins are known to sense zinc ions within their respective bacteria. SmtB and CzrA require their α5 site, composed of imidazole and carboylate ligands, to coordinate zinc ions to cause the conformational change required for de-repression (Arunkumar et al. 2009), whilst ZiaR requires the occupation of both the α5 and α3N sites (Thelwell et al. 1998). Upon de-repression of the *czr* and *zia* operons the diffusion transporter CzrB and the ATPase exporter ZiaA are expressed respectively; promoting zinc efflux from the cytosol. The sequence alignment, Figure 1.3, reveals that the α3N site is not conserved in CzrA probably because it is not necessary for zinc ion sensing. Interestingly SmtB has maintained an α3N site. This site, however, has effectively been deemed redundant for Zn^{2+} sensing from experimental data produced this far but might be involved in Cd^{2+} sensing (VanZile et al. 2002a).

Alpha 3N sensors

The metalloregulatory proteins ArsR and CadC are found to be chromosomally and plasmid encoded in numerous bacteria. They sense non-essential metal ions with a thiolate rich α3/α3N sites, and regulate resistance proteins capable of toxic metal ion removal and/or detoxification (Busenlehner et al. 2003). A number of different metal ions, Table 1.4, can bind to the α3N site in CadC, adopting different coordination geometries. The primary sequence of CadC also includes the conserved α5 binding site, Figure 1.3, however the de-repression and thus the production of CadA, a P-type ATPase efflux pump (Endo and Silver 1995), depends only upon α3N occupancy (Busenlehner et al. 2002). The presence of the more complex *ars* operon protects bacteria from arsenic. ArsR usually regulates the production of two different resistance proteins, ArsB and ArsC, a diffusion transporter and asenate reductase respectively. The *ars* operon can sometimes, however, encode an extra metalloregulator ArsD and a further resistance protein ArsA which acts as an ATPase exporter (Busenlehner et al. 2003).

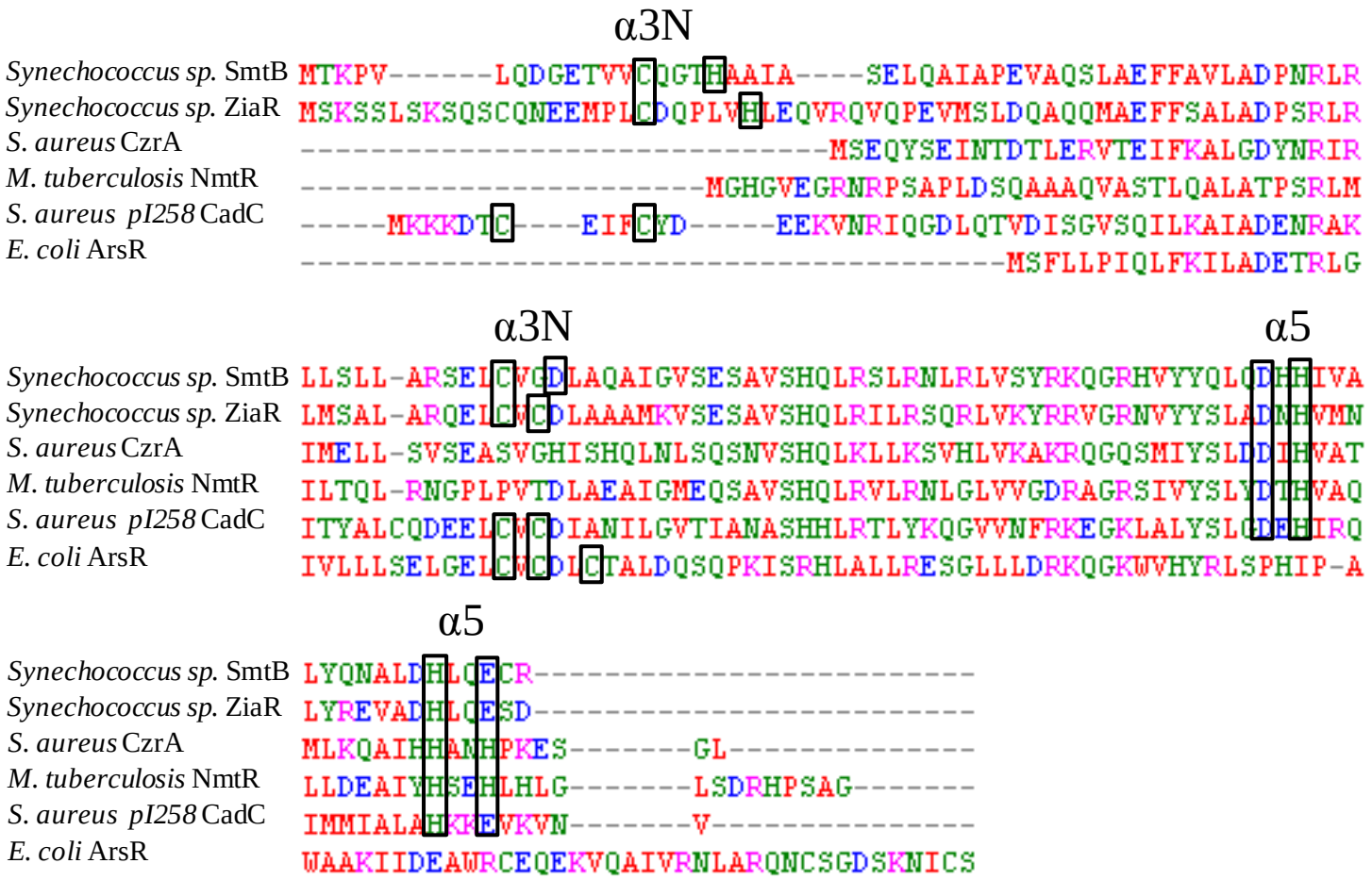


Figure 1.3: Multiple sequence alignment of members of the ArsR/SmtB family.
Created by Cluster Omega (Goujon et al. 2010; Sievers et al. 2011). Sites involved in metal ion binding are boxed in black and grouped into α3N and α5 sites. Residues are colour coded; small hydrophobic amino acids (red), acidic (blue), basic (pink) and other (green).

1.2.8 SmtB

The function of SmtB

The *smt* operon, first isolated and characterised in 1993 (Huckle et al. 1993) , was found to consist of two open reading frames (ORFS), one encoding SmtB and the other encoding the metallothionein SmtA. From SmtB sequence homology and operon orientation it was predicted that the *smt* locus was involved in metal ion sensing (Huckle et al. 1993).

To investigate whether this prediction was correct, genetic mutations were created and cells were exposed to metal ions. Cyanobacterial cells lacking the *smt* operon showed a reduced tolerance to both Zn^{2+} and Cd^{2+} (Turner 1993). No growth inhibition was found in non metal supplemented media and tolerance was restored by re-addition of the operon via transformation, indicating the *smt* operon was only vital upon the cells exposure to metal stress. Rearrangements of the *smt* operon were shown to occur in Cd resistant strains leading to an increase in *smtA* (Gupta et al. 1992). This arrangement was shown to delete part of *smtB* and therefore inhibit SmtB production (Gupta et al. 1993). Maintaining the *smt* operon with an *smtB* deletion in a further cell line also resulted in increased levels of *smtA* in all growth media, both non metal supplemented and metal supplemented. This showed that SmtB was solely required for the transcriptional control of *smtA* (Turner et al. 1995). SmtB is now confirmed as a member of the SmtB/ArsR family of metalloregulators, and regulates the homeostasis of zinc, and possibly cadmium, ion levels in the cyanobacterium *Synechococcus elongatus* PCC 7942, Figure 1.4.

At normal free Zn^{2+} levels apo-SmtB represses the transcription of itself and the *smtA* gene. Upon a rise in the concentration of free Zn^{2+}, caused either by changes in the external environment or cellular requirements, metal ions bind to SmtB causing a de-repression of transcription. This in turn causes the production of the metallothionein SmtA. This protein removes excess zinc ions from the cytosol, curbing their toxic effect. Once the Zn^{2+} concentration reaches an adequate level, newly expressed apo-SmtB rebinds to DNA and once again represses transcription.

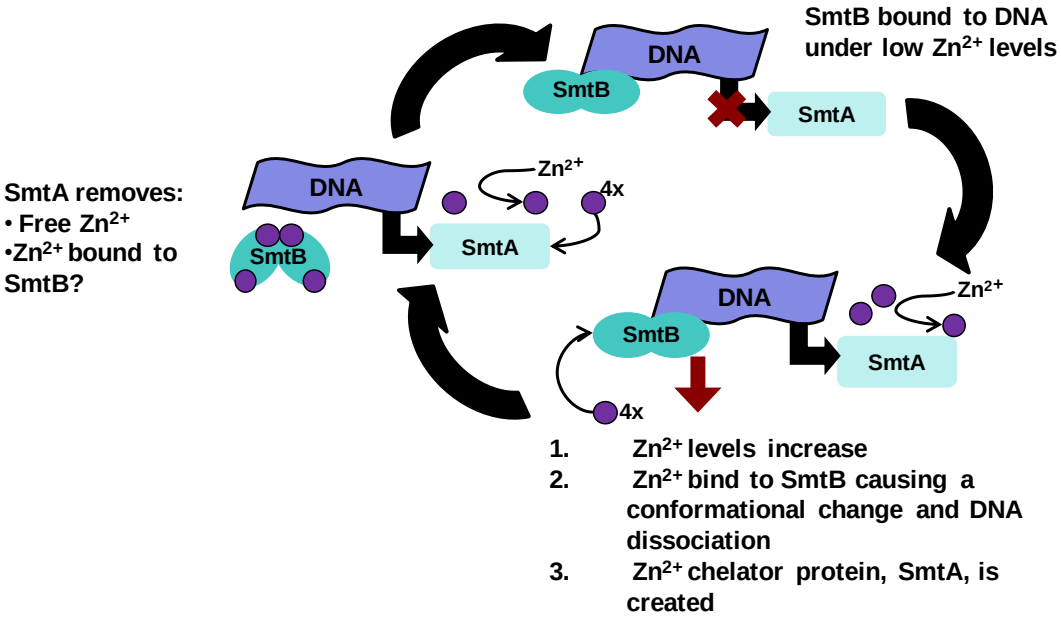


Figure 1.4: The mechanism of regulation of zinc ions in *S. elongatus* PCC 7942

The structure of SmtB

In solution (25 mM Tris, 2 mM DTT, 2 mM EDTA at pH 7.4), apo-SmtB has been shown by sedimentation equilibrium analysis to exist in equilibrium between monomer, dimer and tetramer (Kar et al. 1997). The biologically active form of this repressor, however, is thought to be the predominant dimeric species, which was crystallized providing atomic resolution structural data (Cook et al. 1998; Eicken et al. 2003). These data presented SmtB as a 20Å x 30Å x 60Å elongated dimer, with each monomer containing two β strands and five α helices in the conserved secondary structural fold, αααααββα, Figure 1.5. The helix-turn-helix DNA binding motifs in each monomer were thought to be orientated by the remaining six dimeric helices. These remaining helices were identified as being part of the dimeric interface, an area formed from the hydrophobic pockets on the monomers. In this instance (Cook et al. 1998), and in all subsequent crystal structures (Eicken et al. 2003), the N-terminal arms (approximately 20 amino acids) of both monomers within the dimer are missing, highlighting an element of disorder (Eicken et al. 2003).

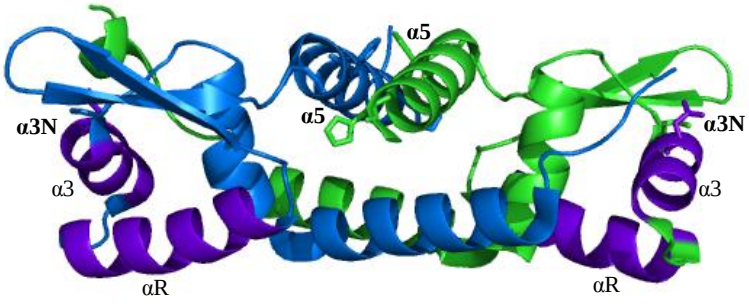


Figure 1.5: The homodimeric structure of SmtB (1R1T)
The two monomers are highlighted in blue and green. Ligands from these monomers make up the potential zinc binding sites (bold type). The DNA recognition HTH regions are purple and the helices which create this motif are labelled (normal type).

Equilibrium analysis, along with X-ray crystallography data, has identified four potential zinc ion binding sites within the dimer (Kar et al. 1997; Cook et al. 1998). Cook *et al*. soaked an apo-SmtB crystal in mercuric acetate and predicted that the four mercury binding sites would have the potential to bind zinc ions. This X-ray data predicted the involvement of C61, D64 and H97 and possibly a water molecule in one site whilst the other was predicted to consist of D104, H106, H117 and E120 (Cook et al. 1998).

These four sites were later shown to be two sets of distinct zinc binding sites and were termed the α3N and α5 sites due to the secondary structural position of the ligating residues (VanZile et al. 2002b), Figure 1.5. Each binding site was predicted to contain four amino acid residues, which were often found to be conserved amongst other members of the SmtB /ArsR family.

The α3N site is proposed to consist of C14 and H18 from the N-terminus of one monomer and C61 and D64 from the third alpha-helix of the adjacent monomer, (VanZile et al. 2002b). The ligating residues within the α3N site were determined by observing changes in amide backbone chemical shifts between apo-SmtB and zinc loaded SmtB. Clear changes were observed for two out of the three residues observed in the Hg^{2+} crystallography data, residues C61 and D64. Chemical shift changes were also observed at C14, however the adjacent residue was also highly perturbed, indicating either residue could be involved (VanZile et al. 2002b). The

final ligating residue, His18, was predicted although there was no supporting NMR data. X-ray data indicates, however that the His18 side chain is close enough to be involved in binding (Eicken et al. 2003), although no complete structure for the α3N site has ever been observed.

The proposed metal sensing sites (α5) are at the dimeric interface and involve the fifth alpha helices of both monomers. One helix provides D104 and H106 and the other H117 and E120 (VanZile et al. 2002b). This order is reversed in the symmetrical site (Cook et al. 1998). Although crystal structures of apo-SmtB, $SmtB_2Zn$ and $SmtB_2Zn_2$ exist, a structure with all four sites occupied has never been obtained. It is currently believed that all four zinc binding sites can not be occupied simultaneously. For further background of SmtB, Zn^{2+} binding stoichiometry and conformational changes see Chapter Three.

The operator-promoter sequence

The *smt* 100 bp operator-promoter sequence is located between the two genes, *smtB* and *smtA*. Within this sequence two sets of promoter -10 and -35 sequences have been identified, one set on each strand, allowing for the transcription of each gene, Figure 1.6. Due to the presence of this divergent and overlapping promoter the operon was predicted to be a divergon (Huckle et al. 1993). This is where two genes within an operon are transcribed from a common operator/promoter sequence in different directions. Often these gene products work within the same system, with one gene product regulating the other by binding to the operator promoter sequence. The -10 sequence, also referred to as a Pribnow box (Pribnow 1975), along with the -35 sequence interact with the detachable RNA polymerase subunit known as the sigma factor. This ensures a tight interaction between the RNA polymerase and DNA, which promotes an energetically favourable unzipping of the DNA, allowing transcription to commence. The RNA polymerase then travels along the sense strand producing an mRNA molecule which is terminated once the RNA polymerase reaches a termination site.

The operator-promoter SmtB binding sites

The first complexes involving the *smt* operator/promoter and a *Synechococcus* PCC 7942 protein, were identified by electrophoretic mobility shift assays (EMSA). One

of these protein:DNA complexes was identified as being dependent on the presence of SmtB and the absence of zinc ions (Morby et al. 1993). Two SmtB binding sites were later proposed after further EMSA, using both cyanobacterial proteins extracts and purified SmtB, revealed three protein:DNA complexes.

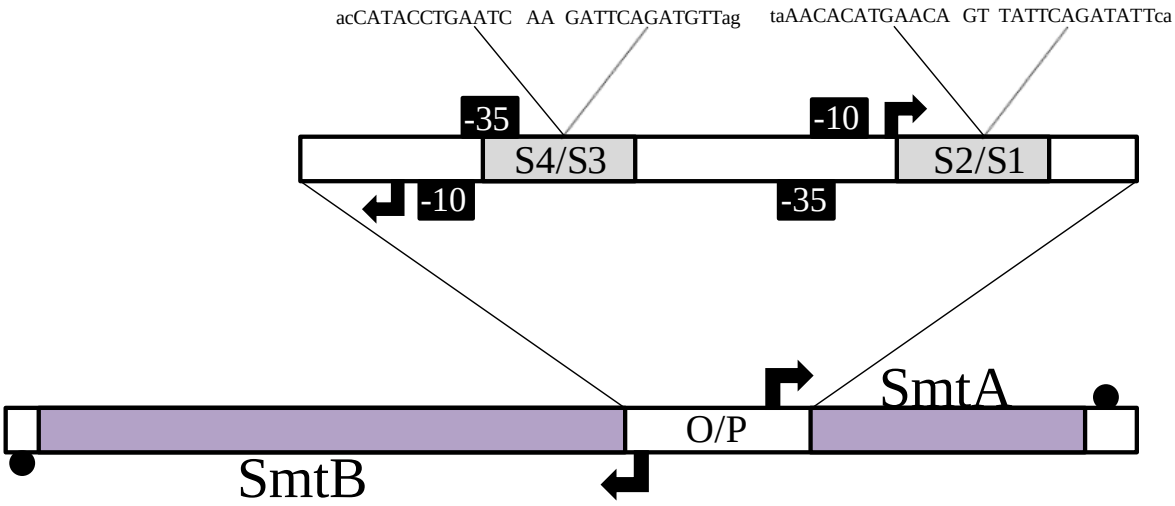


Figure 1.6: A schematic of the *smt* operon
showing the divergently transcribed genes, shaded in purple, with initiation (arrow) and termination sites (dot). The operator promoter region has been blown up to reveal the position of the two proposed SmtB binding sites. The DNA sequences of these two 12-2-12 inverted repeat binding sites are also shown. Adapted from (Huckle et al. 1993; Turner et al. 1996; VanZile et al. 2002a)

The location of these two sites were identified after the implementation of methylation interference assays, a high resolution guanine modification technique (Erbe et al. 1995). S2/S1 was shown to be within an imperfect 6-2-6 inverted repeat located close to the *smtA* initiation site, whilst S3/S4 was located within a 7-2-7 inverted repeat downstream of the S2/S1 site, Figure 1.6, (Erbe et al. 1995). Mutations to the operator-promoter sequence showed SmtB interactions were not restricted to the proposed 6-2-6 inverted repeat but extended to a wider 12-2-12 repeat (Turner et al. 1996), Figure 1.6.

Both sites have the ability to bind at least one SmtB dimer (Kar et al. 2001), although S2/S1 can, with enough flanking sequences, bind two. To visualise binding at the S4/S3 site higher concentrations of SmtB have to be present (Turner et al.

1996), revealing a lower binding affinity (VanZile et al. 2002b). This evidence combined with the observation that deleting the S3/S4 has no implications on SmtA expression suggests this site is more likely to be involved in auto-regulation of the *smtB* gene instead of SmtA regulation (Turner et al. 1996). The proposed models for SmtB:DNA interactions, including SmtB dissociation upon zinc binding is discussed in Chapter Four.

The comparison of 16 cyanobacterial SmtB DNA binding sites reveals if, and to what extent, nucleotides are conserved within the 12-2-12 inverted repeat, Figure 1.7. The main area of conserved residues is within the 6-2-6 inverted repeat, however there are also highly conserved A/T residues flanking this core region. All these conserved nucleotides are highly likely to be involved in non-covalent interactions required to bind SmtB.

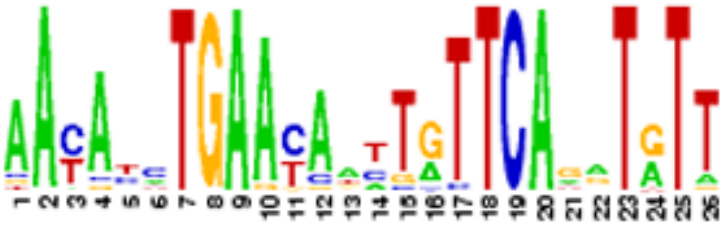


Figure 1.7: A sequence logo of the S2/S1 SmtB binding site
The complete binding site is displayed with large letters showing conserved nucleotides. Taken from RegPrecise (Novichkov et al. 2010)

The comparison of DNA binding sites between SmtB/ArsR α5 metalloregulators provides a core DNA binding motif (aAtAxxTGAaca xx tatTCAxaTxtt) (Busenlehner et al. 2003). This sequence agrees very closely with the sequence logo generated for SmtB, indicating a high level of conservation between all of the binding regions of this form of sensor. A comparison to the α3/α3N core DNA motif (TAxAxTCAAAta xx taTTTGaxTxTA) (Busenlehner et al. 2003) exposes the differences between the α5 and α3N regulators binding interactions.

1.3 Mass spectrometry

1.3.1 Mass spectrometry

Mass spectrometry (MS) is a measurement technique which is capable of measuring the mass to charge ratio (*m/z*) of ions in the gas phase. The main components of a mass spectrometer include an inlet system, source, analyser and detector, Figure 1.8. Varying these components within a system results in instruments which can focus on different applications (Cristoni and Bernardi 2003), resulting in the use of MS in many areas of science, in both academia and industry.

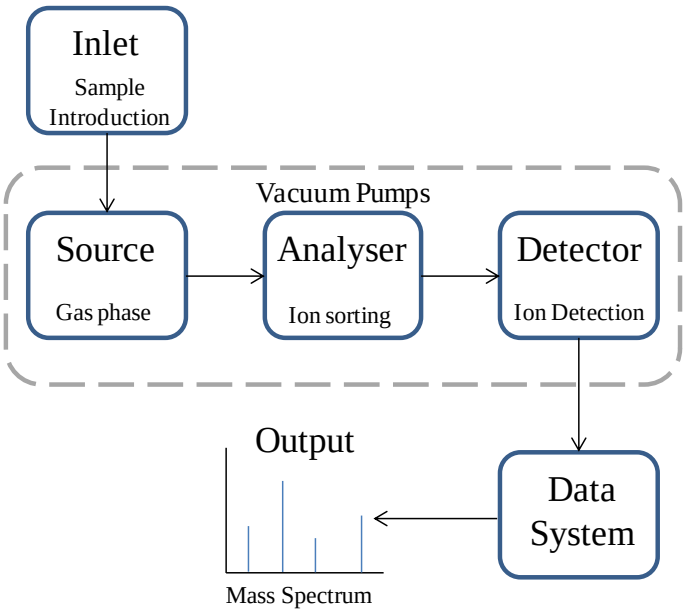


Figure 1.8: A schematic of a mass spectrometer

1.3.2 Ionisation

The ionisation of a sample normally occurs in the source of a mass spectrometer. During this process the sample becomes charged. In many cases, however, the analyte of choice also has to be converted into the gas phase from either a solution or a solid. There are many different processes which result in ionisation, although some are more suited then others to certain sample types. Initially MS was restricted to the study of small volatile thermally stable molecules due to the implementation of ionisation techniques, such as electron impact/electron ionisation (EI), chemical ionisation (CI) and field ionisation (FI). The energies and/or thermal regime used within these ionisation techniques are incompatible with the study of biomolecules; often covalent bonds do not survive the process intact, resulting in fragmentation

(Feng et al. 2008). Advances in ionisation resulted in ionisation processes, such as electrospray ionisation (ESI), matrix assisted laser desorption ionisation (MALDI) and the preceding fast atom bombardment (FAB). These processes are capable of ionising non-volatile thermally unstable molecules, without fragmentation. The invention of these 'soft' ionisation techniques have bought MS to the forefront of biomolecular analysis.

Electrospray

ESI is the most suitable ionisation technique for the study of biological samples; and is also compatible with inlet systems incorporating liquid chromatography (Feng et al. 2008). A solvent dissolved analyte is ejected from a capillary upon the application of a strong electric field. A fine spray of charged analyte containing droplets is produced under atmospheric pressure, which go on to desolvate (Fenn et al. 1989). The evaporation of the solvent gradually reduces the radius of the droplet, until the electrostatic repulsion within the droplet overcomes the surface tension. This is known as the Rayleigh limit. Under these conditions the droplet undergoes fission and splits into smaller droplets. This process of solvent evaporation and fission then repeats itself until either the ion evaporation model (IEM) (Iribarne and Thomson 1976) or the charge residue model (CRM) (Schmelzeisen-Redeker et al. 1989) produces a broad range of multiply-charged ions, Figure 1.9, (Heck and Van den Heuvel 2004). The IEM proposes that after a number of evaporation and fission events, the analyte ion is lifted from the droplet surface into the gas phase, due to the high field strength on the droplet surface. The CRM, however, argues that the evaporation fission cycle proceeds until a droplet contains a single analyte. The evaporation of the final solvent leaves this analyte with the charge previously carried by the droplet (Kebarle and Peschke 2000). It is now widely accepted that different analytes undergo different processes during ESI, with the IEM favoured for small molecules and possibly unfolded proteins (Ahadi and Konermann 2012). Compact proteins are more likely to undergo the CRM (Heck and Van den Heuvel 2004).

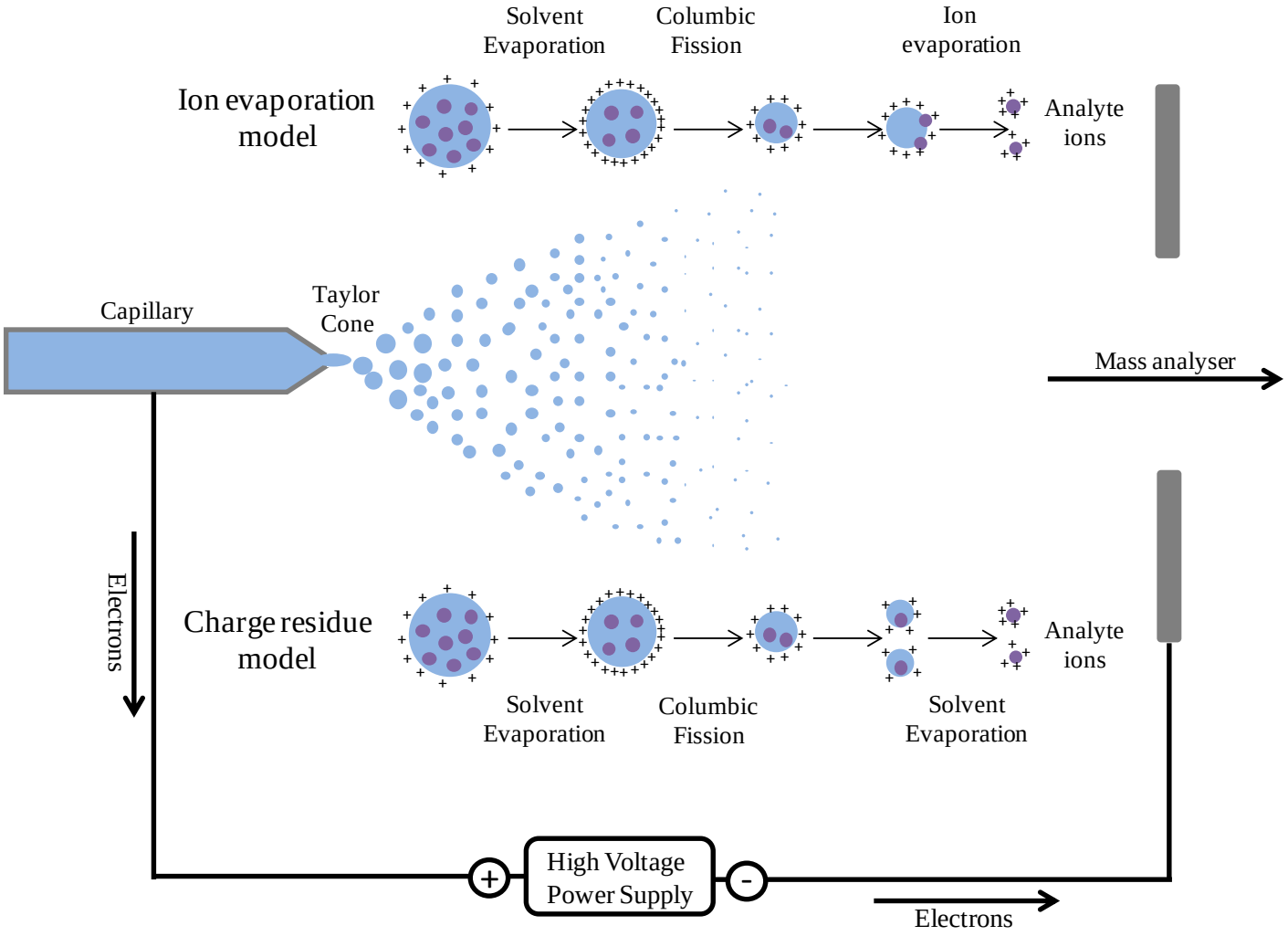


Figure 1.9: Electrospray schematic

Presenting the two models of analyte ion formation from multiply-charged analyte containing droplets.

Matrix assisted laser desorption ionisation

The other ionisation technique used to study biological macromolecules is MALDI. The principles behind this technique were first developed by Karas and Hillenkamp (Karas et al. 1987). They were able to show that laser desorption had the ability to transfer bimolecular ions into the gas phase, without the disruption of covalent bonds (Karas and Hillenkamp 1988). Ions produced by MALDI, however, are thought to retain more energy than in ESI (Feng et al. 2008) and the sample preparation involved is often considered to expose proteins to non-physiological conditions (Heck and Van den Heuvel 2004). Although Karas and Hillenkamp were pioneers in this work, it was Tanaka (Tanaka et al. 1988) who was jointly awarded the noble prize, along with Fenn, for his contributions to *"the development of methods for identification and structure analyses of biological macromolecules"*. Tanaka was able to transfer a number of intact proteins and polymers, into the gas phase, including the proteins lysozyme and chymotrypsinogen (Tanaka et al. 1988).

In brief, the sample is co-crystallised with an excess of an appropriate matrix, with a suitable absorption spectrum, and then subjected to a pulsed laser. This is often an ultraviolet (UV) laser, however other lasers for example infrared (IR) lasers have been shown to be applicable to the study of proteins (Zhang et al. 1998). The energy from the laser pulse is absorbed by the matrix. This causes the matrix to vaporise and the anaylate to be transferred into the gas phase and ionised. The actual mechanism of analyte ionisation is now thought to follow a two step mechanism. This involves the formation of primary ions, which then under go reactions forming, the detected, secondary ions. The mechanism, by which the primary ions are formed, is still widely debated (Knochenmuss 2006). Ions produced from MALDI are largely singly charged and therefore require a mass analyser capable of measuring over a broad *m/z* range, such as a time of flight (TOF) analyser.

1.3.3 Analysers

After such significant improvements to ionisation in the 1980's, allowing for effectively a whole new area of science to be studied by MS, analysers had to be improved and redesigned to be applicable for the new MS era.

A variety of different mass analysers currently exist, including, quadrupoles (Q), ion traps (IT), time of flight (TOF), Orbitrap and Fourier transform ion cyclotron resonance (FTICR). Although all analysers act to measure the ions produced from ionisation, according to their mass to charge, each type can be more applicable to different ionisation methods and areas of research. In many cases instruments now contain more than one mass analyser (hybrid instruments), so that they can be used in conjunction with each other.

Quadrupoles

The initial work carried out on the development of the quadrupole mass analyser was by Wolfgang Paul in the 1950's. In a simplified summary, a quadrupole mass analyser contains four parallel rods, attached to a direct current (DC) so that opposing rods have the same charge. Radio frequencies (RF) are then also applied to the rods allowing only a certain mass to charge to have a stable trajectory through the rods, whilst all the other mass to charge ratios have unstable trajectories and are lost, Figure 1.10. A quadrupole therefore acts as an ion filter, where the RF and DC voltages can be varied to scan through the mass range or set to select for a certain ion.

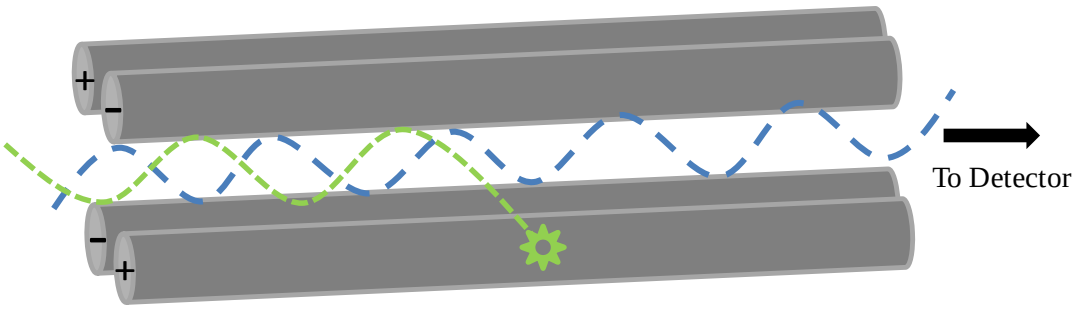


Figure 1.10: A schematic of a quadrupole analyser
The blue ion has a stable trajectory and therefore passes through the quadrupole whereas the green ion has an unstable trajectory and is lost.

Ion Traps

Quadrupole ion trap mass analysers, invented alongside the quadrupole mass filter by Wolfgang Paul (Paul 1990), consist of a ring electrode and two end cap

electrodes. Ions enter through the aperture of one of the end caps and are trapped in the space between the electrodes by the application of oscillating electrical fields. To detect an ion, the ion must first be ejected from the trap, through the other end electrode. To do this the amplitude of the RF voltage, placed upon the ring electrode, is increased (Stafford Jr et al. 1984). Ions with a lower *m/z* are ejected first, whilst ions with the highest *m/z* are ejected last (Jonscher and Yates 1997). The trapping nature of this analyser enhances the sensitivity of the instrument, however trapping too many ions within the cell leads to space charge effects, which effect performance.

There is another form of ion trap, know as the linear ion trap. This form of trap closely resembles the quadrupole mass analyser, Figure 1.10. To trap the ions, RF voltages are applied to the rods confining them radially and stopping potentials are placed upon the end electrodes. This form of trap can not only be used to trap ions but can also act as an ion guide, an *m/z* selector or a collision cell. It is therefore often combined with other mass analysers in hybrid instruments, although it can also be found in isolation (Douglas et al. 2005).

Time of Flight

The first commercial time of flight (TOF) mass analyser, designed in the 1950's, was of a linear orientation. Although providing a potentially endless mass range and having high sensitivity, high resolution could not be achieved with only a few hundred being observed (Mamyrin 2001). As ions enter the TOF analyser they are subjected to an electric field which provides the ions with kinetic energy. The ion then enters a field free flight tube and the time it then takes for the ion to reach the detector (*t*) is measured. This time is dependant on the *m/z* of the ion, Equation 1.1, with a low *m/z* will travel through the TOF faster than ions with high *m/z* and so separation occurs.

$$t = \sqrt{\frac{m}{z} \times c}$$

Equation: 1.1

Initially MALDI, and not ESI, was combined with TOF analysers, this was due to the massive sensitivity loss associated with combining a continuous ionisation method and a pulsed analysis (Mirgorodskaya et al. 1994). Many different research groups developed TOF analysers, with the inclusion of orthogonal acceleration, independently. In this process a region before the TOF collects ions before accelerating them orthogonally into the flight tube. This development allowed many different ionisation sources to become compatible with this form of analyser. It was however the work of Dodonov *et al* in the mid 1980's who developed an ESI-TOF mass spectrometer with sufficient sensitivity and went on to show its application to biomolecules, such as peptides and proteins (Mirgorodskaya et al. 1994).

The introduction of delayed extraction (DE) (Wiley and McLaren 1955) and the reflection greatly improved the resolution and mass accuracy of TOF analysers and thus broadened their applicability.

DE, often implemented in MALDI-TOF instrumentation (Vestal et al. 1995) (Takach et al. 1997) also helps correct for kinetic differences. In this process a high voltage pulse is applied to the newly formed ions after a defined time has elapsed. This time delay allows ions with initial higher velocities to move further away from the source region. When the extraction pulse is applied the ions further away from the source region experience a lower potential than those close by.

The reflectron, developed by Mamyrin, is a decelerating field, which ensures that if two ions have the same *m/z*, but enter the TOF with different amounts of kinetic energy they will still impact on the detector simultaneously (Mamyrin 2001).

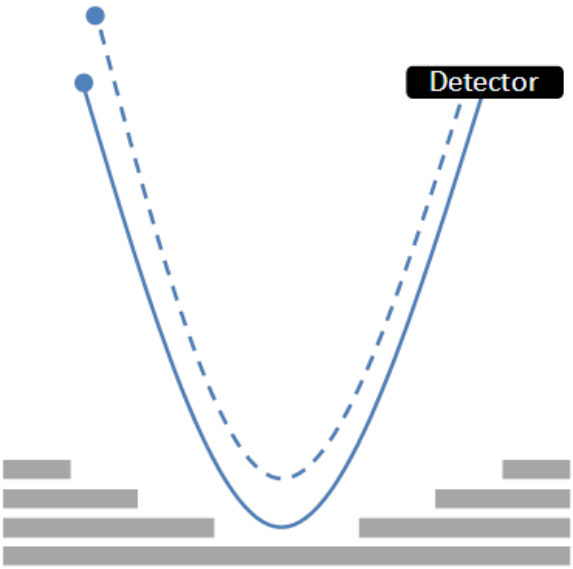


Figure 1.11: A representation of a V-shaped reflectron in a time of flight analyser Both ions have the same *m/z*, however one (solid line) has more kinetic energy than the other (dashed line). Adapted from (Cotter 1999).

Ions with higher kinetic energy travel further into the decelerating field than slower travelling ions, of the same *m/z*. The extra time spent in the decelerating field corrects for any kinetic disparity between ions. The increased flight tube length, created by the reflection, also assists in providing high resolution but to a lesser extent. Reflections can either exist in V-shaped orientation, Figure 1.11 as described above, or a W-shaped orientation. The latter orientation includes an extra ion mirror, leading to higher resolution; however, additional ion loss reduces sensitivity.

1.3.4 Detectors

Once ions have been sorted by the mass analyser they need to be registered by the detector. Here a suitable signal is produced; an electrical current which can be converted into a mass spectrum by a data system. The detection of ions can be based upon one of three different physical properties the ions possess: mass, charge or velocity. It is important that not only the *m/z* of the ion can be detected but also the abundance of the ion.

Early detection involved using double focusing mass spectrometers. These instruments used magnetic and electric fields in a Mattauch-Herzog geometry to

ensure all ions of the same *m/z* reached the same point on a photographic plate (Squires 1998). The spots created could then be interpreted after calibration, with only the shade of the spot revealing the abundance of certain ions. Detection of ions has moved on greatly since these early days. Modern mass spectrometers either contain detectors which record the charge that reaches the detector or they amplify the ion signal. The latter often involves the use of a secondary electron multiplier, examples of which include channeltron detectors and micro-channel plates (MCP)

The MCP detector is a form of continuous dynode electron multiplier. It is made up of many separate channels lined with a semi-conductive material. These all act as individual electron multipliers, allowing for simultaneous ion detection. An ion enters the channel and begins to travel along it. Eventually this ion impacts on the channels surface, known as the conversion dynode. This impact forms secondary electrons which are driven through the channel by an electric field. As these secondary electrons move along the channel they impact again, upon the surface, forming more electrons which eventually results in the production of many electrons by a cascade effect (Koppenaal et al. 2005) (De Hoffmann and Stroobant 2007). The resulting flow of electrons (electrical current) from each channel can then either be combined or kept separate.

This type of detector relies heavily on the velocity of the ion to convert it into a secondary particle. Due to this large ions, which have a lower velocity produce a much lower signal. After the invention of soft ionisation techniques, followed by the drive to get larger and larger molecules into the gas phase, this loss in signal became a significant problem, as an increase in mass leads to an exponential decrease in signal. To overcome this and to increase the overall detection efficiency all ions passing into the detector now are accelerated.

Before the output from the mass spectrometer can be processed by a computer it has to be converted into a digital signal, by the data acquisition system. These are usually either an analogue-to-digital converter (ADC) or a time-to-digital converter (TDC). The former takes the total analogue signal produced by ions hitting the detector, at frequent time points, reduces the high frequency background noise and stores the intensity as a digital value. TDC converters are often used on TOF instruments due

to their pulsed nature (McIntyre 2004). They work by recording an arrival time for each signal produced by the detector, as long as the intensity of the signal is over a pre-determined threshold (De Hoffmann and Stroobant 2007). The consequence of measuring the arrival time for each signal, per TOF push, is a limited dynamic range. Effectively a single arrival time is recorded whether multiple ions, of the same mass, or a single ion created the detector signal (McIntyre 2004).

1.3.5 Tandem Mass spectrometry

Tandem mass spectrometry (MS/MS) experiments can be carried out in space or time, using successive MS analysers or ion trapping approaches respectively.

In MS/MS experiments conducted in space two mass analysers are required and so experiments are carried out on hybrid instruments, often a triple quadrupole (QQQ) time of flight time of flight (TOF-TOF) or a quadrupole time of flight (QTOF). The successive mass analysers are separated by a collision cell, full of a neutral gas such as argon. Ions passing through undergo collisional-induced dissociation or collisionally-activated dissociation (CID/CAD). This form of ion activation, pioneered by Mclafferty and Jennings in the 1960's (McLafferty and Bryce 1967; Jennings 1968), fragments polypeptides through predominantly amide cleavage, generating characteristic b and y ions (Roepstorff and Fohlman 1984), Figure 1.12.

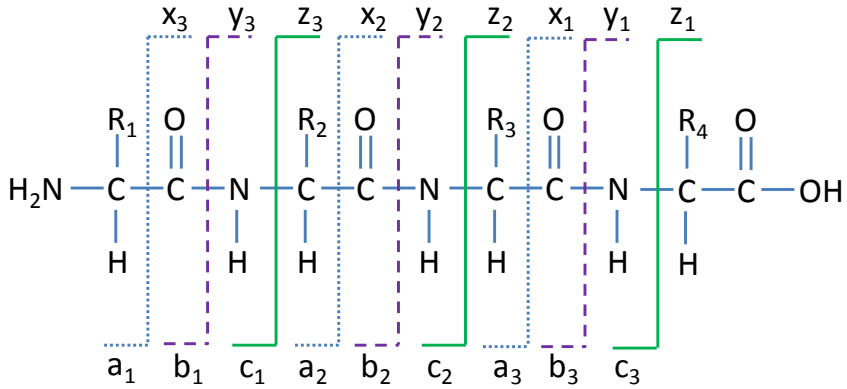


Figure 1.12: Peptide fragment ion nomenclature by Roepstorff and Fohlman

The use of a QQQ allows for a number of different modes of tandem MS to be conducted. These modes include a precursor ion scan, a product ion scan, neutral loss scanning and selected reaction monitoring (De Hoffmann and Stroobant 2007).

The most commonly used mode in structural biology, is the product ion scan. In this mode the first quadrupole selects a specific *m/z* and allows it into the collision cell for CID to occur. The second mass analyser then scans the ions produced, and detection of the different fragment ions occurs.

Tandem mass spectrometry experiments conducted in time are usually carried out in mass analysers capable of trapping ions, such as FTICR and IT (Dodds 2012). The two main forms of this type of MS/MS, in addition to CID, are the non-ergotic mechanisms of electron capture dissociation (ECD) (Zubarev et al. 1998) and electron transfer dissociation (ETD) (Coon et al. 2004). In ECD, often but not exclusively used in conjunction with ESI-FTICR, the multiply charged protein/ peptide captures low energy electrons. This induces cleavage and the formation of c and z ions Figure 1.12, (Zubarev et al. 1998), whist maintaining any post-translational modifications.

The same cleavage pattern is also achievable through ETD. In this experiment, singly charged anions, often flouranthene radical anions, transfer electrons to the protein/ peptide (Coon et al. 2004), causing rapid, random cleavage to the peptide backbone. A negative chemical ionisation source is used to produce the flouranthene radical anions in a two step reaction. Firstly low energy electrons are produced by reacting high energy electrons with methane (M). These low energy electrons then attach to flouranthene (A).

$$M + e^- \rightarrow M^{+\cdot} + 2e^-$$

$$A + e^- \rightarrow A^{-\cdot}$$

After the *m/z* of the protein/peptide of interest is trapped within an ion trap, flouranthene is then accumulated within the cell to allow ETD to occur.

In addition to this protons can also be removed from the ETD products, in a proton transfer reaction (PTR). The flouranthene anion can be used in this reaction because of its high proton affinity. In this reaction protons are removed from the product ions, reducing their charge and thus simplifying the spectra and allowing product ions to be more easily resolved (Hartmer et al. 2008).

1.3.6 Ion mobility Mass Spectrometry

Ion mobility spectrometry (IMS) can and often does exist as a standalone analytical technique and is used the world over, for the detection of explosives and drugs (Guharay et al. 2008). Within this instrumentation, a drift cell is normally incorporated to separate out a mixture of ions, by their mass, charge and shape (Creaser et al. 2004). The technique is therefore referred to as drift cell ion mobility spectrometry (DCIMS). In this form of IMS, ions pass through a gas filled drift tube, under a uniform low electrical field. As an ion packet diffuses through the tube, larger ions collide more frequently with the buffer gas compared to small ions. These collisions therefore increase the time it takes for the large ions to traverse the collision cell, effectively separating them out from the smaller ones (Clemmer and Jarrold 1997). This form of ion mobility can be carried out at ambient or reduced pressure. At an ambient pressure higher separation often occurs due to a higher rate of collisions between the buffer gas and the ions of interest. Much like many gas phase techniques IMS was revolutionised by the application of soft ionisation techniques (Kanu 2008) with DCIMS, run at reduced pressure, being coupled to both MALDI and ESI.

The coupling of IMS to MS, forming ion mobility mass spectrometry (IMMS), added an extra dimension to the separation of analytes. In these instruments the IMS separation of size to charge precedes the mass to charge separation, effectively separating ions that possess the same *m/z* but are a different shape. This opened the door for the use of IMMS for the separation of complex mixtures and the study of molecular structure (Kanu 2008). Its main application in determining structure comes from IMMS's ability to separate out ions which have the same molecular weight but vary in conformation. The extra information obtained from IMMS experiments has resulted in its application to the study of biomolecules, including lipids (Jackson et al. 2007), polysaccharides (Clowers et al. 2005), nucleic acids (Koomen et al. 2002) and proteins, such as Hepatitis B virus capsids (Uetrecht 2008), P22 tail gp4 (Lorenzen et al. 2008), and TRAP (Ruotolo et al. 2005).

IMS and MS were found to be highly compatible as a combined instrument. The time required for IMS separation is milliseconds whereas the MS separation of a TOF is significantly faster, in microseconds. This allows for many mass spectra to be

produced per IMS spectra, resulting in no loss of information (Matz et al. 2002) (Eckers et al. 2007). This compatibility resulted in the production of many in house machines, created by a number of research groups, followed by the release of a commercial instrument (Synapt HDMS system-Waters Corp., Milford, USA.). This was followed recently by the Synapt G2 HDMS system in 2009.

There are many different types of IMS in the literature, for a review on the methodologies behind some of the less well known types of IMS refer to Kanu et al. There are three main types of IMS that frequent the literature most often, these are Drift cell IMS (DCIMS), Travelling wave IMS (TWIMS) and field asymmetric waveform IMS (FAIMS).

DCIMS, as mentioned previously, is the traditional form of IMS and measures the time it takes for an ion to traverse a drift tube under a low electrical field. This form of IMS it is still the only method in which the collisional cross-section of an ion can be obtained directly from the ions measured drift time.

In FAIMS, often carried out at ambient temperature and pressure, an asymmetrical waveform is applied to electrodes comprising the separating region. Different compounds have a different mobility at high field strength in comparison to low field strength, causing them to migrate toward the electrodes at different rates (Purves et al. 1998). The application of a DC voltage, known as a compensation voltage (CV), to an electrode will stabilise the trajectory of a certain ion, preventing it from impacting on the electrodes. Different ions have different CV's; due to their differing mobility and so scanning this voltage allows a mixture of ions to be separated. Unfortunately, as of yet, the information obtained from these types of experiments can not be related to a physical property of the ion, limiting its application (Mukhopadhyay 2008).

1.3.7 Travelling Wave Ion Mobility Mass Spectrometry

The basic process of ion separation by TWIMS is comparable to drift cell devices; an ion which poses a larger collisional cross-section collides with the buffer gas more frequently than a smaller ion, slowing its transit in comparison. In TWIMS however,

the ions are not exposed to a continuous, uniform low electrical field. Instead ions are driven through the cell by voltage pulses applied to a stacked ring ion guide.

In TWIMS instrumentation a stacked ring ion guide (SRIG) forms the IMS cell. The SRIG is composed of a number of circular electrodes stacked so that the ions can pass through the central cavity, Figure 1.13. To improve sensitivity alternating electrodes have an opposing RF current applied to them which reduces the radial diffusion of ions. On top of this RF current, a DC voltage is applied to a pair of electrodes at the start of the SRIG, this current is then transferred after a defined time to the adjacent pair of electrodes. The process of DC current transfer continues down the SRIG forming a travelling wave. These waves, of high potential, drive ions away from them and therefore transfer the ions through the cell. As previously mentioned ions with a larger collisional cross-section encounter more drag and so roll over the waves more often than smaller ions, thus providing separation, Figure 1.13. (Giles 2004; Thalassinos et al. 2004; Pringle et al. 2007).

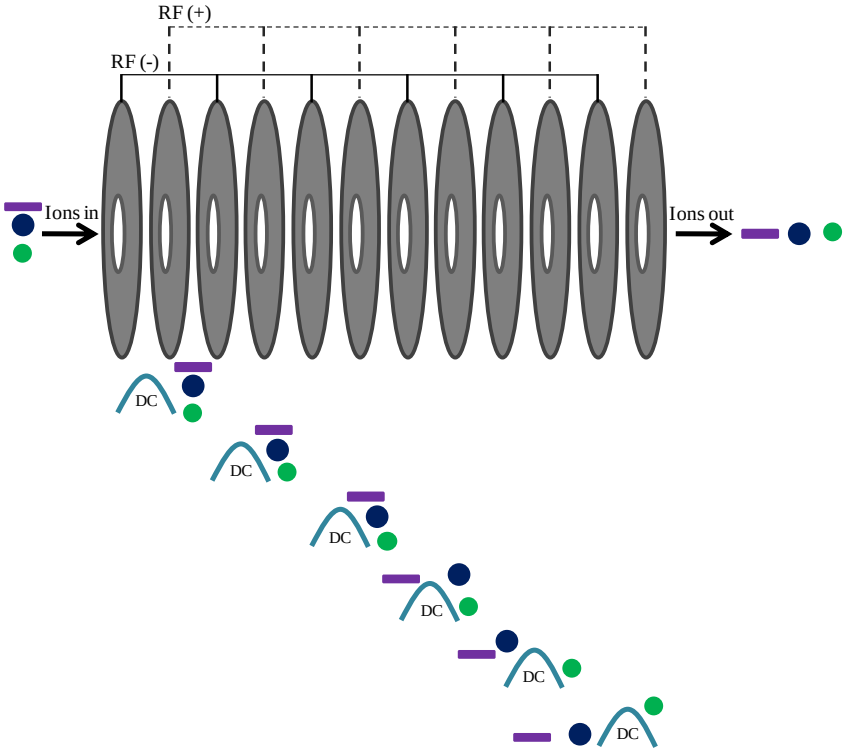


Figure 1.13: The stacked ring ion guide
production of a travelling wave and consequent separation of ions with the same *m/z* but different shape. Adapted from (Giles 2004).

The sensitivity and the resolving power achieved by the Synapt were comparable to the traditional drift cell devices, however the ion mobility resolution was relatively low, around 10 (Ω/ΔΩ). This meant that in some cases if the conformation of the ions were quite similar one broad arrival time would be recorded instead of two. The arrival of a second generation Synapt, Figure 1.14, has gone some way to improving this resolution problem.

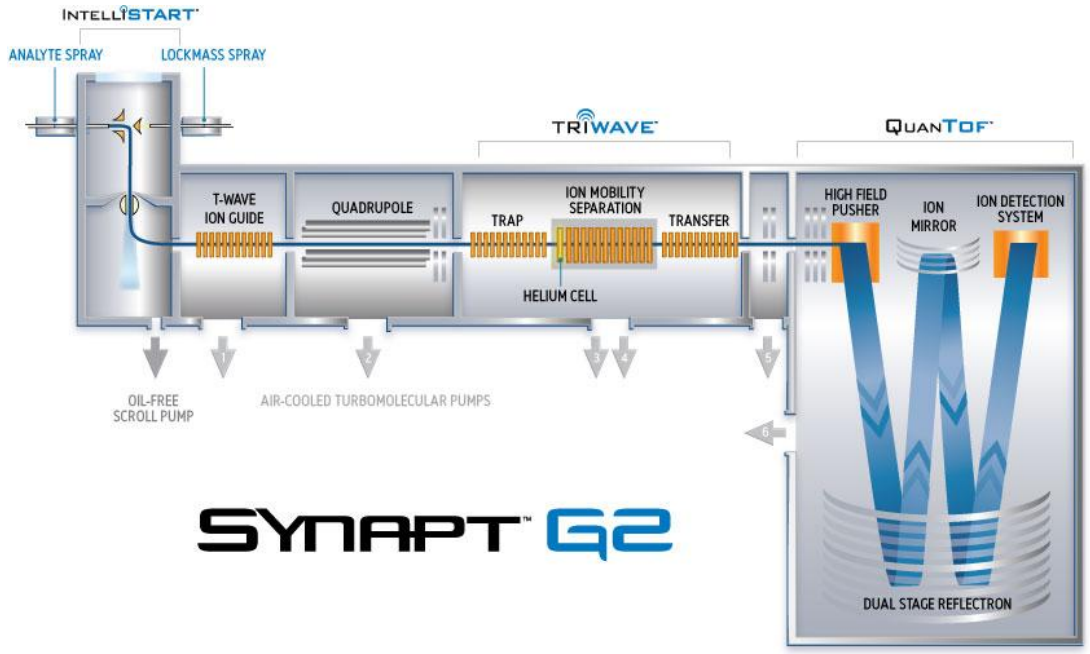


Figure 1.14: A schematic of the Synapt G2
The path of the ions can be followed from the source, through the quadrupole and into the TRIWAVE regions, which contains the IMS separation cell. From here the ions then enter the TOF before being detected.

There were a number of modifications made to the T-Wave cell within the Synapt G2 HDMS system (Waters). These are summarised by Giles *et al*. Many of the changes were centred around trying to increase the nitrogen gas pressure within the IMS cell, which would allow for the use of higher electric fields, resulting in better IMS resolution (Giles et al. 2011). The use of higher gas pressures would however require more energy to drive the ions into the IMS cell, against the outflow of gas. The injection of ions into the IMS cell is a critical step in maintaining native like structures. If the voltage used to accelerate the ions from the trap into the IMS cell is too low a reduction in transmission occurs. If the applied voltage is too high then the resulting ion heating can cause a different population of conformers to be observed

or may even result in ion fragmentation (Morsa et al. 2011). The introduction of a short helium cell preceding the mobility cell reduced ion fragmentation in two separate ways. Firstly, as ions have a heightened mobility in helium, less energy is required to introduce the ions into the cell and therefore a lower injection voltage could be used. Secondly the high helium pressure was shown by Merenbloom *et al*. to collisionly cool the ions. Their work revealed that a reduction in the helium gas flow in the helium cell from 180 mL/min to 75 mL/min resulted in an increase of leucine enkephalin dimer dissociation by 72.5 % (Merenbloom et al. 2012).

The IMS cell was also extended, from 122 to 168 ring electrodes and the DC current, required to make the travelling waves, was applied to four adjacent electrodes instead of two. All of these improvements lead to an increased ion mobility resolution of 40 ($\Omega/\Delta\Omega$) (Giles et al. 2011).

1.3.8 Ion Mobility Calibration

The path of ions through the T-wave is complicated and not completely understood, and therefore TWIMS instruments do not offer a directly proportional relationship between drift time and cross-section, like the drift cell instrument (Mukhopadhyay 2008). Estimated collisional cross-sections of ions can, however, be produced using internal standards with a known collisional cross-section (Thalassinos et al. 2008). Until recently these internal calibrants were often species of low molecular weight; in the case of proteins, denatured myoglobin or cytochrome c were often used (Scarff et al. 2008; Leary et al. 2009). These produced accurate estimated collisional cross-sections for many proteins (Scarff et al. 2008), when both sample and calibrant were run under the same conditions (Leary et al. 2009). In the study of large protein complexes, however, the mobility of the complex would fall outside of the high mobility values of these calibrants. In order to determine a cross-section the calibration would have to be extrapolated, resulting in significant errors (Shvartsburg and Smith 2008) and collisional cross sections which can be dependent on instrument settings, in particular wave height and velocity (Bush et al. 2010) (Salbo et al. 2012). To produce accurate collisional cross-sections for large proteins and their complexes the absolute collisional cross-section of a new set of standards were derived and implemented. These were a range of natively folded proteins ranging from cytochrome c (12 kDa) to GroEL (801 kDa) (Bush et al. 2010). Further

research into T-wave calibration of protein complexes has highlighted the importance of optimising the native calibrants used in IMS calibrations. This work revealed that choosing native-like calibrants which only loosely bracket the mass of the analyte ion, in this instance insulin monomer, dimer and hexamer, still leads to T-wave velocity dependent collisional cross sections (Salbo et al. 2012). Once calibrants were used, which closely bracketed the masses of the different insulin oligomers cross sections became independent of settings and absolute errors reduced to less than 2 % .

The is no universal calibration procedure for TWIMS, however the method outlined below was created by the Scrivens group (Scarff et al. 2008; Thalassinos et al. 2008; Williams and Scrivens 2008), based on the work of Wildgoose *et al*. Originally created for the Synapt HDMS, it was adapted for the second generation instrument, the Synapt G2 HDMS and can be used in conjunction with both the old and the new calibration standards.

The arrival time, measured in scan number (n) relates to the push of the TOF. This is converted to arrival time (t_d) by multiplying it by the pusher time, Equation 1.2.

$$t_d = n \text{ x pusher time (ms)}$$

Equation: 1.2

In previous calibration procedures this value would then be corrected for *m/z* independent flight (time in transfer cell) and *m/z* dependent flight (time between transfer cell exit and TOF pusher). The *m/z* independent flight was shown not to significantly affect the calculated collisional cross-section and so has been emitted for simplicity. To correct for *m/z* dependent time, the instrument dependent enhanced duty cycle (EDC) delay coefficient (*c*) is used, Equation 1.3.

$$t'_d = t_d - \frac{c\sqrt{(m/z)}}{1000}$$

Equation: 1.3

The published collisional cross-sections (Ω) of the chosen calibrant is corrected for reduced mass and charge (*z*), using the molecular weight of the calibrant (m_i) and the molecular weight of the buffer gas (m_n) Equation 1.4.

$$\Omega' = \frac{\Omega}{z\sqrt{(\frac{1}{m_i}+\frac{1}{m_n})}}$$

Equation: 1.4

From the values calculated, a plot of Ω' against t'_d is produced and is fitted to either a power ($y = AX^B$) or linear ($y = Ax + B$) series fit, with an R^2 value ≥0.99. This plot provides the values A and N required for the production of the estimated cross-section (CCS), Equation 1.5 or Equation 1.6 for power or linear fit respectively.

$$CCS = A\ t'_d{}^N\ z\sqrt{(\frac{1}{m_i}+\frac{1}{m_n})}$$

Equation: 1.5

$$CCS = [A\ t'_d + B]\ z\sqrt{(\frac{1}{m_i}+\frac{1}{m_n})}$$

Equation: 1.6

1.4 The study of biological molecules

1.4.1 Biophysical techniques used in structural biology

Many biophysical techniques exist which can be applied to the study of native proteins. These include fluorescence spectroscopy, circular dichroism (CD) spectroscopy, nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography. Each technique often provides complementary data which, when used in combination, can help determine a proteins structure and function.

Fluorescence spectroscopy

Fluorescence spectroscopy measures the emission of fluorescence from intrinsic fluorophores present within proteins, each of which has certain excitation and emission wavelengths. Tryptophan, tyrosine, phenylalanine and disulphide bonds are all known fluorophores present within proteins. Often the excitation and emission wavelengths used within experiments are optimised to limit the study to one of these, instead of a complex mixture of all of them. Tryptophan is often the chosen fluorophore as the imidazole ring present in its side chain has the strongest emission out of all the natural fluorophores. The emission from tryptophan is extremely sensitive to its environment. A red shift in λmax is observed upon the exposure of a previously buried tryptophan to a hydrophilic environment, thus studies on protein conformation and folding can be carried out (Smith et al. 1991).

CD spectroscopy

The difference between a chiral molecule's absorption, of left and right circularly polarised light, measured by a CD spectrophotometer can provide structural information. The peptide bond present within all proteins makes them chiral. This allows for the study of a protein's secondary structure by far-UV (190 – 250 nm) CD, and tertiary structure by near-UV (250 - 350 nm) CD. The electronic transitions of the peptide bonds within secondary structures are different, providing distinctive maxima and minima within the spectrum. From the final spectrum, the average secondary structure of the solution to be obtained, allowing conformational changes to be observed and followed.

In structural biology two of the most widely used and highly regarded techniques are X-ray crystallography and NMR. These are often considered to be highly valuable because the information collected from samples can be at atomic resolution. The results can therefore be used to form a three dimensional structure of the protein of interest.

Nuclear magnetic resonance

NMR experiments produce signals dependent on the environment of NMR active atomic nuclei, which are then assigned to amino acid residues. The most commonly studied nuclei are ^{1}H. These nuclei yield the most information but due to the number present in proteins, especially ones ≥ 10 kDa, resonances overlap even when multi-dimensional techniques are used. To improve this size constraint, ^{13}C and ^{15}N are often looked at along with ^{1}H, (Cooke 1997). As these isotopes are not naturally high in abundance a labelling technique is employed. The inclusion of three dimensions enabled the study of ~ 30 kDa proteins containing approximately 240 residues (Clore and Gronenborn 1998), however most exciting of all is the implementation of 4D experiments with isotope substitution, ^{1}H to ^{2}H (Chary and Govil 2008). In these experiments deuterium labelling is used on a proportion of the proteins side chains, (Luh et al. 1997), improving both sensitivity and resolution. This has allowed the study of proteins in excess of 100 kDa (Chary and Govil 2008).

X-ray crystallography

Often due to the size constraints of NMR, X-ray crystallography is the technique of choice for the study of large proteins (≥35 kDa) (Pusey et al. 2005). As of the time of writing this approximately 88 % of all protein structures in the Protein Data Bank have arisen from X-ray crystallography. To create a structure, the X-ray diffraction pattern produced from an X-ray beam striking a crystal is extensively studied, eventually resulting in the production of an electron density map revealing the atomic coordinates of atoms within molecules. Although many advances have taken place to increase the throughput of X-ray crystallography the technique is still dependent on the production of protein crystals. One of the numerous problems associated with creating protein crystals is homogeneity, even the presence of flexible proteins in a number of conformations can prevent crystallisation (Pusey et al. 2005).

The techniques outlined above can all provide an invaluable insight into protein structure, but they all have limitations. Some proteins do not crystallize which excludes the use of X-ray crystallography. NMR requires a high concentration of labelled recombinant sample, can only cover a relatively small mass range and along with the other spectroscopic techniques can only visualise the average protein solution structure. Mass spectrometry is fast, requires small sample volumes and is not affected by sample heterogeneity (Loo 1997; van Duijn and Heck 2006). These qualities make it a useful and complementary biophysical technique in the study of structural biology.

1.4.2 The evolution of biological mass spectrometry

The introduction of soft ionisation techniques such as, matrix assisted laser desorption ionisation (MALDI) but most notably electrospray ionisation (ESI), greatly enhanced mass spectrometry's involvement in structural biology.

The emergence of MS in conformational protein studies

One of the first instances presenting the application of ESI-MS in the study of protein conformational changes, discussed the observed charge state envelope of cytochrome c at different pH values. This work highlighted that cytochrome c's unfolded conformation has a wider charge state distribution, at a lower *m/z* ratio, compared to its native conformation. The number of charges which are accepted by the protein molecule is thought to be related to the number of basic amino acids, mostly Arg, His and Lys, exposed on the protein's surface (Chowdhury et al. 1990). It is therefore proposed that a denatured protein will carry more charge than a natively folded protein, as not only will it have more exposed sites but there should be less electrostatic repulsion between sites (Heck and Van den Heuvel 2004). The pioneering work of Chowdhury *et al*, and further subsequent work (Loo et al. 1991; Mirza et al. 1993), suggested that the unfolding of a protein in the solution phase could be captured by ESI-MS. This phenomenon has led to a method of investigating quaternary and tertiary structure, ion conformation and folding/unfolding, by looking at the charge distribution observed (Kaltashov et al. 2006), under varying conditions such as metal ion concentration and denaturant concentration. The denaturation of myoglobin by Konerman *et al.* highlights the complementary use of ESI-MS in such

experiments. In this instance in combination with absorption spectroscopy. During the unfolding process a transient intermediate was observed, using both techniques. It was from the mass spectrum, however, that the species was calculated to be unfolded holo-myoglobin by combining information from the observed *m/z* and charge state distributions (Konermann et al. 1997).

These experiments were only the start of the utilisation of mass spectrometry in the study of biomolecules. Further improvements since this pioneering work have led to continual advances in this area over the past two decades, to a point where mass spectrometry can now, not only study native biomolecules, but their complexes too.

Advances in MS and their application in the study of biomolecules

The sensitivity achieved by mass spectrometers has improved significantly, increasing its compatibility to biological samples. Advances in this area include the introduction of orthogonal ionisation sources (Z-Spray), reducing the neutral solvent passing into the skimmer and nano-ESI, implementing low flow rates (20-40 nL/min) (Wilm and Mann 1996). Further advantages of nano-ESI, over conventional ESI, include increased desolvation and reduced metal ion carry over, when used in conjunction with the study of metalloproteins (Mattapalli et al. 2009).

The mass ranges achieved by mass analysers are another important aspect required for the study of native complexes. Often native complexes are studied using TOF analysers due to their effectively unlimited mass range, high speed and sensitivity. The *m/z* range of quadrupoles has also, however, become of heightened importance whilst studying these complexes. Tandem MS experiments often require the filtering of a specific *m/z*, before CID (Heck and Van den Heuvel 2004) and with the molecular weight of samples forever increasing; along with their ions *m/z*, the range achieved by quadrupoles needed to increase. Quadrupoles are often limited up to between 4000 - 8000 *m/z*; however custom made (Sobott et al. 2002) and more recently commercial quadrupoles can now reach 32000 *m/z*. The transmission of ions through the mass analysers has also been greatly improved by implementing collisional cooling (Loo et al. 2005). In this process the pressure regime of certain regions of the mass spectrometry is increased. The increase in pressure creates more

neutral gas collisions with the large ions; which reduces their kinetic energy and improves their focussing (Chernushevich and Thomson 2004).

Ion mobility mass spectrometry

As described previously the implementation of IMMS, in particular drift cell and TWIMS adds a further dimension to the study of proteins. As IMS, and in particular TWIMS, operates over a short time frame, approximately 30 ms, the conformation observed is a reflection of the biologically relevant, solution-phase, conformation. The ability to observe a proteins conformation and how this can change under different solution (Smith et al. 2007) or experimental conditions (Ruotolo et al. 2007) has provided greater insight into intermediate species involved in protein unfolding/folding and disassembly. A further application allows for the observation of the conformational changes, occurring in biological mechanisms, such as the change in conformation of a protein upon ligand binding. These conformations can then be compared to structures determined from other techniques, including X-ray and NMR (Scarff 2008), using software approaches such as MobCal (Mesleh et al. 1996; Shvartsburg and Jarrold 1996).

1.4.3 Gas phase structure

To be able to confidently study the conformation of a protein accurately by mass spectrometry, one vital piece of information had to be proven. This was that the conformation that exists in the solution phase had to be shown to survive in the gas phase, for at least as long as a mass spectrometry experiment. The gas phase structure of proteins and their complexes have, as a result been studied at great lengths, yielding a plethora of supporting evidence, especially involving large proteins and complexes (Loo et al. 1993; Heck and Van den Heuvel 2004; Loo et al. 2005; Ruotolo et al. 2005).

Evidence has now started to emerge that the structure of small globular proteins, in the gas phase, evolves over time, Figure 1.15. Breuker and McLafferty summarised findings that show a number of stabilisation effects help preserve the solution conformation in the gas phase for a millisecond timescale. These include the aggregation of water molecules around charged side chains until the final stages of desolvation and the collapse of charged side chains, upon complete desolvation,

forming a stabilising electrostatic network (Breuker and McLafferty 2008). This type of electrostatic network has been shown by ECD experiments to compensate for the loss of hydrophobic interactions for as long as 4 s from the study of the helical KIX protein (Breuker et al. 2010). Studies, utilising a quadrupole ion trap IMS set up, on cytochrome c (Badman et al. 2005) and ubiquitin (Myung et al. 2002) demonstrate that the unfolding of compact structures produced from ESI, closely resembling the solution phase structures, do not occur within 30 ms and 20 ms respectively. More recent studies on the same native ubiquitin charge state indicate that under gentle instrument conditions this natively folded protein species can in fact be preserved for at least 100 ms (Wyttenbach and Bowers 2012). After this millisecond timescale however, unfolding starts to occur due to the loss of electrostatic interactions as well as hydrophobic ones. Eventually after seconds to minutes refolded gas phase structures are formed (Badman et al. 2005), which may have no reflection upon the solution structure, Figure 1.15.

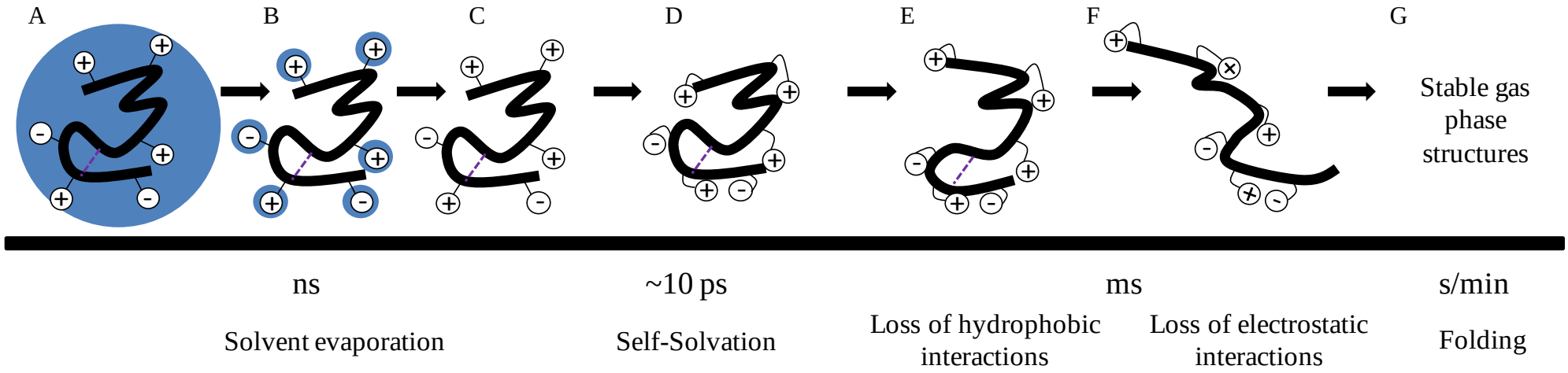


Figure 1.15: The structure of a globular protein, after ESI, evolving with time.
(A) Native protein, containing a hydrogen bond, surrounded by a monolayer of water, which undergoes water evaporation and associated cooling after a nanosecond time scale. (B) Exposed ionic groups on the protein remain hydrated, until further ns evaporation and cooling. (C) Solvent free protein self-solvates by collapsing exposed ionic groups. (D) After ms hydrophobic interactions are lost, leading to partial unfolding. (F) Further unfolding occurs by the loss of electrostatic interactions leading to transient unfolded ions. These ions go on to form more gas stable structures in seconds (G), which stabilise to energy minima conformation in minutes. Adapted from Breuker and McLafferty (2008).

The maintenance of proteins or complexes, in the gas phase, predominantly held together by hydrophobic interactions is an area of interest to mass spectrometrists. One would assume that upon desolvation of an ion during electrospray; the water, which drives hydrophobic interactions, evaporates along with any stabilising affect it may impart. Research in this area is particularly difficult due to the strengthening of the other interactions, in the mass spectrometer; known to stabilise proteins and complexes, such as electrostatic interactions and hydrogen bonds (Robinson et al. 1996). There have been many instances, summarised by Barylyuk *et al*, where a protein known to be predominately stabilised by hydrophobic interactions remains intact in the gas phase due to a small number of other interactions which only play a very small part in solution stabilisation (Boeri Erba et al. 2011). There are also many examples which demonstrate the loss of hydrophobic interactions in the gas phase, lead to complex degradation. A good example of this involves the proteins complexes RNase S and LZ. These complexes both have high solution stability, shown by their low K_d values, however once transferred into the gas phase, by ESI, either a very small amount or no complex was observed (Bich et al. 2010).

1.4.4 The study of DNA complexes

Mass spectrometry

The advances made in mass spectrometry, which has led to its application to the study of proteins, has also led to an increase of MS based studies on other biological macromolecules. These include oligonucleotides and their non-covalent complexes. Oligonucleotide studies using hard ionisation techniques were not able to produce intact molecular ions corresponding to two nucleotides and even the implementation of FAB, in the 1980's, only produced intact molecular ions of up to twelve nucleotides (Nordhoff et al. 1996). Although ESI and MALDI allowed for the ionisation of large oligonucleotides, up to megadaltons by ESI in the early 1990's, there were still many problems associated with the study of this type of biomolecule.

Oligonucleotides can be studied either using positive or negative ion mode. In positive mode protons attach to the nucleobases, with certain bases having higher proton affinities than others (Nordhoff et al. 1996), whereas the site for de-

protonation in negative ion mode is thought to be from the phosphate group (Moradian et al. 2002). It is therefore possible to visualise the molecular ions, however, the presence of negative phosphate groups within oligonucleotides, especially longer ones, can result in extensive binding of cations. This often leads to ion suppression and a reduced signal-to-noise (S/N) ratio in either mode, explaining the drop in signal intensity at higher masses. Sample preparation is therefore extremely important; ammonium acetate precipitation, the use of metal chelators, the inclusion of ammonium acetate in spraying solutions and the use of high performance liquid chromatography (HPLC) were all shown to significantly clean up spectra (Nordhoff et al. 1996).

There are now many examples of the preservation of non-covalent nucleic acid interactions in the gas phase. Examples in, positive mode, include the visualisation of a 16 base pair double stranded DNA (dsDNA) helix (Gupta et al. 2001) and its complexes with three different drugs. In negative mode observations have been made of duplex, triplex and quadruplex DNA (qDNA) (Rosu et al. 2002). Since 2002 numerous publications, summarised by Jennifer Beck, involve the study of qDNA, a structure formed from DNA sequences rich in G bases. These G rich sequences are of high biological importance as they are often found in genes of immortal cells and thus qDNA has become an anticancer drug target (Beck 2011).

Ion mobility mass spectrometry on DNA

Con-current with protein literature after the visualisation of these complexes in the gas phase, the conformation of nucleic acids and their complexes then began to be investigated (Hoaglund et al. 1997; Moradian et al. 2002; Gidden et al. 2005). The aim was to determine if the gas phase structures observed were representative of their solution phase counterparts. It had been suggested that desolvation of nucleic acids would expose the negative charges associated with the phosphate backbone, leading to charge repulsion and structural re-organisation. The initial work carried out by Haogland *et al.* on a 10 mer thymine oligonucleotide revealed that high de-protonation of oligonucleotides leads to elongation of its conformation, in comparison to the same oligonucleotide with fewer charges (Hoaglund et al. 1997). This observation suggested that oligonucleotides may behave in a similar way to proteins in the gas phase. Taking this idea further, cross-sectional comparisons

between DNA oligonucleotides and proteins of similar molecular weights were made. These indicated that DNA species were over 20 % smaller than their comparable protein species (Moradian et al. 2002). Although these experiments provided information on oligonucleotides in the gas phase they did not reveal whether solution structural elements were maintained, such as DNA helices. Ion mobility coupled with molecular dynamics has since indicated that DNA sequence length, charge and content affects the conservation of DNA structural elements in the gas phase (Gidden et al. 2005; Baker and Bowers 2007; Baker et al. 2009).

As indicated above the increase in charge carried by DNA in the gas phase leads to the elongation of DNA conformers due to charge repulsion. This charge repulsion also disrupts any secondary structural elements, such as helices, hairpins or pseudoknots and therefore the lower charge states will retain the most native conformation and are therefore studied.

Small DNA helices (< 8mer) were shown to quickly adopt globular conformations independent of base content. As in solution, however, DNA stability depends greatly on C·G content, with A·T pairs breaking more frequently than C·G pairs. Helical structures can be persevered, on a ms time scale, in C·G only DNA, in helices as small as 8mers, whilst the length of an A·T helix has to reach a 10mer before any helical structure is maintained (Gidden et al. 2005), Figure 1.16.

Although DNA helices are conserved in sequences as small as 8mers, they are not the B-helix found in solution. Upon desolvation they convert to the shorter wider structure of an A-helix. The preservation of the B-helix, although slightly bent, is not achieved until 22bp (C·G only) or 18bp (C·G and A·T mix) (Baker and Bowers 2007), Figure 1.16.

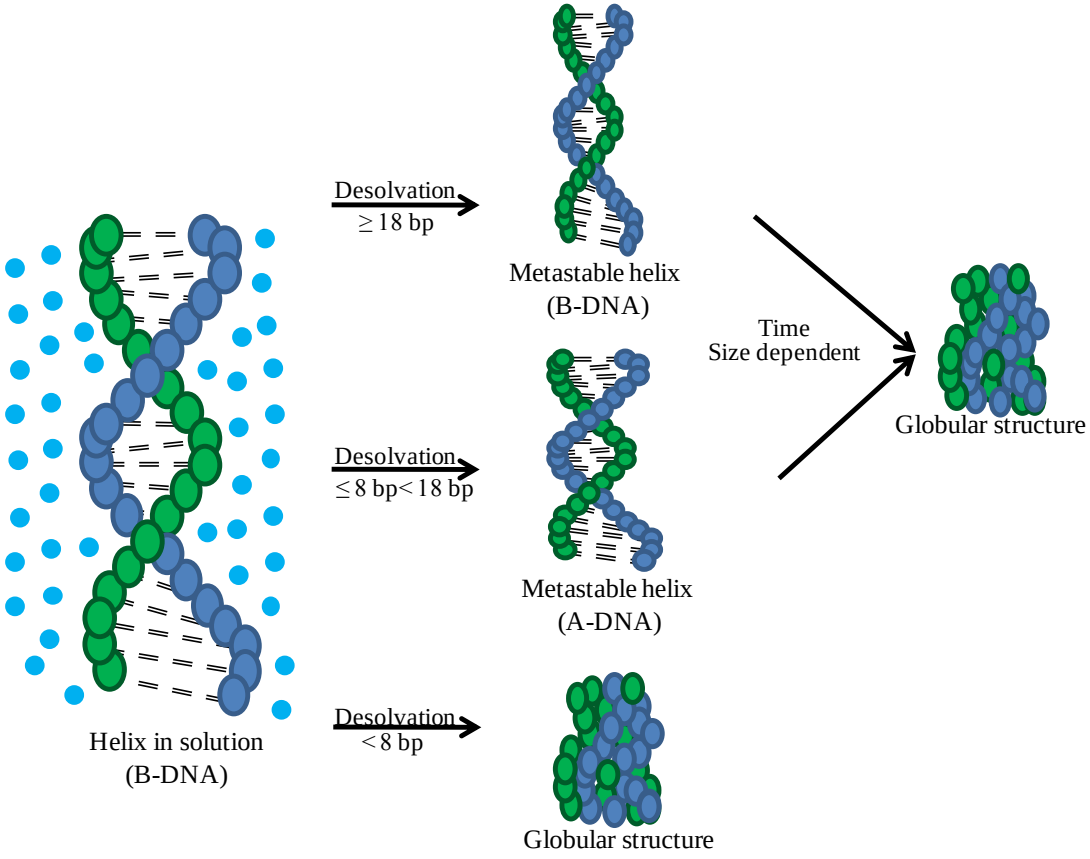


Figure 1.16: Duplex DNA in the gas phase
The potential sequence of events occurring upon DNA helices upon desolvation after electrospray ionisation. Adapted from (Gidden et al. 2005) and (Baker and Bowers 2007).

Studies have also investigated the structural effects of transferring DNA G-quadruplex and triplexes into the gas phase. G-quadruplexes were found to be transferred well and stabilised by ammonium ions (Gabelica et al. 2007), whereas triplexes were found to be distorted but like duplexes still maintained a helical structure (Arcella et al. 2012).

1.4.5 Protein complexes

The reason why proteins perform such a diverse range of functions, within biology, is down to their ability to interact with numerous binding partners. These partners can include other proteins or peptides, oligonucleotides, small molecules and metal ions (Loo 1997). The time scale of interaction can also vary; allowing long lived complexes, such as the multimeric iron containing protein haemoglobin or transient complexes, for example the transport of ions or proteins across membranes using protein transporters.

1.4.5.1 Protein-protein complexes

Interactions between proteins are seen in many areas of biology, from forming functional cellular complexes to aggregation of proteins in disease. The identification of subunits, within a multimeric protein complex, can be obtained from a denatured mass spectrum whilst the interaction between these subunits is often investigated using tandem mass spectrometry (Heck 2008). Its application can provide information regarding topology, structure and structure stability of subunit binding and is discussed in more detail elsewhere (Heck and Van den Heuvel 2004). The use of CID has provided information on subunit positioning in a number of proteins including P22 tail (Lorenzen et al. 2008) and RNA polymerase III (Lorenzen et al. 2007).

1.4.5.2 Protein-metal complexes

Metallomics, the study of metal ion containing proteins is often carried out by the combined use of elemental analysis, ESI-MS and to a lesser extent X-ray absorption or fluorescence spectroscopy (Shi and Chance 2008).

Elemental analysis

The quantitative evaluation of metals within samples, ranging from ceramics (Tagle et al. 2011) to blood plasma (Lane et al. 1999), can be conducted by; inductively coupled plasma optical emission spectroscopy (ICP-OES), inductively coupled plasma mass spectrometry (ICP-MS), or atomic absorption spectroscopy (AAS) (Kaltashov et al. 2006). ICP-MS/ICP-OES are highly sensitive techniques capable of multi-elemental analysis. Although AAS has lower metal ion detection limits it often only has the ability to quantify one element at a time (Shi and Chance 2008). Although some of the newer machines have the capability to analyse more elements consecutively (Lane et al. 1999), generally the lower throughput of AAS, means that the complexity within biological samples are more suited to ICP-MS/ICP-OES analysis.

ICP-OES

The acidified sample is introduced into the ICP-OES instrument, often by an automated sample manager. A nebuliser transforms the sample into micro droplets

which are desolvated. Once they enter the plasma, all of the atoms within the sample are excited and upon relaxation they emit light. This light is then detected and because each element emits at a specific wavelength, many elements can be identified and quantified (Lane et al. 1999). For each element of interest a set of standards, covering the required concentration range has to be made up and run.

As many elements, from an individual sample, can be detected and quantified simultaneously ICP-OES is often the technique of choice for determining metal ion binding stoichiometries (Vincenzetti et al. 1996) (Stola et al. 2006). For this application a non-metal element found in proteins must also be quantified often sulphur or phosphorus (Brewer and Marcus 2007). In an example where sulphur is chosen, the number of sulphur atoms within the protein must be known. This allows for the conversion from sulphur concentration, calculated from the ICP-OES results, into a protein concentration. A ratio can then be produced from the protein concentration and the concentration of the metal ion of interest.

As the ICP-OES plasma excites all the atoms within the sample, protein: metal ion stoichiometries are highly dependent upon a pure protein sample, with a single stoichiometry. This is because any contaminating sulphur, from proteins or protein fragments, leads to the calculation of a synthetically high protein concentration and so a low stoichiometry; whilst a mixture of stoichiometries will always result in a ratio representing the average of the solution. Due to these drawbacks the results gathered from ICP-OES are compared with other techniques, mentioned previously.

Mass spectrometry

Although elemental analysis can quantify and identify metal ions within proteins, due to its destructive nature further information concerning the complexes structure and dynamics are lost. ESI-MS not only has the capability to maintain non-covalent interactions, allowing structural studies to be undertaken, but also has the unique ability to visualise all species, within a sample independently. The advantage of this is that exact stoichiometries can be determined for all the species present, and not an average of the sample. These qualities have seen the application of ESI-MS to many metalloprotein studies; allowing for the determination of metal ion stoichiometry in a vast number of proteins (Yu et al. 1993; Hu et al. 1994). With investigations into

metal ion acquisition and dissociation, revealing some instances of metal binding cooperatively (Jensen et al. 1998; Lei et al. 1998; Gehrig et al. 2000) and changes in protein conformation, in conditions comparable to cellular compartments, upon metal ion release (Gumerov and Kaltashov 2001; Gumerov et al. 2003). Recently the application of ESI-MS has advanced to measuring the metal ion stoichiometry of SOD1 directly from diseased tissue (Rhoads et al. 2011).

1.4.5.3 Protein-DNA complexes

There are many cellular processes which depend upon the formation of complexes between proteins and DNA. Examples of these processes include DNA replication, protein production and transcriptional control.

The presence of conserved secondary structural motifs, indicative of DNA binding allows for the identification of potential DNA binding proteins using bioinformatics. These motifs include; the ubiquitous helix-turn-helix, leucine zipper and zinc finger. All these motifs allow proteins to interact with either ssDNA or dsDNA through a varying combination of non-covalent interactions, including hydrogen bonds, van der waals forces, ionic interactions and hydrophobic interactions.

Protein-DNA complex mass spectrometry

The application of ESI-MS techniques for the visualisation of protein:DNA complexes began in the mid 1990's, with mass spectra obtained for a number of proteins bound to ssDNA and dsDNA. In many of these instances MS was used as a method to confirm complex formation and to accurately determine the protein:DNA stoichiometry (Beck et al. 2001). Some of the complexes spectra were obtained in positive and some in negative mode, revealing the individuality of the species in question. One of the first instances showing specific interactions between a DNA binding protein and dsDNA, observed by ESI-MS, used the DNA binding domain of human PU.1 (PU.1-DBD). This protein was incubated with a mix of two different dsDNA species, with only one containing the GGAA binding sequence. FTICR spectra revealed a 1:1 stoichiometry of PU.1 with the dsDNA containing GGAA, but no binding to the DNA without this sequence. There are many further examples exhibiting specific interactions between DNA and proteins (Kapur et al. 2002), revealing the formation of these complexes in the solution phase and not un-

specifically during ionisation. Although in some instances low abundant non-specific interactions are identified (Deterding et al. 2000).

More recently TWIMS has been employed to investigate protein: DNA complexes. An increase in collisional cross-section corresponding to 5 % has been observed upon the binding of fructose-1, 6-bisphosphate to a CggR: DNA complex (Atmanene et al. 2010) This evidence shows the capability of studying protein: DNA complexes and their associated conformational changes by ion mobility mass spectrometry.

1.5 Aims and objectives

The work in this thesis investigates the biomolecular components and interactions within a zinc ion homeostatic system, by means of mass spectrometry-based approaches.

The use of ESI mass spectrometry for the study of stoichiometry is a favoured technique, as it enables the visualisation of all species within a mixture simultaneously but individually. This technique was applied to the study of purified SmtB protein, cytosolically expressed in the presence of zinc, and apo-SmtB in the presence of a DNA sequence, containing a characterised SmtB binding site.

The application of ion mobility mass spectrometry to the study of protein conformational changes has become an established technique. The introduction of a second generation Synapt G2 HDMS system in December 2009 has increased the resolution of ion mobility allowing for the investigation of smaller conformational differences. The different conformations adopted by SmtB and SmtA in the presence of varying zinc ion stoichiometries was investigated.

The aims of the project were:

- Express SmtB in the presence of metal ions, usually Zn (II) but Cd (II) was also investigated. Purify the resulting protein in the absence of further metal ions and characterize the protein produced by ICP-OES, denatured MS and far-UV CD (Chapter Two).

- Determine the SmtB binding stoichiometry of SmtB expressed in the bacterial cytosol by native MS. Compare the conformation of SmtB with different stoichiometries of zinc ions bound by chelation or competition experiments combined with ion mobility mass spectrometry measurements (Chapter Three).

- Study the oligonucleotides and the annealed DNA of the S2/S1 SmtB binding site by ion mobility mass spectrometry. Produce apo-SmtB from the previously purified metal bound sample and combine it with the cognate DNA sequence to form a protein: DNA complex (Chapter Four).

- Express and purify SmtA, study the protein using ion mobility mass spectrometry. Look for the possible interaction between SmtA and SmtB (Chapter Five).

1.6 Research and Conference papers

1.6.1 Research Papers

One peer-reviewed paper has been produced from the results presented in this thesis.

Kondrat, F. D. L., Kowald, G. R., Scarff, C. A., Scrivens, J. H. and Blindauer, C. A. Resolution of a paradox by native mass spectrometry: facile occupation of all four binding sites in the dimeric zinc sensor SmtB. (Submitted).

1.6.2 Conference Papers (Peer-reviewed)

Kondrat, F. D. L., Scarff, C. A., Blindauer, C. A. and Scrivens, J. H., Conformational changes associated with the removal of zinc ions from the cyanobacterial regulator SmtB, *Proc. 3rd International Symposium on Metallomics*, 2011, Munster, Germany.

Martin, E. M., Lu, J., **Kondrat, F. D. L.**, Scrivens, J. H., Sadler, P. J., and Blindauer, C. A., Probing allosteric interactions between metal ion and fatty acid binding to human serum albumin: an electrospray mass spectrometry approach, *Proc. 15th international conference for biological inorganic chemistry*, 2011, Vancouver, Canada.

Kondrat, F. D. L., Scarff, C. A., Leszczyszyn, O. I., Kowald, G. R., Blindauer, C. A. and Scrivens, J. H., Conformational changes associated with the removal of zinc ions from proteins in a cyanobacterial Zn^{2+} homeostatic system, *Proc. 58th ASMS Conf. on Mass Spectrometry and Allied Topics*, 2010, Salt Lake City, USA.

Kondrat, F. D. L., Scarff, C. A., Leszczyszyn, O. I., Kowald, G. R., Blindauer, C. A. and Scrivens, J. H., Conformational changes associated with the removal of zinc ions from proteins in a cyanobacterial Zn^{2+} homeostatic system, *5th IMMS MMX*, 2010, Chester.

Scarff, C. A., **Kondrat, F. D. L.**, Booyjzsen, C., Mukherjee, A., Sadler, P. J. and Scrivens, J. H., Probing transferring-metal interactions by means of travelling-wave ion mobility mass spectrometry, *Proc. 58th ASMS Conf. on Mass Spectrometry and Allied Topics*, 2010, Salt Lake City, USA.

Scrivens, J. H., **Kondrat, F. D. L.**, Scarff, C. A., Blindauer, C. A., Sanghera, N., Hilton, G. R., Gill, A. C., Pinheiro, T. and Thalassinos, K., A shape selective study of conformational changes in metal containing proteins, *Proc. 57th ASMS on Mass Spectrometry and Allied Topics, 2010,* Salt Lake City, USA.

Kondrat, F. D. L., Scarff, C. A., Leszczyszyn, O. I., Blindauer, C. A. and Scrivens, J. H., An ion mobility mass spectrometry-based study of the metalloprotein SmtB, *Proc. 18th International Mass Spectrometry Conference*, 2009, Bremen, Germany.

Kondrat, F. D. L., Scarff, C. A., Blindauer, C. A. and Scrivens, J. H., The study of SmtB using ion mobility mass spectrometry-based approaches, *The Royal Society of Chemistry 2009 Analytical Research Forum*, University of Kent, Canterbury.

Chapter Two: SmtB expression and purification

2.1 Introduction

The activation of the *smt* operon, in *Synechococcus elongatus* PCC 7942, which leads to an increase in production of *smtA* transcripts, was achieved by the addition of a number of metal ions; most notably zinc and cadmium ions (Huckle et al. 1993). The strongest response was observed after the introduction of zinc ions, (Huckle et al. 1993) which indicated that zinc ions were the most likely to be the cognate metal ions for this system. This observation was confirmed by two supporting findings; the gene product SmtA was shown to have a high affinity for zinc ions and the deletion of the *smt* gene increased the sensitivity of the cells against zinc and cadmium (Turner et al. 1993).

Previous expression and purification methods for recombinant SmtB, were either partial (Turner et al. 1996) or highly involved and resulted in relatively low expression yields of between 50-100 mg of apo-SmtB protein per 24 L of bacterial culture (Kar et al. 1997). The apo-SmtB produced in these instances was only short lived, even when stored at -80 °C; with the oxidation of the free cysteines observed after a few months (VanZile et al. 2000). The reduction of pH to pH 6, and the inclusion of DTT was found to prolong the life of the protein, at 4 °C, but only for a few weeks (Kar et al. 1997).

Aims

This chapter outlines the expression and purification of SmtB, employed in this research which was adapted from previous work conducted by Gregory Kowald. SmtB was expressed in the presence of the predicted cognate metal ion, zinc (II). The aim of this was to increase the stability of the protein, which should in turn increase the achieved yield, prevent truncation and aid in long term storage. The production of metal bound SmtB would also allow for an accurate determination of the metal: protein stoichiometry, without the addition of further metal ions. Efforts were also undertaken to express and purify SmtB in the presence of cadmium following the experimental observations outlined above.

2.2 Material and Methods

2.2.1 Chemical and Materials

All reagents were obtained from Sigma-Aldrich UK or Fisher Scientific UK unless highlighted within the text. These were of high quality and purity, to reduce the risk of metal ion contamination. Solvents, including water, used during mass spectrometry were of analytical grade. Any equipment used is presented in the text.

2.2.2 Expression

The expression of SmtB (mononeric Mw 13413.31), Figure 2.1, was carried out in the presence of zinc ions, to ensure relevant protein folding and stability.

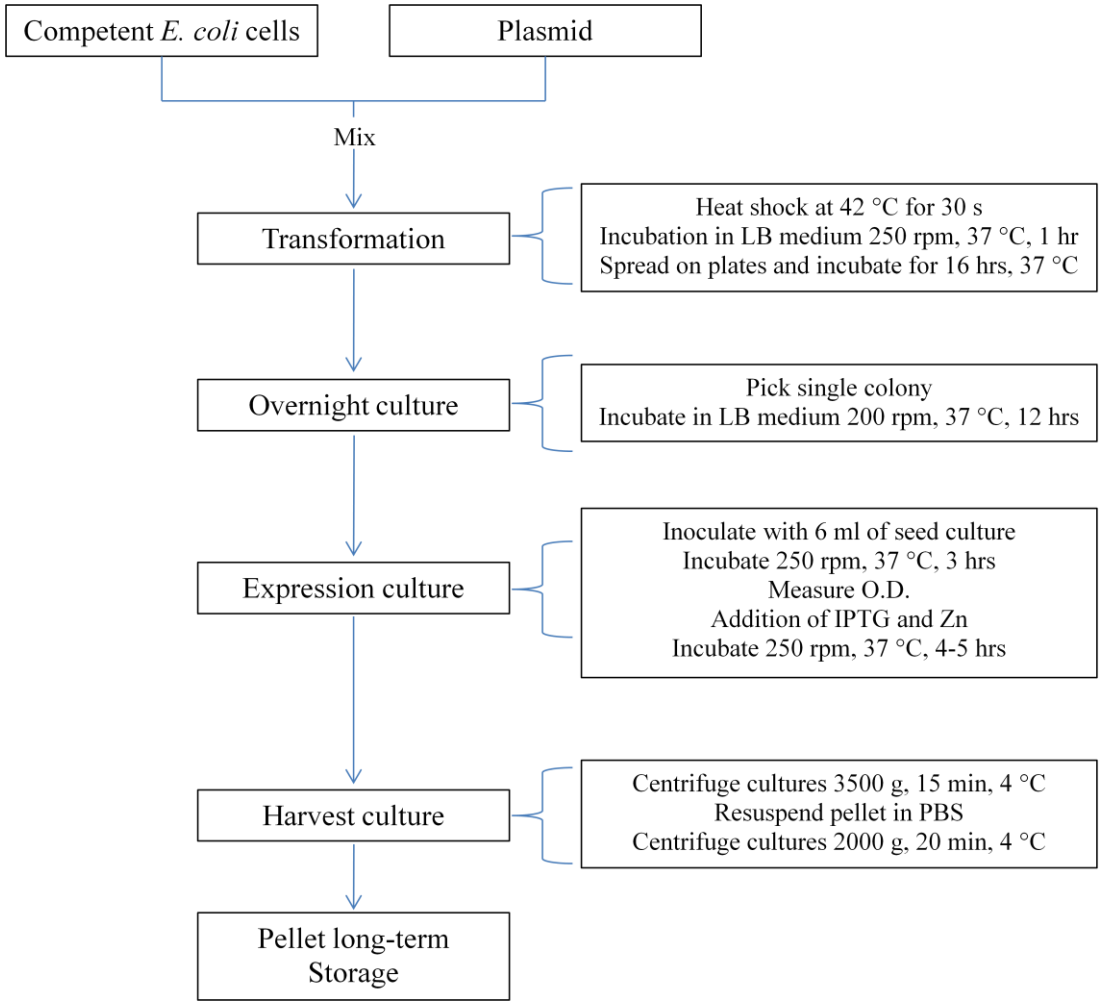


Figure 2.1: A flow diagram of the expression method

In addition to the expression of SmtB with zinc ions, expression was also conducted in the presence of cadmium ions. This was to investigate the effect of a chemically similar non-cognate metal ion binding to SmtB. In this round of expression all procedures were identical to those outlined below, for the expression of SmtB with zinc. During gene expression, however, cadmium chloride solution was added instead of zinc sulphate solution.

2.2.2.1 Transformation

Competent *E. coli* BLR(DE3)pLysS (Novagen) were transformed following the Novagen® Competent Cells protocol (Novagen 2004). A SmtB plasmid pSRK15-1, kindly donated by N. Robinson, was used as the DNA source. Two 20 µL cell aliquots were thawed on ice before the addition of 1 µL of DNA to one of the tubes leaving the other as a control. After this step both tubes were treated identically. The tubes were then gently mixed, placed on ice for 5 minutes before being heat shocked in a 42 °C water bath for precisely 30 seconds. Once out of the water bath the tubes were placed back on ice for a further 2 minutes before the addition of 80 µL of room temperature Luria Broth (Miller's LB broth). The tubes were then transferred into a 37 °C incubator (New Brunswick Scientific Innova 44) and shaken at 250 rpm for an hour.

For each transformation four agar plates, containing LB, kanamycin 25 µg/mL, chloramphenicol 34 µg/mL and tetracycline 12.5 µg/mL, were used. BLR(DE3)pLysS competent cells contain selectable resistance markers against chloramphenicol and tetracycline. This allowed for selection against contaminating cells. The SmtB plasmid pSRK15-1 carries a resistance marker to kanamycin, which provides a selection against non-transformed cells. Transformed cells, in volumes of 10, 30 and 50 µL, were spread upon three of these selective plates to ensure single colonies could be obtained. Two control plates were created by aliquoting 30 µL of the control cells onto a LB only plate (positive control) and a LB and antibiotic plate (negative control). These plates were then allowed to dry and then were placed upside down in a 37 °C incubator for 16 hours. The plates were then sealed and stored at 4 °C before use.

2.2.2.2 Overnight cultures

Overnight cultures were created, in duplicate, by scraping a single plate colony into 50 mL of sterile Miller liquid broth pH 7.5 containing kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL). These cultures were shaken at 200 rpm for 16 hours at 37 °C

2.2.2.3 Gene expression

For protein expression 6 mL of overnight culture was added to 2-4 x 600 mL of sterile Miller liquid broth pH 7.5 containing kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL). These cultures were shaken at 250 rpm at 37 °C until the optical density of the culture reached 0.6-0.8, at 600 nm; this took approximately 3.5 hours. Sterile, filtered zinc sulphate was added to the culture to a final concentration of 5.0×10^{-4} M, followed by the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Melford, UK), to a final concentration of 6.25×10^{-4} M. These additions allowed for the induction of protein expression at sufficient but not toxic zinc ion levels.

In the expression of SmtB with cadmium, sterile filtered cadmium chloride was added to a final concentration of 2.0×10^{-4} M. The cultures were then shaken for another 4-5 hours before harvesting.

2.2.2.4 Culture harvesting

The bacterial cultures were harvested by centrifugation, at 3500 g for 15 minutes at 4 °C (Sorvall RC-6 Plus). The pellets produced were then re-suspended in 30 mL of phosphate buffered saline (PBS) pH 7.4 and transferred into 50 mL falcon tubes before being centrifuged at 2000 g for 20 minutes at 4 °C (Heraeus® Biofuge® Primo R). The pellets were then stored at -80 or -20 °C.

2.2.3 Purification

The purification of SmtB, described in Figure 2.2, was carried out in the absence of added zinc ions to prevent un-specific metal ion binding to SmtB.

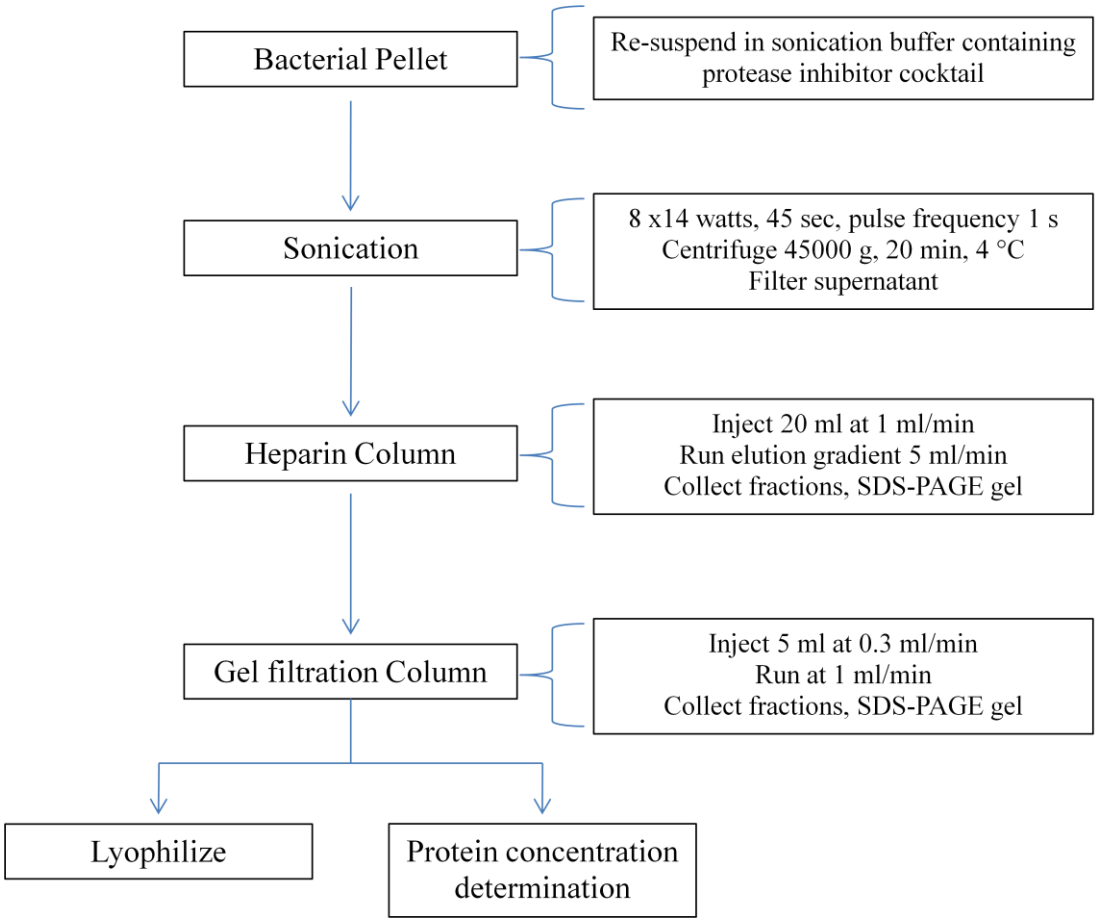


Figure 2.2: A flow diagram of the method used to purify zinc bound SmtB

2.2.3.1 Sonication

In order to purify SmtB protein from *E. coli* the cells had to be first broken open, by sonication. A 5 mL volume of HEPES buffer A (20 mM HEPES, 50 mM NaCl, 0.5 mM DTT (Melford,UK) pH 7.6) was added to the bacterial pellet along with an EDTA free protease inhibitor cocktail solution (Sigma). This mixture was sonicated on ice (Sonics Vibra Cell) at 14 watts for 45 seconds using a pulse frequency of 1 second; this was repeated eight times with sufficient time gaps to prevent the solution over-heating, since this could result in protein degradation. Once the pellet was completely re-suspended, the lysate was transferred to centrifuge tubes and centrifuged at 45000 g for 20 mins at 4 °C. The supernatant produced was diluted up

to 20 mL, using HEPES buffer A, before being syringe filtered (Sartorius Minisart pore size 0.2 µm). The resulting sample was then ready for chromatography.

2.2.3.2 Fast protein chromatography

All fast protein liquid chromatography (FPLC) steps were carried out using a Pharmacia Äkta™ Purifier system coupled to a F950 fraction collector.

2.2.3.2.1 Heparin affinity chromatography

The first step in the purification procedure was affinity chromatography, exploiting the DNA-binding ability of SmtB and utilising a Hi Trap™ heparin HP column. The sample was placed into a superloop, which had a maximum volume of 50 mL, and runs were preformed with 20 mL injections being loaded onto the column. Previous work had indicated that this was the maximum loading volume before peak broadening and sample loss occurred. The run included equilibrating the column with 5 column volumes of HEPES buffer A running at 5 mL/min after which the rate was slowed to 1 mL/min during sample injection. A linear gradient of HEPES buffer B (20 mM HEPES, 1 M NaCl, 0.5 mM DTT, pH 7.6) was then set up to elute the heparin binding proteins; this started at 0 % and reached 100 % over 20 column volumes. At the end of the run the column was washed with 3 column volumes of HEPES buffer B to make sure all proteins were removed.

2.2.3.2.2 Gel filtration / size exclusion chromatography

The fractions which showed absorbance at 280 nm were run on a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gel to confirm the presence of SmtB. The fractions were then pooled and poured into a super-loop and loaded onto a size exclusion chromatography (SEC) column, (16/60 HiLoad™ 75 Superdex™ prep grade, Amersham Biosciences). The buffer used was 20 mM ammonium bicarbonate, which had been previously filtered and degassed. Whilst sample injection was taking place the rate was 0.3 mL/min, which was increased to 1 mL/min for the remainder of the run. Confirmation of SmtB within fractions was then accomplished by the use of SDS-PAGE gels. These fractions were pooled and a Cys assay, A_{280} assay or inductively coupled plasma absorption emission spectroscopy (ICP-OES) was conducted to determine protein and metal ion

concentrations. Samples were then aliquoted and lyophilised, and once dry they were stored at -80 °C until further use.

2.2.4 Protein concentration determination

A number of different approaches were employed to determine the concentration of SmtB in FPLC fractions.

2.2.4.1 Absorbance at 280 nm

FPLC fractions containing SmtB were concentrated from 4 mL to 1 mL using an Amicon Ultra-4 10 kDa cut-off centrifuge filter units (Millipore). They were then loaded into a quartz cuvette, with a path length of 1 cm. The samples were measured in triplicate and before each absorbance reading at 280 nm a blank was measured. The average absorbance was determined and then the concentration of the sample (C) was determined using the Beer-Lamberts law. A pathlength (L) of 1 cm and an estimated extinction coefficient (ε) for SmtB, 5960 M^{-1} cm^{-1} (VanZile et al. 2000) were used.

$$A_{280} = C \times L \times \varepsilon$$

Equation 2.1

2.2.4.2 Cysteine assay

The Cys assay is a modified version of the Ellman's assay (Ellman 1959) and entails the inclusion of the chelator EDTA at a final concentration of 1 mM. EDTA was used to remove the bound metal ions from the cysteine residues within SmtB. A set of standards were prepared including 2.6 mL of 0.1 M Tris buffer (pH 7.05) and 2.5 mM DTNB in 50 mM ammonium acetate (pH 5.0) but with five different concentrations of cysteine. A standard curve was then created by measuring the absorbance of the solutions at 412 nm.

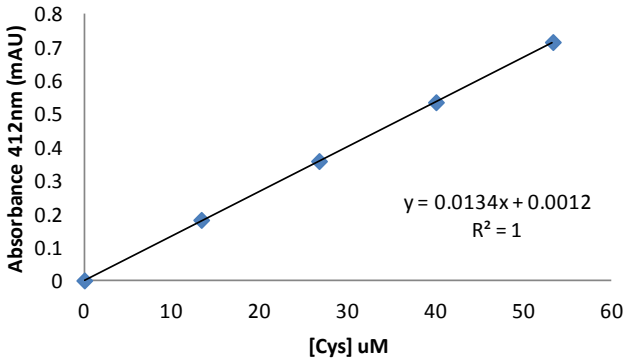


Figure 2.3: An example of a cysteine assay standard curve

The gradient of the standard curve was used to determine the concentration of free cysteines within the sample. A test solution was prepared using a SmtB containing fraction in place of cysteine. This was left to incubate at room temperature for 15 minutes and then the absorbance measured.

$$A_{412} \text{ x dilution factor/standard slope} = [\text{free cysteine}]$$

Equation 2.2

This value can then be converted into the concentration of monomeric SmtB, with the knowledge that each SmtB monomer contained three cysteines.

2.2.4.3 ICP-OES

Samples were either taken directly from the FPLC and concentrated using Amicon Ultra-4 10 kDa cut-off centrifuge filter units (Millipore) or lyophilised samples were re-suspended into 20 mM ammonium bicarbonate. Once an appropriate concentration was reached these were then diluted 1 in 5 into 0.1 M nitric acid. Standard curves were obtained from a set of standards, by measuring at the specific wavelength for each element. The standards, which contained 0, 0.2, 0.5, 0.7 1.0 and 2.0 ppm of the elements; zinc, sulphur, cadmium and copper, were monitored at 213.857 nm, 180.669 nm, 228.802 nm and 324.752 nm, respectively. Sample data was then obtained, in counts, and was converted into molar concentrations before the ratios of the elements were calculated.

2.2.5 Chacterisation

Characterisation of SmtB was carried out at different stages to confirm the presence of SmtB in FPLC fractions, and to determine the molecular weight and secondary structure of the isolated protein.

2.2.5.1 SDS-PAGE

To determine which fractions from the FPLC runs contained SmtB, SDS-PAGE gels (4-12% gradient NuPAGE® gel, Invitrogen) were used. Samples were loaded from each fraction corresponding to a peak in absorbance along with molecular markers (SeeBlue Plus2 pre-stained standard). Prior to loading, the 6.5 µL samples with the addition of 2.5 µL loading buffer and 1 µL reductant were denatured by boiling for two minutes. After running at 200 V for 40 minutes, the gel was stained with colloidal coomassie blue for at least an hour. De-staining required incubation and agitation in distilled water with numerous exchanges, until bands could be distinguished from the background. Gels were then photographed using a BioDoc-it™ imaging system.

2.2.5.2 ESI mass spectrometry

Denatured samples to be exaimened by mass spectrometry were concentrated and desalted up to five times using 10 kDa cut-off Amicon Ultra-0.5 centrifugal filter devices (Millipore) in 10 mM ammonium bicarbonate. The concentrated sample was diluted in 30 % acetonitrile, 0.2 % formic acid to a concentration of approximately 10 µM.

2.2.5.3 Circular Dichroism

A freeze-dried sample of SmtB was re-suspended in 10 mM ammonium bicarbonate and desalted using 10 kDa cut-off centrifugal filter devices (Millipore). The sample was diluted to 13 µM and placed in a 1mm path length quartz cuvette. A JASCO J-815 spectropolarimeter, maintained at 20 °C, was used in continuous scanning mode with a scanning speed of 100 nm/sec, between 190 nm and 250 nm in 0.5 nm increments. A total of 16 acquisitions were averaged and a buffer scan was subtracted resulting in the final spectrum.

2.3 Results and discussions

Zinc bound SmtB was successfully produced utilising a pET expression system after transformation of the SmtB plasmid into BLR(DE3)PLysS cells. This was followed by a two step purification which resulted in a decent yield of un-truncated protein. Yields varied between expressions; but typically yields for SmtB varied between 2.79-3.79 mg per litre of cell culture. These values are comparable with Kar *et al*, who achieved 2.08-4.17 mg (Kar et al. 1997). The purification process here was simpler and produced a more stable product.

2.3.1 Heparin affinity chromatography

The crude extract, obtained from sonication, was filtered and loaded onto a heparin column. Heparin is a glycan, which can either act as a cation exchanger or an affinity ligand, especially to DNA binding proteins, steroid receptors and growth factors (GEHealthcare 2007). The elution profile of proteins within the crude extract was monitored by absorbance at 280 nm and 220 nm, Figure 2.4.

A constant elution of proteins was observed from just after the injection of the sample onto the column, time zero, until an elution volume of approximately 140 mL. The 220 nm absorbance was saturated and so could not be used to provide any useful information. The 280 nm absorbance provided a clear elution profile of the proteins. Whilst the sample was loaded onto the column the A280 observed was high, at around 3000 units. This UV absorbance is caused by the proteins which have no affinity to heparin and therefore pass straight through the column. After the initial high absorbance the value dropped off rapidly until the introduction of the NaCl gradient. The salt within the buffer competed with the proteins binding to the heparin, and eluted them at specific concentrations. This produced a number of separate smaller peaks, each containing only a few proteins, with an absorbance of approximately 800 units.

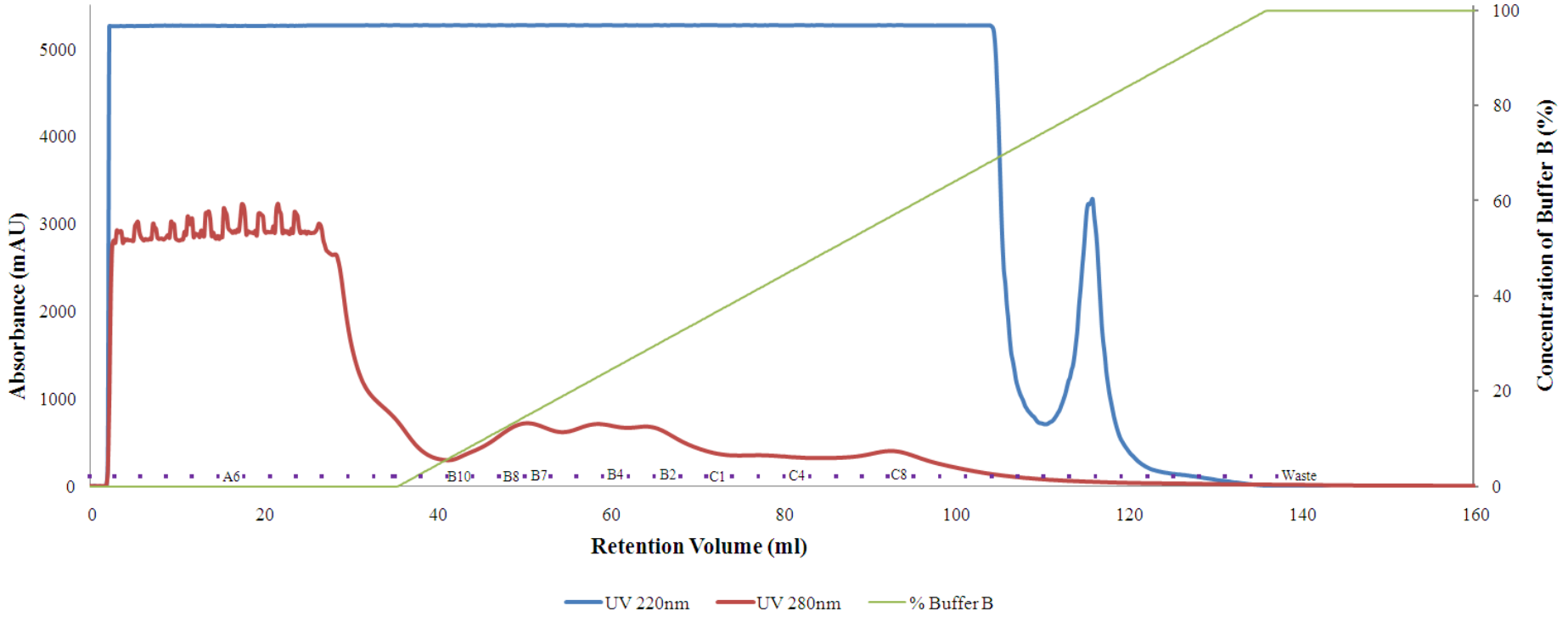


Figure 2.4: The elution profile of *E. coli* crude extract proteins from a heparin column.
Upon the increase of percentage of buffer B (green). The elution was followed by monitoring the absorbance at two different wavelengths, 220 nm (blue) and 280 nm (red). 3 mL fractions were collected across the profile.

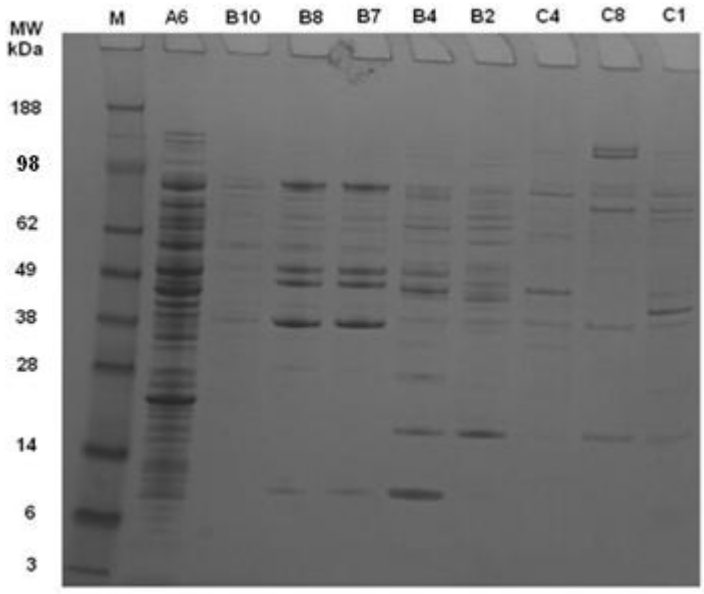


Figure 2.5: SDS-PAGE gel of the fractions collected from the heparin column. Lane M was loaded with pre-stained molecular weight markers, their weight is indicated to the left of the gel. The subsequent lanes were loaded with 10 µL of sample from the fractions indicated above the lanes.

An SDS-PAGE gel was run from the fractions collected across the elution profile, Figure 2.5. This was used to determine if zinc bound SmtB had a sufficient affinity to heparin and a narrow enough elution profile to separate it from the highly complex mixture of proteins within the crude extract. Fraction A6 shows the proteins which had no affinity to heparin. The elution profile and the gel both show that a high quantity of crude extract proteins did not bind to the column. The SmtB monomer has a molecular weight of around 13 kDa, and although there were a few bands around this molecular weight in A6, they are not intense enough to suspect the presence of SmtB. An intense band, at the correct molecular weight in A6, would have either indicated that SmtB had a low affinity for heparin or that the column had been over-loaded with extract. The only sample which had a clear band around 13 kDa was fraction B4, whilst fraction B7+8 also had a faint band at the same molecular weight. In each of these fractions contaminating proteins were also highly abundant. The band likely to correspond to SmtB appears to run slightly lower than its expected weight, although within an acceptable region for the accuracy of this technique. Size exclusion chromatography was then used for further purification.

2.3.2 Size exclusion chromatography

The fractions containing SmtB were then run through a size exclusion column and the elution of proteins from the column was followed by absorbance at 220 nm and 280 nm, Figure 2.6.

The SmtB purification procedure used did not initially contain a size exclusion step but instead used PD-10 Desalting columns (Amersham Biosciences) to desalt samples before analysis. The elution profile, Figure 2.4, as well as the SDS gel, Figure 2.5, indicated high levels of sample impurity. These other protein species would not be removed by PD-10 columns (Amersham Biosciences), since they only had the capacity to separate species of 5000 Da or higher from species less than 1000 Da (Biosciences 2003). They therefore would not only interfere with protein concentration estimation, but also with the mass spectra which could contain a number of species which could lead to the suppression of SmtB peaks. The size exclusion column was run in a mass spectrometry compatible buffer, 20 mM ammonium bicarbonate containing no sodium chloride. This meant that samples would be simultaneously purified, buffer exchanged and desalted whilst they ran though the column. This buffer exchange and desalting step removed the requirement for PD-10 columns (Amersham Biosciences) which were time consuming and led to sample dilution without further purification.

Due to the lower concentration of proteins observed on the SEC run, when compared to the heparin run, the absorbance collected at 280 nm, from aromatic residues within the proteins, was very low and therefore the absorbance from the peptide bonds at 220 nm was used. There was an initial void volume before the elution of three main peaks, at approximately 40, 60 and 100 mL. The initial peak from the SEC elution profile was likely to correspond to high molecular weight contaminants, which had been observed in the fractions by SDS-PAGE, Figure 2.5. SmtB would be expected to interact with the column more, due to its low molecular weight, and therefore should elute nearer the end of the run. A SDS-PAGE gel was run to confirm the presence of SmtB within the fractions eluted from the size exclusion column.

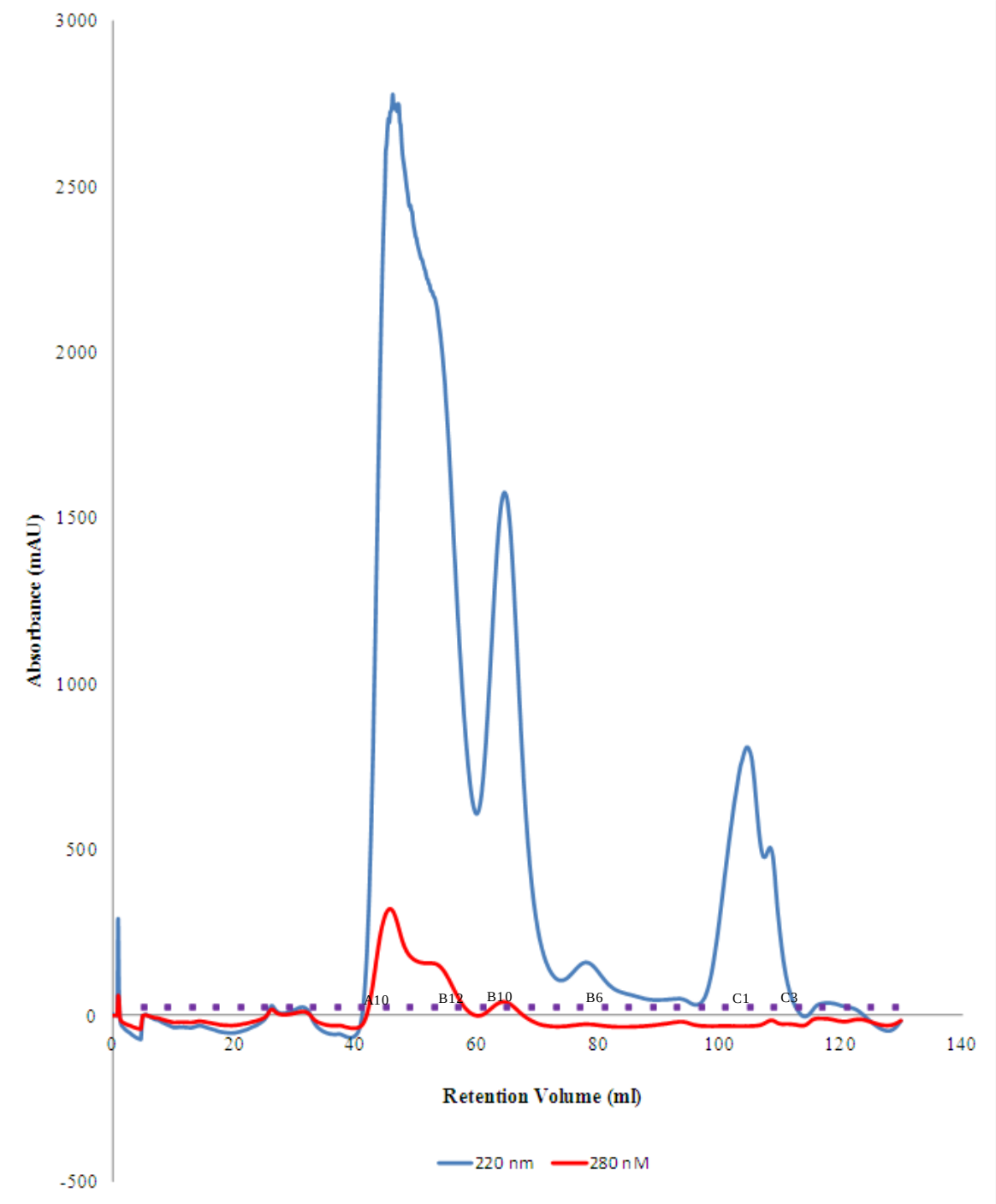


Figure 2.6: An elution profile of pooled heparin fractions from a SEC column
The elution was monitored following the absorbance at 220 nm (blue) and 280 nm (red). SDS-PAGE was conducted on the fractions indicated.

The SDS-PAGE gel, Figure 2.7, shows weak bands corresponding to low concentrations of the high molecular weight proteins in fraction A10. Stronger bands were observed in the next lane; however these are still too high in molecular weight

to correspond to SmtB. The most intense band was located in fraction B10. This was estimated at between 6 and 14 kDa which was in the correct range for the SmtB monomer. No subsequent fractions produced bands on the SDS-PAGE gel. The absorption peaks could be attributed to low molecular weight proteins at concentrations below the limits of coomassie gel detection (Wilkins et al. 1996) or small molecules within the buffer that have an absorbance at 220 nm.

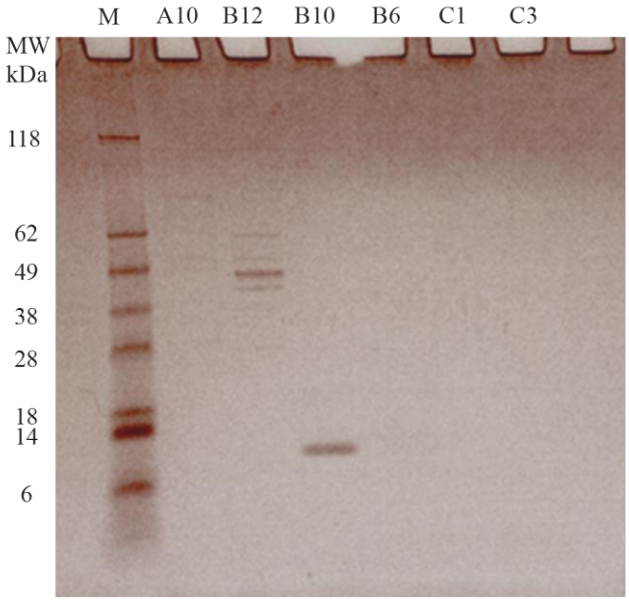


Figure 2.7: A SDS-PAGE gel of the fractions collected from the SEC column
Lane M was loaded with pre-stained molecular weight standards; their weight is indicated to the left of the gel. The subsequent lanes were loaded with 10 µL of sample from the SEC fractions indicated above the lanes.

2.3.3 Denatured mass spectrometry

To confirm that the purified product corresponded to SmtB, a sample from B9/B10 was denatured and analysed using ESI mass spectrometry, Figure 2.8.

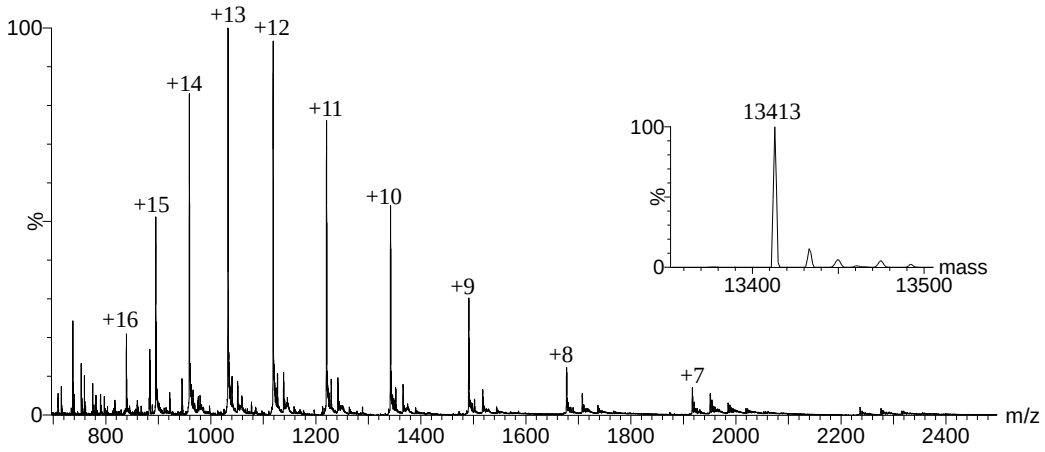


Figure 2.8: Non native ESI mass spectrum of SmtB
SmtB concentration was 10 µM in 30 % acetonitrile, 0.2 % formic acid. Inset shows the deconvoluted mass.

The large charge state distribution observed was typical of a denatured protein studied by ESI mass spectrometry. As the protein is unfolded there are more avalible basic sites for proton attachment (Chowdhury et al. 1990) . Each peak corresponds to the monomeric SmtB bound to a different number of protons. These peaks were deconvoluted onto a true mass scale providing an observed neutral mass of 13413 Da, see Figure 2.8 inset. This mass agrees with the mass observed by Kar et al (Kar et al. 1997) and corresponded to that of the SmtB peptide after cleavage of the N-terminal methionine, 13544.5 Da – 131 Da = 13413.3 Da. This form of peptide processing has been shown to occur in bacterial enzyme systems, after removal of the N-formyl from the initiating N-formylmethionine. It is performed by a protease known as methionine aminopeptidase. This enzyme cleaves the N-terminal methionine if it is next to an amino acid with a small side chain, such as alanine, valine or threonine (Sherman et al. 1985), the amino acid next to methionine in the SmtB polypeptide. N-terminal sequencing has previously shown that the N-terminal methionine can be often cleaved from the SmtB polypeptide (Turner et al. 1996).

2.3.4 Circular Dichroism

The secondary structure of the purified SmtB was determined using far-UV circular dichroism (CD). The spectrum produced, Figure 2.9, had one maximum and two prominent minima at 193 nm, 208 nm and 222 nm, respectively. These measurements correspond to specific n to π* and π to π* transitions which were indicative of a predominantly alpha helical structure.

The CD results which indicated a predominantly alpha helical structure agree with the X-ray crystallography structures present within the protein data bank (PDB). These structures contain 0 (1R1T), 1 (1r23) and 2 (1R22) (Eicken et al. 2003) zinc ions. A study of the published secondary structures reveals that there are no secondary structural changes upon zinc ion binding and so no indication of stoichiometry could be observed from a CD spectrum.

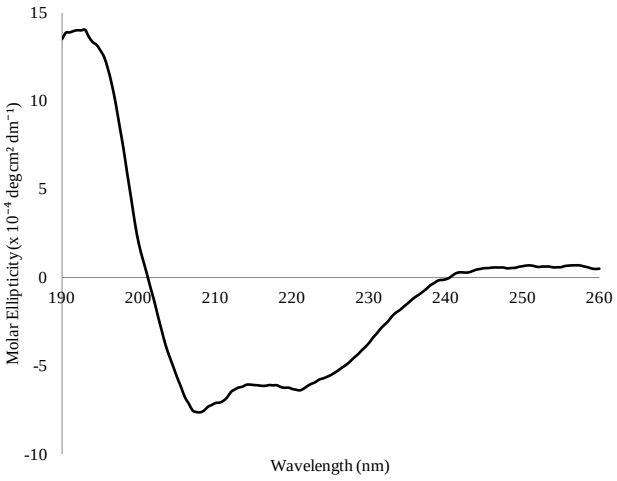


Figure 2.9: Far-UV spectrum of recombinant zinc loaded SmtB

Although no conclusions could be drawn regarding the stoichiometry or tertiary structure of SmtB, the secondary structural information suggested that the protein had not been misfolded during expression and freeze drying had not impaired the secondary structure to a significant extent.

2.3.5 ICP-OES and absorbance at A_{280}

A number of methods were used to calculate the concentration of the protein within the samples produced. The A_{280} assay used in other literature concerning SmtB, relies heavily on the extinction coefficient. The extinction coefficient, at 280 nm, for SmtB was calculated as 5960 M^{-1} cm^{-1} (Kar et al. 1997; VanZile et al. 2000). Prediction of a proteins extinction coefficient is relatively reliable but becomes less so when it contains no tryptophan residues (Pace et al. 1995), as in the case of SmtB.

To investigate whether the estimated extinction coefficient for SmtB was correct, a comparison of the concentration obtained by ICP-OES to the concentration obtained from an A_{280} assay was made, Table 2.1.

Table 2.1: Comparison of two different techniques for the determination of protein concentration

	SmtB concentration determined by each method (μM)	
Sample	A_{280} assay	ICP-OES
1	37.47	29.75
2	37.56	30.27
3	53.70	48.45

The three samples used to compare the two techniques, were in each case estimated to have a higher concentration when the A_{280} assay was used in comparison to ICP-OES. This indicated that the estimated extinction coefficient used may be too small.

From the concentrations determined by ICP-OES the extinction coefficient for the three zinc bound SmtB samples were calculated as 7496 M^{-1} cm^{-1}, 7400 M^{-1} cm^{-1} and 6605 M^{-1} cm^{-1}. These values were 10 – 20 % higher than the predicted value. The extinction coefficient was likely to be inaccurate due to the lack of tryptophan residues. Pace *et al* estimated and measured the extinction coefficient for 80 proteins; and their findings reported that proteins which contained no Trp could cause estimations, in extinction coefficients, to be inaccurate by more than 10 % and up to 17 %. It was concluded that these inaccuracies were related to the reliance on tyrosine residues, whose molar absorbance can be sensitive to their environment (Pace et al. 1995). This apparent error in the estimated extinction coefficient could

lead to inaccuracies within experiments where a precise protein concentration is required, for example in the determination of the SmtB zinc ion stoichiometry.

2.3.6 The expression of SmtB in the presence of cadmium

To determine the effect of expressing SmtB in the presence of cadmium ions instead of zinc ions, the cadmium sample was purified utilising the methods outlined above.

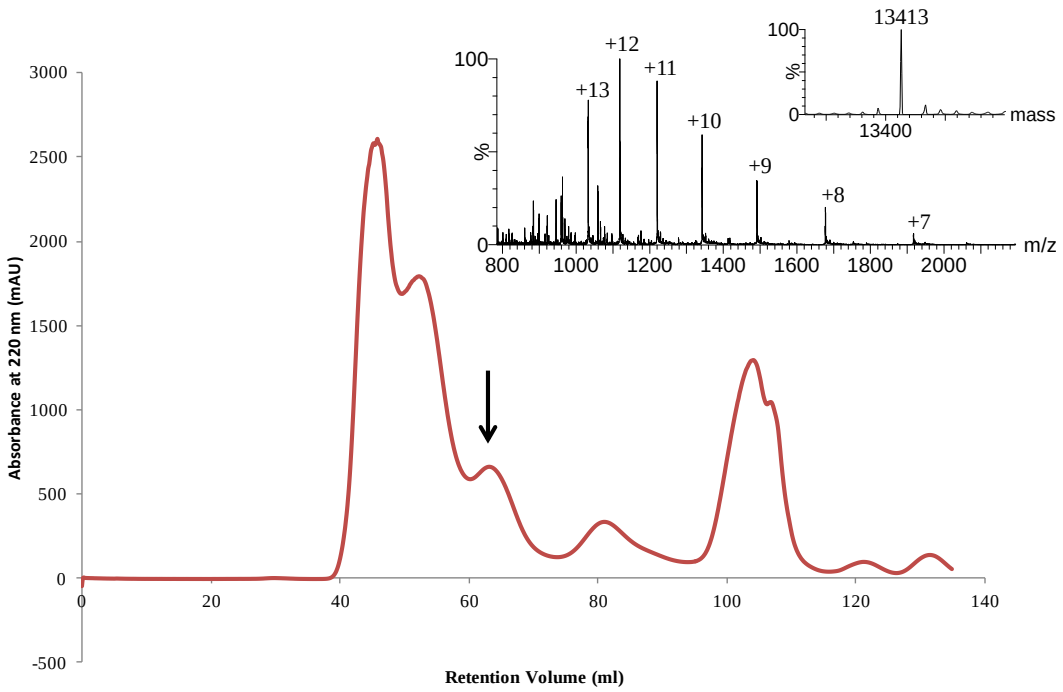


Figure 2.10: The SEC elution profiles from a sample expressed in the presence of cadmium.
The arrow indicates the known elution peak of SmtB when expressed in the presence of zinc. The insert presents the denatured mass spectrum and accompanying deconvolution from a sample obtained from the highlighted peak.

The heparin (data not shown) and SEC, Figure 2.10, chromatograms produced from a SmtB cadmium pellet were comparable to their zinc equivalents. The Zn^{2+} and Cd^{2+} pellets were expressed from the same overnight culture under the same conditions. The only difference was that a lower concentration of cadmium was used since it has been often observed to have a lower minimum inhibitory concentration (MIC) (Brocklehurst and Morby 2000), although the exact MIC may differ between different cell types.

The similarities between chromatograms were expected. The majority of the observed peaks corresponded to partially purified proteins originating from the expression host, *E. coli*, and their production would not rely on the availability of added zinc ions. The SEC chromatogram obtained from the cadmium sample did contain the peak usually associated with SmtB from zinc pellets; however it was low in intensity and partially shifted to an earlier retention time. These results indicated that a lower amount of SmtB protein had been expressed in total, in the presence of Cd, when compared to Zn. To ensure the protein within the fractions was still SmtB, and not another protein, a denatured mass spectrum was obtained to help to identify the protein. This produced a mass spectrum, Figure 2.10, which was comparable to Figure 2.8 and an estimated mass which supported the fact that the product was N-terminally cleaved SmtB. Although this mass spectrum demonstrated that SmtB could be produced in the presence of Cd ions it did not reveal if Cd was bound since any protein metal interaction would be lost in highly inorganic buffer. To accurately compare the amount of SmtB produced the SmtB yield was calculated from both conditions. In the presence of zinc ions 2.79 mg/L of SmtB was produced whereas in the presence of cadmium ions 1.86 mg/L was produced. This suggested that, either SmtB did not bind Cd or that, even with bound Cd the protein was still unstable which led to degradation and therefore sample loss throughout expression and the purification procedure. The metal speciation of this sample was determined. The result will be discussed in Chapter Three.

2.4 Conclusions

The expression and purification procedure outlined above produced comparable yields to other purification methods outlined in the literature. The method produced a product which was not only pure but, due to its exposure to zinc ions in the expression culture, was more stable and less prone to oxidation. This allowed for easier sample preparation and storage, although care still had to be taken to ensure limited exposure to oxygen. The incorporation of a size exclusion column into the purification procedure allowed a mass spectrometry compatible buffer to be used, thus a desalting set was incorporated whilst further sample purification was achieved. The expression and purification of SmtB in the presence of cadmium was shown to produce SmtB but at a lower yield. This suggested that Cd was less able than Zn to stabilise SmtB during purification.

Chapter Three: Conformational changes associated with zinc ion removal

3.1 Introduction

3.1.1 Metal binding to SmtB

Four potential zinc binding sites (2 x α5 and 2 x α3N) have been identified within the sequence of the SmtB dimer, Figure 3.1. There is some debate, however, as to the biological occupancy of these sites.

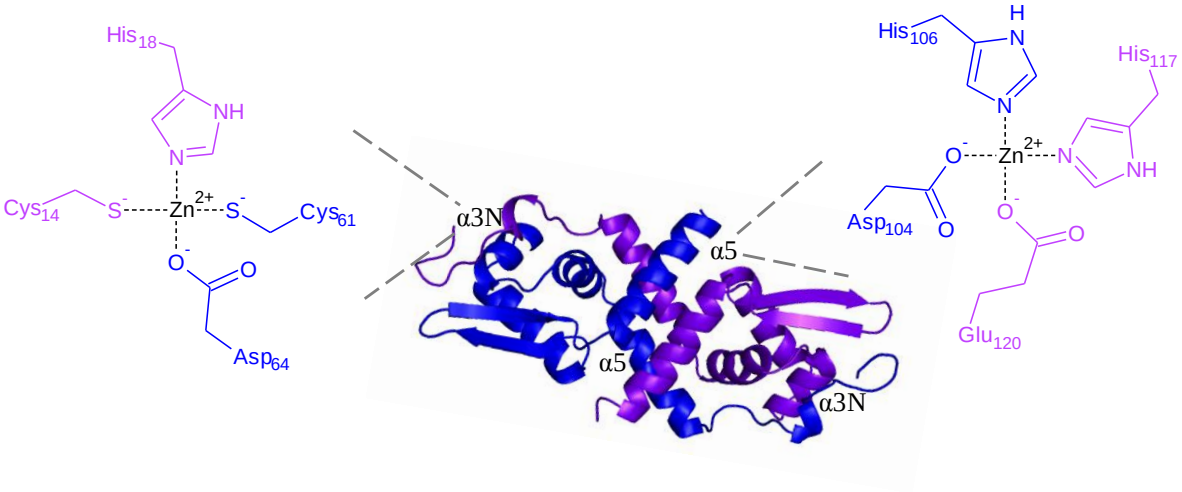


Figure 3.1: A schematic showing the location and coordination environment of the four potential zinc binding sites.
One monomer is displayed in blue whilst the other is purple, residues in the binding sites are coloured to identify which monomer they belong to.

Early work involving the double mutation of the residues T11 and C14 to Ser, (C14 was later found to belong to the α3N site), and the residues H105 and H106 to Arg (H106 was later found to belong to the α5 sites) revealed that the α5 site was highly involved in the zinc sensing capabilities of SmtB. The mutation to the α3N site did not prevent SmtB sensing zinc, although the resulting de-repression of the operon upon zinc exposure was reduced compared to the wildtype. Disruption to the α5 site, silenced sensing i.e. no de-repression was observed in response to zinc (Turner et al. 1996). Sedimentation equilibrium experiments (Kar et al. 1997) later indicated that apo-SmtB existed at equilibrium between monomeric (16%), dimeric (81%) and tetrameric (3%) forms. Upon addition of Zn^{2+} two observations were made; firstly SmtB had the ability to bind four zinc ions per dimer, and secondly this binding event not only improved the stability of the dimeric species but also resulted in a

compaction in the dimeric structure. The results obtained, however, could not identify each class of binding sites separately (Kar et al. 1997).

UV-Vis spectroscopic titrations carried out more recently disagree with the proposed stoichiometry of 4 zinc ions per dimer (VanZile et al. 2000). Competitive titration experiments used the metallochromic indicators mag-fura-2 and pyridyl-azo-resorcinol (PAR) to determine the zinc binding stoichiometry. In brief, a mixture of apo-SmtB and apo mag-fura-2 was produced, to which zinc ions were titrated. After the SmtB: zinc ion ratio reached 1:1; zinc ions bound to mag-fura-2 were indicated by a shift in absorbance maxima. The PAR assay involved titrating apo-SmtB into the PAR_2-Zn complex. As SmtB removed zinc ions from PAR_2-Zn a decrease in the absorbance at 500 nm was observed. This decrease in absorbance continued until the concentration of monomeric SmtB matched the concentration of zinc ions, resulting in the complete dissociation of the PAR_2-Zn complex. These results led to the conclusion that SmtB contained only one high affinity zinc binding site per monomer (VanZile et al. 2000). Interestingly EXAFS data, presented in the same publication, suggested the presence of at least one cysteine thiolate ligand in the coordination environment, of both zinc and cobalt bound SmtB. Coupling this data with the slightly reduced zinc sensing of the T11S/C14S double mutant (Turner et al. 1996) and NMR data (Kosada et al. 1999) the involvement of the C14 thiolate in zinc coordination was suggested. The authors discussed the possibility that the identified high affinity site may be involved in dimeric stabilisation and not regulation; however the absence of a second lower affinity zinc binding site led them to believe that this theory was unlikely. It was concluded that the identified site, containing a mixture of thiolate, imidizole and carboxylate ligands, was indeed the repressor site. The proposed regulatory sites, containing the known zinc sensing residue His106 (Turner et al. 1996) do not contain ligating Cys residues but only His, Asp and Glu residues (Cook et al. 1998).

Later publications (VanZile et al. 2002a; VanZile et al. 2002b) calculated that, in the absence of DNA, the α3N sites had a 20 fold higher affinity to zinc ions than the α5 sites. NMR perturbation spectrometry (VanZile et al. 2002b) also showed that the binding of zinc ions to wildtype SmtB caused large shifts in the amide backbone resonances around the α3N site, although shifts around the α5 site were also

observed and zinc ion binding to this site could not be ruled out. When the α3N sites were disrupted by S-methylation, only shifts in the α5 site were observed. Relating these observations to their earlier work, which displayed the binding of two zinc ions per SmtB dimer and the involvement of a thiolate ligand in the primary coordination sphere, the authors came to the surprising conclusion that metal ion binding to $SmtB_2$ undergoes strong negative cooperativity. This means that when SmtB is free in solution the α3N sites are not preferentially occupied but solely occupied. As the α5 sites, and not the α3N sites, are vital for zinc ion sensing the lower affinity α5 sites must be filled in the SmtB:DNA complex for the sensor to function. The observation that Co, which bound exclusively to the α5 sites within an oxidised form of SmtB, could partition between the α5 and α3N sites upon reduction of the disulphide bonds led to the proposal of an intermolecular migration model shown in Figure 3.2.

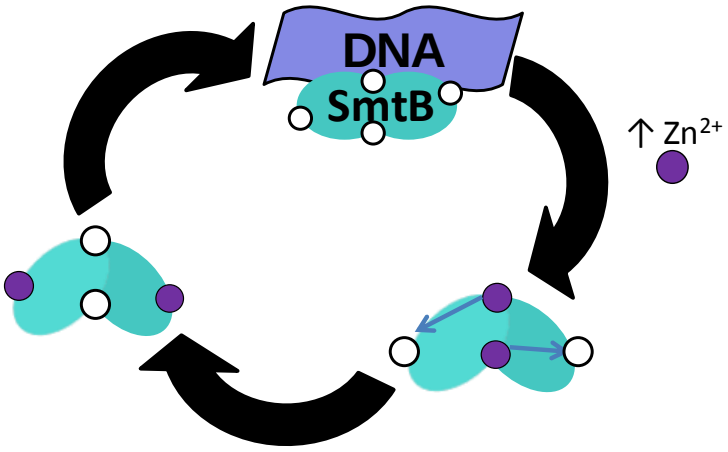


Figure 3.2: Proposed mechanism of zinc ion binding and transfer within SmtB. White circles correspond to empty binding sites whereas purple circles represent zinc ions. Figure adapted from (VanZile et al. 2002a).

In this model apo-SmtB is proposed to be bound to the operator/promoter sequence, causing repression of SmtA transcription. In the event of a cytosolic rise in zinc ion concentration, zinc ions initially bind to the α5 sites, permitting the conformational change required for de-repression. These zinc ions then migrate from the α5 to the α3N sites upon SmtB's dissociation from DNA, Figure 3.2. Unlike Co, Zn ions are not thought to equilibrate between the α3N and the α5 site due to the substantial difference in their zinc binding affinities. Instead Zn is thought to solely occupy the

α3N sites (VanZile et al. 2002b). The previously published EXAFS data suggested that Co and Zn had identical coordination spheres and this is inconsistent with that result.

The retention of the α3N site, within SmtB, although it appears to have no active role in Zn^{2+} sensing, led Vanzile *et al*. to suggest that its conservation could be due to Cd^{2+} sensing. This theory is supported by the evidence that exposure of the *smt* operon to Cd^{2+} leads to *smtA* transcription (Huckle et al. 1993) suggesting Cd^{2+} is also recognised by SmtB and an associated conformational change occurs.

3.1.2 SmtB conformational changes

The binding of metal ions to metalloregulatory proteins can often lead to a conformational change, which either increases or decreases the ability of the protein to bind to specific operator/promoter site. In the case of SmtB, zinc binding has been shown to cause compaction of the dimeric structure (Kar et al. 1997) which is thought to lead to de-repression of itself and the *smtA* gene.

Recent X-ray crystallographic studies have not only confirmed the presence of this compaction but have also provided detail on how zinc ions binding to the α5 site may promote a conformational change (Eicken et al. 2003). Crystal structures have been obtained for the apo, one and two zinc ion bound SmtB dimers. To produce these structures wildtype SmtB and a triple Cys to Ser mutant were expressed and purified in their apo forms. To the mutant, containing the disrupted α3N sites, 1 mol equiv. of Zn^{2+} per monomer of SmtB was added, forming the $SmtB_2Zn_2$ species with zinc ions bound in both α5 sites. The $SmtB_2Zn$ species with only one α5 site filled was created by the addition of zinc to the wildtype protein. It is important to note that the binding of a zinc ion to one α5 site in the $SmtB_2Zn$ structure is in itself extremely interesting (Eicken et al. 2003). Excluding experimental error, such as protein oxidation disrupting the α3N site, there is no explanation as to why a single zinc ion should be bound to a α5 site, instead of to one of the α3N sites, which are proposed to have a much stronger zinc binding affinity.

The superposition of these structures revealed a structural change between the apo dimer and the $SmtB_2Zn_2$ structure, whereas the change upon binding a single zinc

ion was not as pronounced, Figure 3.3. The conformational change, caused by both α5 sites being occupied, involves the movement of the helix-turn-helix (α3-turn-αR). This is predicted to be involved in DNA binding (Huckle et al. 1993; Cook et al. 1998), and is found to move towards the opposing promoter by as much as 4.8 Å (Eicken et al. 2003).

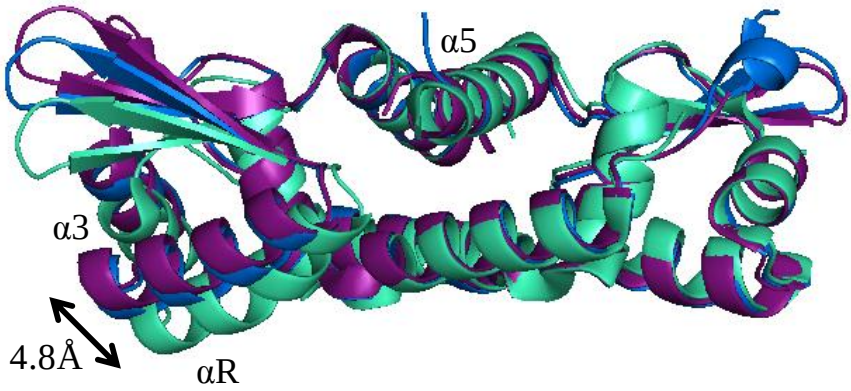


Figure 3.3: Overlaid crystal structures of SmtB dimers, with varying zinc ion stoichiometries.
The α5 helices were superimposed allowing the change in conformation, indicated with an arrow, between crystal structures, apo (blue) Zn (pink) and Zn_2 (green), to be visualised. The Figure was adapted from (Eicken et al. 2003)

In the $SmtB_2Zn_2$ structure a hydrogen bond network is thought to be set up between the α5 binding site, and the αR DNA recognition helix. The network stretches from the non-ligating $N^{\varepsilon 2}$ of H117 across the subunit interface to Arg87'. The nearby amide of Leu88' then forms a hydrogen bond to Leu83', present within αR. This network is absent from both the apo and $SmtB_2Zn$ species, indicating that it may well play a pivotal role in the conformational change associated with the complete occupancy of the α5 sites.

A further conformational change associated with the binding of Zn^{2+} to the α3N site, after the conformational change associated with de-repression has not been ruled out. This is due to the fact that, if intermolecular movement of zinc ions does occur, SmtB must retain a conformation with a reduced affinity for DNA, to prevent rebinding of the repressor (VanZile et al. 2002a).

Conformational changes associated with other SmtB/ArsR members

The study of individual SmtB/ArsR sensors family members can often provide important information which can be applied to the remainder of family members. One such member is the α5 sensor CrzA, which as discussed in Chapter One, is a homologous protein to SmtB. The free structure of CzrA was found to be almost identical in the presence and in the absence of bound zinc ions (Eicken et al. 2003). It was not until the CzrA:DNA complex was studied that the zinc induced conformational change became clear. Recently the binding of zinc ions and subsequent dissociation of the CrzA:DNA complex has been investigated by molecular dynamic simulations (Chakravorty et al. 2012) and were found to be in good agreement with the NMR data. The data suggests that the DNA compatible closed conformation of CzrA is only achieved in the presence of DNA. Binding of zinc ions to apo-CrzA:DNA, orders the metal binding α5 helices and increases flexibility in the proteins β-sheet regions. These changes cause a four fold reduction in the affinity of the protein for its DNA binding sequence finally causing the open conformation to be adopted in the CrzA dissociated state.

Aims

The following work investigates the stoichiometry of SmtB within the gas phase and compares it to solution phase data. SmtB was recombinantly produced from *E. coli* grown in media supplemented with zinc ions; allowing for the study of the protein as produced in the cytosol and not a reconstituted protein. As in the case of SmtB dimers, changes in Zn ion stoichiometry go hand in hand with changes in conformation. The use of a zinc chelator and variation of the pH, to produce a range of SmtB stoichiometric species, coupled with the shape selective technique of ion mobility mass spectrometry has led to the majority of the results and conclusions outlined in this chapter. Further work investigating SmtB cadmium binding and the re-addition of zinc to apo-SmtB, along with attempts to locate the zinc binding sites, are also discussed.

3.2 Materials and Methods

All reagents were obtained from Sigma-Aldrich UK or Fisher Scientific UK unless highlighted within the text. These were of high quality and purity, to reduce the risk of metal ion contamination. Solvents, including water, used during mass spectrometry were of an analytical grade. Any equipment used is presented in the text.

3.2.1 Sample preparation of $SmtB_2Zn_4$ for ESI experiments

A lyophilised SmtB sample was re-suspended in 10 mM ammonium bicarbonate, adjusted to pH 7.4 and concentrated using Millipore centrifugal filter units, 10 kDa cut off. The concentrated sample was washed a total of five times to ensure the removal of excess sodium ions. Each of the washes involved diluting the concentrated sample 10 fold in 10 mM ammonium bicarbonate pH 7.4 and then concentrating it by spinning the centrifugal filter units at 10500 g for approximately 10-15 minutes, depending on the concentration of the sample.

3.2.2 ICP-OES sample preparation

ICP-OES samples were taken directly from the FPLC, with no further clean up, to calculate the presence of free metal ions within SEC samples. To determine the protein: zinc ion stoichiometry of SmtB, FPLC samples were concentrated ~5x using Amicon Ultra-4 10 kDa cut-off centrifuge filter units (Millipore). All ICP-OES samples were then diluted 1 in 5 using 0.1 M nitric acid and data obtained using the procedure outlined in Chapter Two.

3.2.3 Production of ammonium salt of EDTA

A 50 mM stock of $(NH_4)_4$EDTA was produced by dissolving EDTA (free acid) in a 1 M stock of ammonium hydroxide in a 1:4 molar ratio. Once dissolved the sample was made up to contain 10 mM ammonium bicarbonate and the pH was adjusted with 2 M acetic acid to pH 7.4.

3.2.4 EDTA time course

A concentrated sample of purified, in-vivo metallated SmtB, was diluted down with $(NH_4)_4$EDTA so that the final reaction sample contained 0.8 mM SmtB (based on

monomer) and 1.6 mM EDTA. This sample was left to incubate at room temperature and aliquots were taken after 7, 22, 30, 60, 90, 180 minutes. Each aliquot was diluted down to 13 µM in 10 mM ammonium bicarbonate, pH 7.4 and kept on ice before introducing to the mass spectrometer. This process usually took approximately 5-10 minutes but the dilution and cooling on ice should have slowed any further reaction taking place. A control sample of concentrated SmtB, without EDTA, was also left at room temperature for the duration of the experiment and was diluted down to 13 µM before mass spectrometry analysis.

3.2.5 EDTA titration

A concentrated SmtB sample was aliquoted into five different reaction vessels. Each aliquot was diluted down so there was a final concentration of 0.8 mM SmtB (based on monomer) in the reaction vessel using different concentrations of EDTA solution. This resulted in 0, 0.4, 0.8, 1.2 and 1.6 mM EDTA being present in the final reaction mixtures; these were then incubated at room temperature. After four hours an aliquot from each reaction mixture was diluted down so the final SmtB concentration was 20 µM, this was higher than usual as some protein degradation was expected. The samples were then kept on ice before analysis using mass spectrometry.

3.2.6 Removal of zinc ions from SmtB via dialysis

To produce a SmtB sample without zinc an aliquot was concentrated and desalted, up to five times, in ammonium bicarbonate utilising Amicon centrifugal filters, 10 kDa cut-off (Millipore). The resulting sample was then loaded into a pre-soaked dialysis cap and dialysed against 10 mM ammonium bicarbonate, 5 mM EDTA, 2 mM DTT for 24-48 hours. Dialysis took place at 24 °C, in an anaerobic cabinet (Don Whitley Scientific Limited), 80% nitrogen, 10% carbon dioxide and 10% hydrogen. All buffers were made up using nitrogen saturated water. Buffer exchanges were conducted after 2, 5, 20 and 28 hours.

3.2.7 Nitrogen saturated water

To ensure anaerobic conditions within the anaerobic cabinet, nitrogen saturated water was produced. Water was heated until boiling, on a hotplate, to remove oxygen. Nitrogen gas was then pumped through the water, in a fume hood, whilst it was cooled on ice.

3.2.8 Sample preparation for Ion mobility mass spectrometry of apo-SmtB

Apo-SmtB had to be cleaned after dialysis to remove excess bound EDTA. This was conducted by centrifugal filters after further dialysis failed to remove all the bound EDTA. Aliquots of the samples were washed a maximum of three times in 10 mM ammonium bicarbonate pH 7.4 before mass spectrometry and ion mobility mass spectrometry experiments were carried out. A control of zinc bound SmtB was also analysed after its incubation in the anaerobic cabinet.

3.2.9 Re-addition of Zn^{2+} to apo-SmtB

An aliquot of the cleaned apo-SmtB, produced from dialysis against 10 mM ammonium acetate, 5 mM EDTA, 2 mM DTT for 48 hours, was incubated with excess zinc acetate for two hours at room temperature. Ammonium acetate was chosen instead of ammonium bicarbonate to remove the chance of producing insoluble zinc carbonate upon the addition of zinc ions. Zinc acetate was chosen, as the source of zinc ions, instead of zinc sulphate to remove the requirement of a sample cleaning step before mass spectrometry, which may lead to sample degradation. Mass spectrometry conditions were kept constant between apo-SmtB and zinc added samples.

3.2.10 pH titration

A stock solution of 10 mM ammonium bicarbonate was adjusted to a range of pH values, between 8.5 and 3.0, using either acetic acid, to lower the pH, or ammonium hydroxide, to higher the pH. A concentrated sample of SmtB was produced, outlined in the procedure above, and diluted down to 13 µM in the different ammonium bicarbonate solutions. These solutions were then left to equilibrate for 30 minutes at room temperature. After the removal of an aliquot for mass spectrometry analysis the pH of the reaction mixtures was measured by a pH meter to determine their final pH value.

3.2.11 Collisionally induced dissociation

A desalted 13 µM sample of SmtB in 10 mM ammonium bicarbonate pH 7.4 was infused via nanospray needles into the mass spectrometer. The mass spectrometry and ion mobility conditions applied were identical to those outlined below. The cone

voltage was maintained at 60 V, whilst the trap collision energy (CE) was ramped from 4 V to 55 V.

3.2.12 Electron transfer dissociation

Electron transfer dissociation (ETD) experiments were carried out in collaboration with Dr. Julia Smith, on an amaZon speed ETD system (Bruker, Coventry, UK). A $SmtB_2Zn_4$ sample was desalted/cleaned as indicated previously and was diluted, using 10 mM ammonium bicarbonate (pH 7.4) to 2 ρmol/µL. This sample was then infused, from a syringe drive, into the system at 3 µL/min. Intact masses were determined using UltraScan mode and MaxEnt deconvolution. The isolation of both monomeric and dimeric SmtB with differing numbers of zinc ions bound was conducted prior to ETD fragmentation which was performed in conjunction with proton transfer reaction (PTR) experiments. Flouranthene radical anions and Flouranthene anions, which are generated in a negative chemical ionization source, were used as the ETD and PTR reagents respectively. ETD ion charge control (ICC) was 250000 with an 8 ms reaction time, whilst PTR ICC was 150000 with a 20 ms reaction time. Compass Data Analysis 4.0 was used in conjunction with BioTools 3.2 and Sequence Editor 3.2 to analyse spectra to determine sequence coverage.

3.2.13 Mass spectrometry and ion mobility mass spectrometry analysis

Mass spectrometry experiments were carried out on a Synapt G2 HDMS system (Waters, Milford, USA) utilising a nanoflow electrospray source in conjunction with fused silica nanospray needles. Instrument parameters were optimised to ensure adequate transmission whilst maintaining near native conditions. All experiments were carried out with 1.2-1.4 kV applied to the capillary, a source temperature of 80 °C, whilst a cone voltage of either 40 or 60 V was applied to reduce any adduct formation. The source backing pressure was optimised at 2.6 mbar to provide a sufficient amount of collisional cooling for the higher molecular weight ions. To ensure the acquisition of accurate masses a calibration using 2 mg/mL CsI, dissolved in 50% isopropanol, was conducted in positive ion, resolution mode, scanning between 500-3500 *m/z*. Ion mobility parameters were optimised to provide the maximum separation within the ion mobility cell. For SmtB these values were; wave height 40 V, IMS gas 90 mL/min, wave velocity 800 m/s. The ion mobility bias was also optimised to a value of 45 V to ensure the sufficient transmission of higher

molecular weight ions through the ion mobility cell. MassLynx software (Waters, Milford, USA) was used for the acquisition of data and was used in conjunction with Driftscope software for processing data.

3.2.14 Ion mobility calibration

The measured arrival time of an ion, in a travelling wave ion mobility device, does not have a direct mathematical relationship with the collisional cross section of that ion. A calibration has therefore to be performed to calculate the collisional cross section. This procedure has been outlined previously (Scarff et al. 2008; Thalassinos et al. 2008). Briefly the procedure involves acquiring data from a calibrant protein with a known collisional cross section and mass, under the same experimental conditions as the sample of interest. Here the calibrant sperm whale myoglobin was chosen, as its mobility space bracketed the mobility space of the SmtB protein. The arrival times of the sperm whale myoglobin were corrected for *m/z* dependent flight, whilst the published CCS were corrected for reduced mass and charge. A plot of these corrected values produced a power fit and the constants within this fit were used to estimate the CCS of SmtB.

3.2.15 MOBCAL

Theoretical CCS's of SmtB were produced from the X-ray crystal structures published in the RCSB protein data bank (PDB). These include the apo-SmtB dimer, $SmtB_2Zn_1$ and $SmtB_2Zn_2$. These allowed comparisons to be drawn from the experimentally deduced estimated CCS. Prior to the X-ray structures being entered into MOBCAL, hydrogen atoms were added to ensure a true representation of the species. This was performed in Pymol. Two different algorithms were implemented, Projection Approximation (PA) and Exact Hard Sphere Scattering (EHSS). Cross sections produced from the simplest and highly computationally efficient PA method were created by projecting the ion onto a perpendicular plane. The orientation of the ion was then changed and the process repeated. The final result is then the average of all of the projections. This model, although suitable for small ions, does not take into account the scattering process (Bleiholder et al. 2011) and therefore under-estimates the CCS for concave surfaces (Shvartsburg and Jarrold 1996) or proteins larger than 2 kDa (Jurneczko and Barran 2011). The EHSS method is slightly more computationally intensive as it considers collisions of the ion with the buffer gas but

does not take long range potential into account. This method produces more accurate CCS's for larger proteins, in comparison with the PA method, and is often within reasonable agreement with the trajectory method (TM). The TM model takes into account long range potentials as well as collisions between ions and the buffer gas. This greatly increases the computational requirement and therefore in this instance was not implemented. A new method has recently been published, the projected superstition approximation (PSA) (Bleiholder et al. 2011). This algorithm is not as computationally intense as TM but provides consistent results which are highly comparable even in the study of complex protein structures. This algorithm could therefore be used in place of the EHSS method in any future work.

3.2.16 Molecular modelling of $SmtB_2Zn_4$

The molecular modelling of $SmtB_2Zn_4$ was carried out in conjunction with Dr. Claudia Blindauer, and is based on pdb file 1r23.

To create an SmtB structural model with two occupied αN3 sites, first the N-terminus of the second monomer (B), in which residues 18-24 were not resolved, was extended, making use of coordinates derived from monomer A. The homology modelling program MODELLER (v. 9.7) (Marti-Renom et al. 2000) was then employed, to extend both N-termini up to Val13. The 10 models generated were inspected manually, and the mode that showed the two extensions in the most frequently observed conformation was selected. Side-chain conformations, with the exception of those of Asp104, His106, His117, and Glu120 (the α5 ligands) were optimised using SCWRL (v. 3.0) (Canutescu et al. 2003). The model was then imported into the program MOE (v. 2.0), hydrogens were added, and four Zn ions were introduced into the four sites. Each site was then energy-minimised separately, using a customised version of the AMBER-94 force-field, in which Zn-specific parameters were incorporated as well as the restraint on omega-angles being increased. This had been found to be unsatisfactory in the original version. After energy-minimisation of the position of the immediate ligands, the adjacent environment was optimised, and finally, the entire molecule was subjected to energy minimisation. All minimisations were terminated based on the steepness of the RMS gradient (<0.05). The final model was then submitted to the WHATIF web interface (WHATIF), ensuring that a physically reasonable model had been produced.

3.3 Results and discussion

3.3.1 ESI mass spectrometry of recombinant SmtB

To investigate the equilibrium of in-vivo metallated SmtB between monomer, dimer and tetramer in the gas phase, along with the zinc binding stoichiometry of these species, nano-electrospray ionisation mass spectrometry was carried out under near native conditions, Figure 3.4. Although the pI of apo-SmtB was calculated as 6.54, meaning it should carry a negative charge at pH 7.4, no negatively charged ions were observable and therefore positive ion ESI was used.

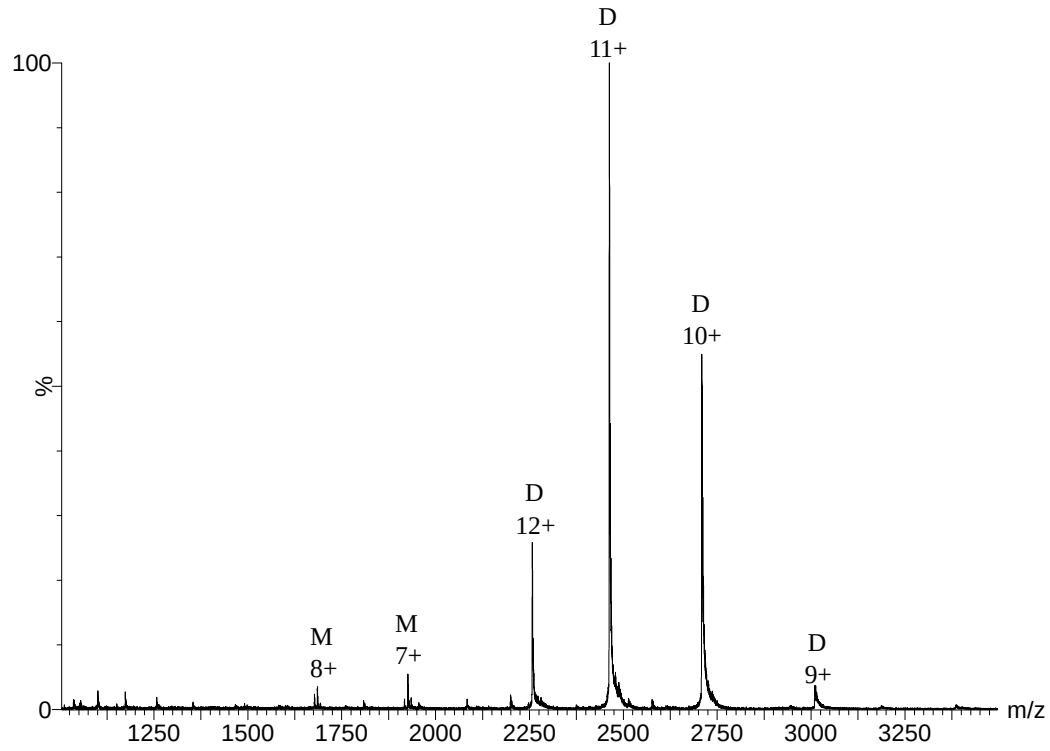


Figure 3.4: ESI mass spectrum of 13 µM SmtB in 10 mM ammonium bicarbonate pH 7.4.
Monomeric and dimeric peaks are labelled with M and D respectively, along with their associated charge state.

In the presence of zinc ions, solution phase sedimentation equilibrium analysis has determined the equilibrium, between zinc-loaded SmtB monomer and dimer, to be mainly dimeric (97%) with a small percentage of monomer (Kar et al. 1997). The results obtained in the gas phase, Figure 3.4, agree reasonably well with these solution phase results, although the ionisation efficiencies of monomeric and dimeric

species remain unknown. Two charge state distributions can be seen, one at a lower intensity when compared to the other. The *m/z* peaks in the low intensity charge state distribution correspond to the +7 and +8 monomeric species whereas the high intensity distribution contains *m/z* peaks which correspond to +9, +10, +11 and +12 dimeric SmtB species, Figure 3.4. SmtB samples were also studied under higher buffer strengths, up to 200 mM ammonium bicarbonate, to determine if these conditions would stablilise the dimeric species. No obvious changes in equilibrium were observed.

Upon deconvolution onto a true mass scale, utilizing the maximum enthalpy algorithm MaxEnt1, one main dimeric species with a neutral mass of 27,080 Da was identified, Figure 3.5 ii and Table 3.1. The subsequent lower intensity peaks observed within the deconvoluted spectrum were calculated to correspond to sodium adducts. With the knowledge that SmtB polypeptides can undergo N-methionine cleavage, the average mass of the apo-SmtB dimer was calculated as 26,826.6 Da. The difference of 253.4 Da, between the apo-species and the isolated species, was attributed to the binding of zinc ions to SmtB during protein expression in the bacterial cytosol. The observed mass corresponded well with the expected mass, 27080.2 Da, of the SmtB dimer with four zinc ions bound, taking into account the release of 8 protons to balance the charge, Table 3.1.

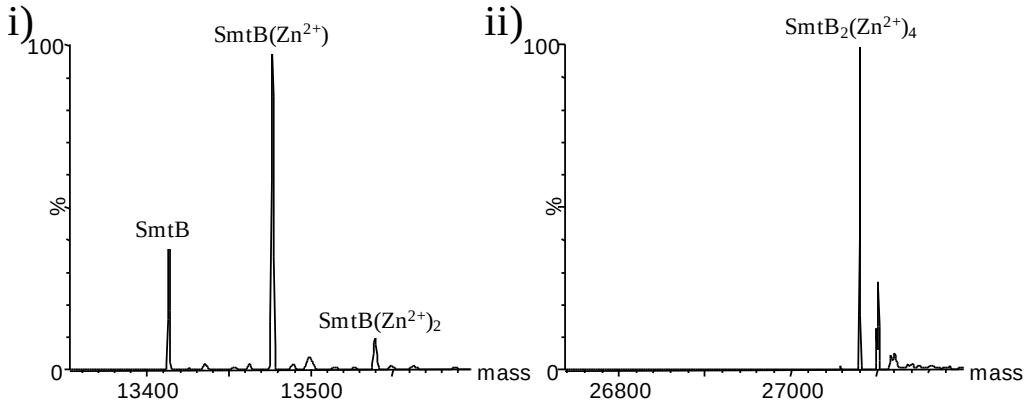


Figure 3.5: Deconvoluted spectra of i) monomeric and ii) dimeric SmtB species revealing the different zinc ion: protein stoichiometries.

The deconvolution of the low intensity charge state revealed that there were at least three different monomeric species present. The observed masses of 13413 Da, 13476

Da and 13540 Da agree with the expected masses of monomeric SmtB bound to 0, 1 and two zinc ions respectively, along with the loss of 2 protons per zinc ion, Table 3.1. Taking into account the determined K_d value for dimerisation of zinc loaded SmtB (80 nM) (Kar et al. 1997) and the solution concentrations used, the equilibrium between monomer and dimer should be 92 % dimeric and 8 % monomeric.

Table 3.1: The observed masses, along with corresponding expected masses, of the prominent species in the near native mass spectrum.
The mass differences of all observed species were within ± 0.2 Da of the predicted masses of SmtB.

Monomeric SmtB	Expected mass (Da)	Observed mass (Da)	Mass difference
Zn_2	13540.1	13540.0	-0.1
Zn	13476.7	13476.5	-0.2
Apo	13413.3	13413.5	0.2

Dimeric SmtB	Expected mass (Da)	Observed mass (Da)	Mass difference
Zn_4	27080.2	27080.0	-0.2

The percentage of monomeric forms observed varied between purified samples from as low as 1 % to as high as 15 %. This variation in monomeric abundance may be a reflection of biological variation caused by zinc availability in expression cultures. In addition any dimeric zinc loss, during purification and clean-up steps, may also contribute to monomer formation. A small amount of mass spectrometry-based dissociation of the dimer cannot be ruled out completely, although attempts were taken to keep this to a minimum, whilst preventing unspecific interactions from occurring. If the monomer observed in the mass spectrum is a reflection of the species present in solution, the incomplete binding of zinc ions is unsurprising upon closer inspection of the proposed binding sites. Both the α5 and the 3αN sites lie on the interface between the two protomers within the dimer and therefore the zinc ions are ligated by residues from both polypeptide chains. If only one of these chains is present the interaction between the zinc ion and the half binding site is likely to be relatively weak, leading to zinc ion dissociation.

3.3.2 Inductively coupled plasma optical emission spectroscopy (ICP-OES)

ICP-OES was carried out to determine the concentrations of zinc, sulphur and copper within purified SmtB fractions. Expression lysates were first subjected to heparin

affinity chromatography followed by SEC to ensure sufficient separation of the lysate proteins. The collection of purified samples from SEC fractions then allowed comparisons to be drawn between the gas phase and the solution phase stoichiometry of SmtB; when the protein was expressed in the presence of zinc ions, Figure 3.6.

The results from the ICP-OES revealed that sulphur was detected, and therefore probably proteins were present, in all but fraction six collected from the size exclusion column. The concentration of sulphur agreed well with the absorbance at 220 nm from peptide bonds, displaying relatively high concentrations when the absorbance peaked and low concentrations in the fractions with low 220 nm readings.

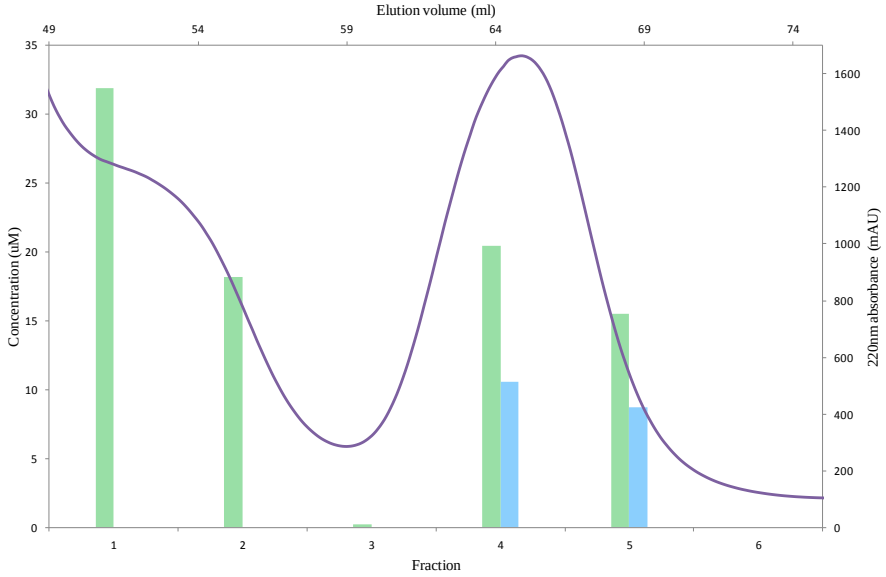


Figure 3.6: An average of two gel filtration chromatograms at 220 nm, purple line, overlaid with the concentrations of sulphur and zinc, green and blue bars respectively, within the displayed fraction.

Zinc ions were only detected from fractions four and five, whilst no significant levels of copper were observed in any SEC fraction, Figure 3.6. The presence of zinc ions in these two fractions could indicate the presence of a zinc binding protein, as free zinc ions would posses a later elution time due to their small dimensions. Comparing the zinc containing fractions with an SDS-PAGE gel revealed that SmtB was found almost exclusively within the zinc containing fractions. These results suggest that the zinc ions found in fraction four and five are exclusively bound to the SmtB protein.

As the concentration of sulphur and zinc within protein samples could be determined simultaneously, it was possible to estimate the number of zinc ions bound to SmtB from ICP-OES data. An average stoichiometry of 1.6 ± 0.2 zinc ions per three sulphur ions was calculated. This corresponded to approximately 3.2 zinc ions per SmtB dimer and clearly indicated that SmtB dimer had the capacity to bind more than two zinc ions in the solution phase as well as in the gas phase. The calculated ratio was not exactly four zinc ions per dimer; however, this may be explained due to the fact that ICP-OES can only determine the average stoichiometry of a sample. The presence of any monomeric species resulting from the dynamic equilibrium of the SmtB protein, which has been shown to bind between zero and two zinc ions, or low concentration protein impurities, would reduce the observed ratio. When this fact is taken into consideration the sulphur to zinc ion ratio in mass spectra including Figure 3.4 was determined. This gave a ratio of 1.9 zinc ions per 3 sulphurs, or 3.8 zinc ions per dimer. The ratio determined from the averaged MS results does not, however, take into account any low molecular weight (< 10 kDa) protein fragments, including any degraded monomeric species, and contaminants. This is because the extensive sample clean up carried out before mass spectrometry would remove the majority of low molecular weight fragments and any remaining fragments below 500 *m/z* would not be observed.

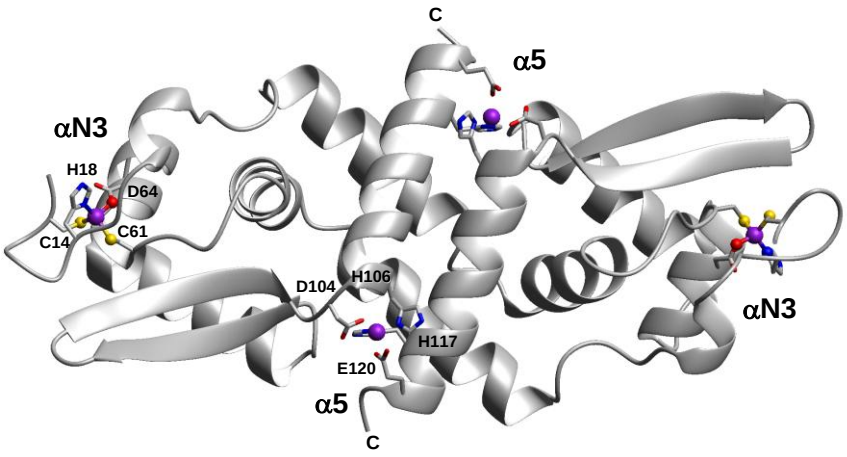


Figure 3.7: A model presenting the proposed occupation of the SmtB dimer with four zinc ions (purple). Ligating oxygen (red), nitrogen (blue) and sulphur (yellow) are highlighted. The model is based on pdb 1r23.

The combination of the mass spectrometry and ICP-OES results reveal a predominant dimeric SmtB species with four zinc ions bound. The absence of zinc ions from all sample buffers and the implementation of a relatively high cone voltage suggests that all the binding observed was specific and could not be contributed to non-specific interactions formed during the electrospray ionisation process.

The results obtained lead to the conclusion that, in the absence of DNA, both sets of metal binding sites within SmtB, α5 and α3N, can be simultaneously occupied by zinc ions, Figure 3.7, as previously observed in the SmtB/ArsR family members CadC (Busenlehner et al. 2002) and ZiaR (Thelwell et al. 1998). This observation disagrees with the negative cooperativity of zinc ion binding that has been proposed in the literature (VanZile et al. 2002b). It is however consistent with previously published, but largely overlooked, data (VanZile et al. 2000; VanZile et al. 2002b). The Zn K-edge EXAFS data (VanZile et al. 2000), which proposed a best fit for a sulphur(imidizole)$_2$carboxylate coordination environment, is likely to correspond to the combined average coordination environment of the α5 and α3N sites and not the coordination environment of a single site. This conclusion could also explains why apo-SmtB, substituted with 2 molar equivalents of Co, can have a mixture of the two sites occupied (VanZile et al. 2002b), but posses the same coordination environment as Zn-bound SmtB. Not only do the results agree with published EXAFS data but they also correlate with the presence of NMR shifts in the α5 site as well as the α3N sites upon zinc binding to wildtype SmtB (VanZile et al. 2002b). The ability of SmtB to bind zinc ions in all four sites simplifies the currently proposed zinc binding model appreciably, allowing both the higher affinity non-sensory α3N sites to be filled as well as the low affinity α5 sensory sites.

3.3.3 ESI-MS of recombinant SmtB expressed in the presence of Cd^{2+}

The conservation of the α3N sites within the SmtB structure may not be vital for zinc sensing, but has led to speculation that this site may be involved in cadmium sensing (VanZile et al. 2002b). Members of the ArsR/SmtB family containing sensory α3N (CadC), or the related α3 sites (ArsR), often sense non-essential heavy metal ions using this thiolate rich site. To investigate the theory that SmtB has retained the α3N site for this purpose SmtB was expressed in the presence of cadmium (Chapter Two). Purified samples were subsequently investigated by using native mass spectrometry.

Since this technique can preserve metal protein interactions in the gas phase, it may reveal if cadmium ions are bound to SmtB and, if so, how many.

Although dimeric SmtB contains four zinc binding sites the chemical nature of the nitrogen and oxygen ligands in the α5 site are not normally associated with the sensing of Cd. The inclusion of soft ligands, in this instance sulphur from two cysteine residues, in the α3N sites however, render these sites more likely to bind Cd (Helbig et al. 2008). When the chemical composition of the binding sites is considered, the observed species could correspond to either $SmtB_2Cd_2$ or $SmtB_2Cd_2Zn_2$. The mass spectrum, Figure 3.8, was at a high enough resolution to distinguish between apo-SmtB with four bound zinc ions (+253.48 Da), two bound cadmium ions (+220.8 Da) and the combination of two zinc and two cadmium ions (+347.54 Da). In comparison to the zinc loaded +11 charge state a difference of approximately -3 *m/z* and 8.6 *m/z* would be expected for the cadmium species or the mixed zinc/cadmium species respectively. In this instance there is no evidence in either the raw or deconvoluted data to suggest the binding of any number of Cd ions to SmtB.

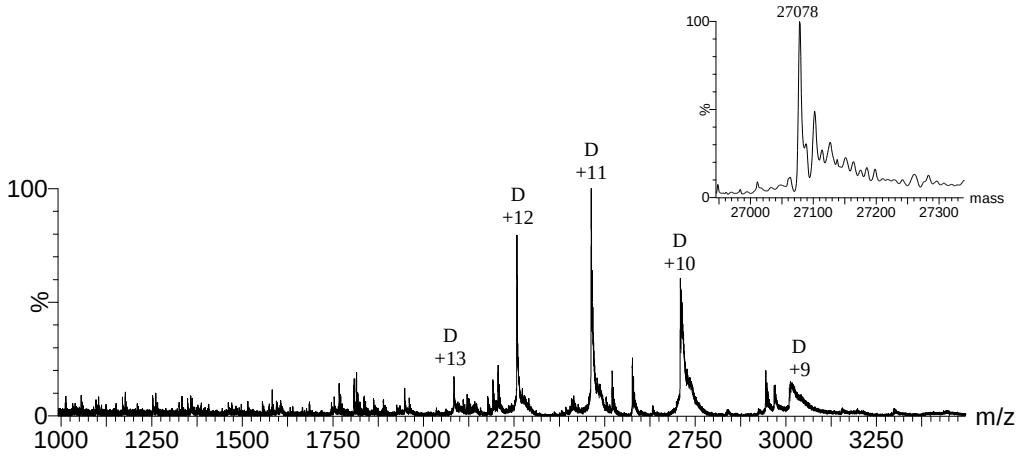


Figure 3.8: ESI mass spectrum of 13 µM SmtB, expressed in the presence of Cd^{2+} in 10 mM ammonium bicarbonate pH 7.4. Insert shows the main species present upon deconvolution with adjacent peaks corresponding to sodiated species.

The spectrum was deconvoluted and produced an abundant peak at 27078 Da. This suggested that four zinc ions were bound to dimeric SmtB, as had been seen previously. The mass spectrum and deconvoluted data contained additional peaks which could correspond to a contaminating protein and possible lower molecular

weight protein fragments. The comparatively low purity of this sample is likely to be due to the lower expression yield, produced from this round of expression, leading to poorly resolved peaks in SEC. The potential protein fragments observed may correspond to degraded proteins.

Since the mass spectrum produced after two rounds of purification showed no cadmium bound species this suggests that either these species are unstable or are not produced. The expression yields achieved from cultures grown in the presence of cadmium were comparatively low compared to SmtB expressed in the presence of zinc ions. One possible explanation is that the majority of expressed SmtB remained as apo-SmtB, whilst a percentage may be stabilised by available zinc ions. These Zn^{2+} are present in the growth media and are taken up and concentrated by *E. coli* at a high efficiency (Outten et al. 2001). The low yield and the absence of SmtB species at lower zinc ion stoichiometries emphasises the prominence of the $SmtB_2Zn_4$ species and could suggest the absence of zinc results in an unstable protein, prone to oxidation and degradation. The absence of cadmium bound SmtB from expression does not rule out the existence of such species, but may highlight the potential of instability of such a species. Attempts could be made to produce a cadmium species from pure apo-SmtB since this could be carried out under anaerobic conditions. This may help to stabilise the species.

3.3.4 Removal of zinc ions from SmtB

In order to investigate the dynamics and binding events associated with the SmtB protein, zinc ions were removed to produce species with a lower zinc binding stoichiometry. To remove zinc ions from the SmtB dimer different methodologies were carried out, these included a pH titration, an EDTA time course, an EDTA titration, dialysis against EDTA, and collisionally induced dissociation (CID).

3.3.4.1 pH titration

Exposure of proteins to changes in pH is a suitable method to investigate protein stability and zinc binding affinities. The high concentration of protons found in a solution of low pH can unfold proteins by disrupting important hydrogen bonds and electrostatic interactions, by protonating residues and competing for interactions. Protons can also compete with metal ions for binding sites, leading to metal ion

dissociation at low pH. If two metal binding sites have different metal binding affinity, as is thought to be the case in SmtB, then different concentrations of protons would be needed to cause metal ion dissociation.

A pH titration has been conducted on SmtB samples in order to investigate the effect on increasing proton concentration, Figure 3.9. The results indicate that a drop from pH 7.6 to 4.9 had no major effect upon the conformation of SmtB. The same charge states, creating a tight charge state envelope reveal a tightly folded protein, with all four zinc ions remaining bound. At pH 5.8 the charge state envelope appears very slightly narrower than at pH 7.6, 6.7 and 4.9, Figure 3.9. This indicates the dimer is perhaps slightly more stable and compact at this pH. In the absence of zinc ions the monomer-dimer association constant was determined to be 100 fold higher at pH 6.0 in comparison to pH 7.4. This increase was suggested to be due to protonation of His residues strengthening non-covalent hydrogen bonds and ionic interactions between the monomers (Kar et al. 1997). This change in stability has not been investigated in the presence of zinc but experimental evidence indicates that a small stabilising effect might still be present, in the gas phase.

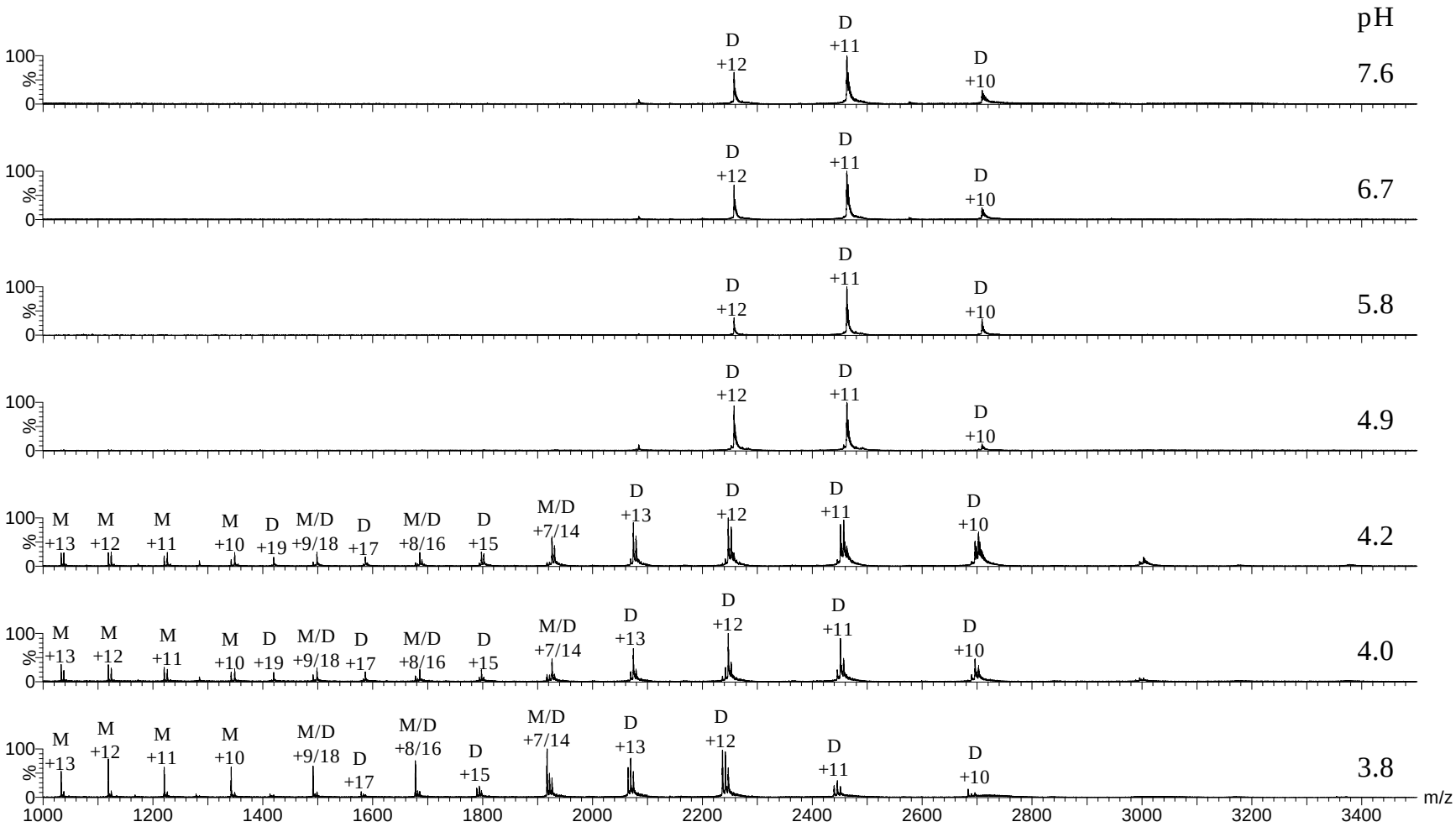


Figure 3.9: A pH titration of SmtB in 10 mM ammonium bicarbonate adjusted to the displayed pH by concentrated acetic acid. Monomeric (M) and dimeric (D) charge states are labelled.

Changing the pH from 4.9 to 4.2 resulted in a significant change in the observed charge state envelope. Peaks corresponding to highly charged dimeric and monomeric SmtB can be seen alongside the folded dimeric SmtB species, which were originally present in the sample. The appearance of these species suggests dimeric unfolding has started to occur, in conjunction with dimeric dissociation. In samples conducted at higher energies, higher abundances of monomeric species are observed (data not shown), which are likely to be produced from the dissociation of unfolded dimer. The abundance of monomeric forms observed within the spectra at pH 4.0 is very similar to that observed at pH 4.2, whereas further dissociation leads to an increase of monomeric apo-SmtB at pH 3.8.

The metal content of the $[M+11]^{11+}$ species, the only solely intense dimeric peak, was observed throughout two different pH titration experiments conducted a few months apart. A reduction in the pH, to pH 5 resulted in the $[M+11]^{11+}$ charge state of dimeric SmtB remaining almost 100 % occupied by the four zinc ions, Figure 3.10 and Figure 3.11.

A reduction in pH from pH 4.9 to pH 4.2 resulted in the loss of up to two zinc ions, Figure 3.10, and a further reduction to pH 3.8 promoted the loss of the zinc ions from the remaining binding sites. The slight build up of the $SmtB_2Zn_2$ species observed at approximately pH 4.0 indicates that there is a difference in affinity between the two sets of binding sites, in agreement with the literature. The loss of the first two zinc ions therefore is suggested to correspond to zinc ion removal from the sensory α5 sites whereas the removal of the final zinc ions is from the α3N sites. Further experiments would be needed to confirm this hypothesis.

Zinc ion depletion was followed from the +11 monomeric charge state, Figure 3.10. At pH 4.9 apo-SmtB monomer and SmtBZn were the only species observed at low abundance. The abundance of both species increased greatly upon the reduction of the pH of the solvent to pH 4.2. As the pH continued to be reduced the relative abundance of apo-SmtB monomer increased, so that at pH 3.8 it was the predominant monomeric SmtB species. These results show that a lower pH can remove zinc ions from both monomeric and dimeric species, although the percentage

of monomer which is formed from a loss of interaction between the dimeric protomers is also likely to be significant.

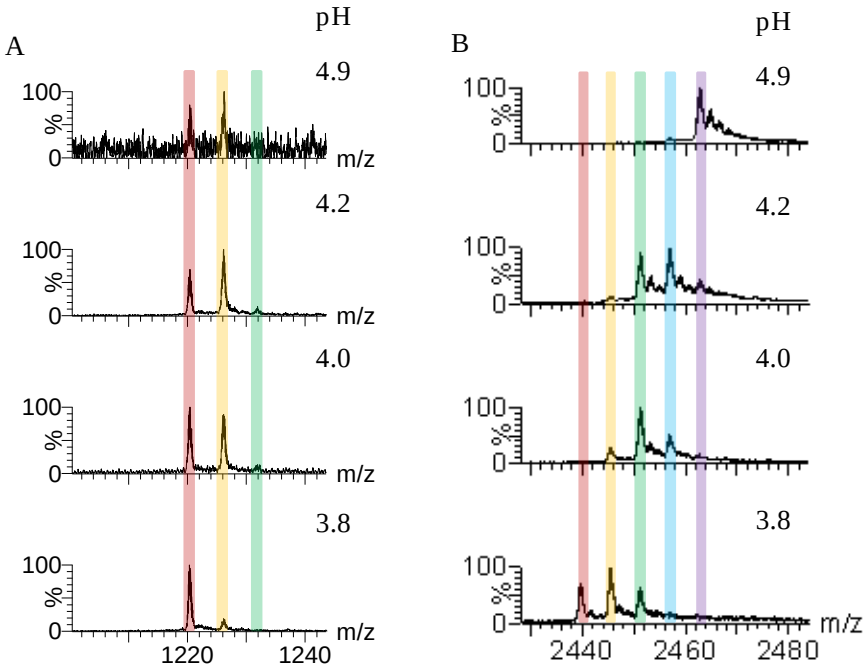


Figure 3.10: A close up of the monomeric (A) dimeric (B) SmtB $[M+11H]^{11+}$ charge state during the latter end of the pH titration.
The loss of zinc ions is observed and different zinc bound species are colour coded, Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red =Apo.

The pH of half dissociation i.e. the pH at which only a half of proteins metal binding sites are occupied, is often used as a measure of the average metal binding affinities for metalloproteins, especially metallothioneins (Shi et al. 1992). In order to investigate the pH of half dissociation, the +11 dimeric SmtB charge state was chosen. At each pH the relative intensity of each +11 species, with four to no zinc ions bound, was measured. These intensities were then used, along with the number of zinc ions bound, to calculate the average percentage of zinc ions bound to the +11 charge state and were then plotted against the pH of the solution, Figure 3.11.

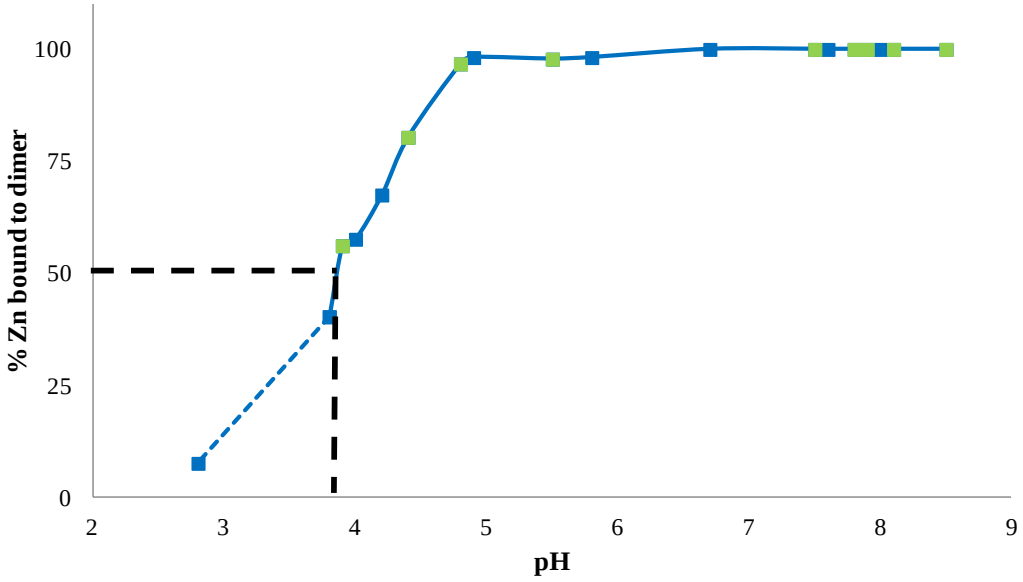


Figure 3.11: Hydrogen ion competition for metal binding sites to SmtB dimer, using $[M+11H]^{11+}$ as an example of the dimeric peak.
Data was gathered from two different pH titrations, shown in green and blue. The pH of half dissociation is also displayed by the thick black dashed line.

Under the experimental conditions employed the pH of half dissociation for dimeric SmtB was found to be between pH 3.85 and 3.90. The titrations were also conducted using a higher cone voltage to investigate whether this would change the percentage of zinc ions bound to the dimer at different pH values. Cone voltage can be an important factor when considering equilibrium constants (Gabelica et al. 2003). The implementation of higher voltages can lead to in-source CID and dissociation of the complex which in turn would reduce its relative intensity. Here all values collected with a cone voltage 60 V were within 3 % of the values obtained using a cone voltage of 40 V. This suggests that the increase in energy was not high enough to remove zinc ions from fully or partially loaded species and that no specific species was degraded. The value for the pH of half dissociation was calculated to be still between pH 3.85 and pH 3.90. This value when compared to pH 4.1 for the GST-SmtA protein (Shi et al. 1992) and pH 4.05 for the cleaved SmtA protein (Daniels et al. 1998) is consistent with the hypothesis that SmtB has a higher average affinity to zinc ions than SmtA. It therefore seems unlikely that SmtA could remove zinc ions from SmtB based on thermodynamics, unless a binding event between the two caused a conformational change which affected the zinc binding sites affinity.

3.3.4.2 EDTA time course

EDTA forms very tight complexes with Zn as shown by its low dissociation constant (10^{-16} M) (Nyborg and Peersen 2004) and was therefore chosen to chelate Zn from SmtB samples. This reaction was followed over time using near native ESI-MS, Figure 3.12. In the absence of EDTA at pH 7.4, SmtB exists in equilibrium between monomer and dimer, where the predominant species was $SmtB_2Zn_4$. Upon incubation at room temperature with 4 molar equivalents of EDTA, the gradual loss of zinc ions was followed by the appearance of a cluster of peaks preceding the fully loaded dimeric peaks, Figure 3.12 and 3.13. As zinc ions were removed by the EDTA, the proportion of dimeric to monomeric species changed, so that throughout the time course there was a notable increase of monomer and an observed broadening of the dimeric charge state envelope, Figure 3.12.

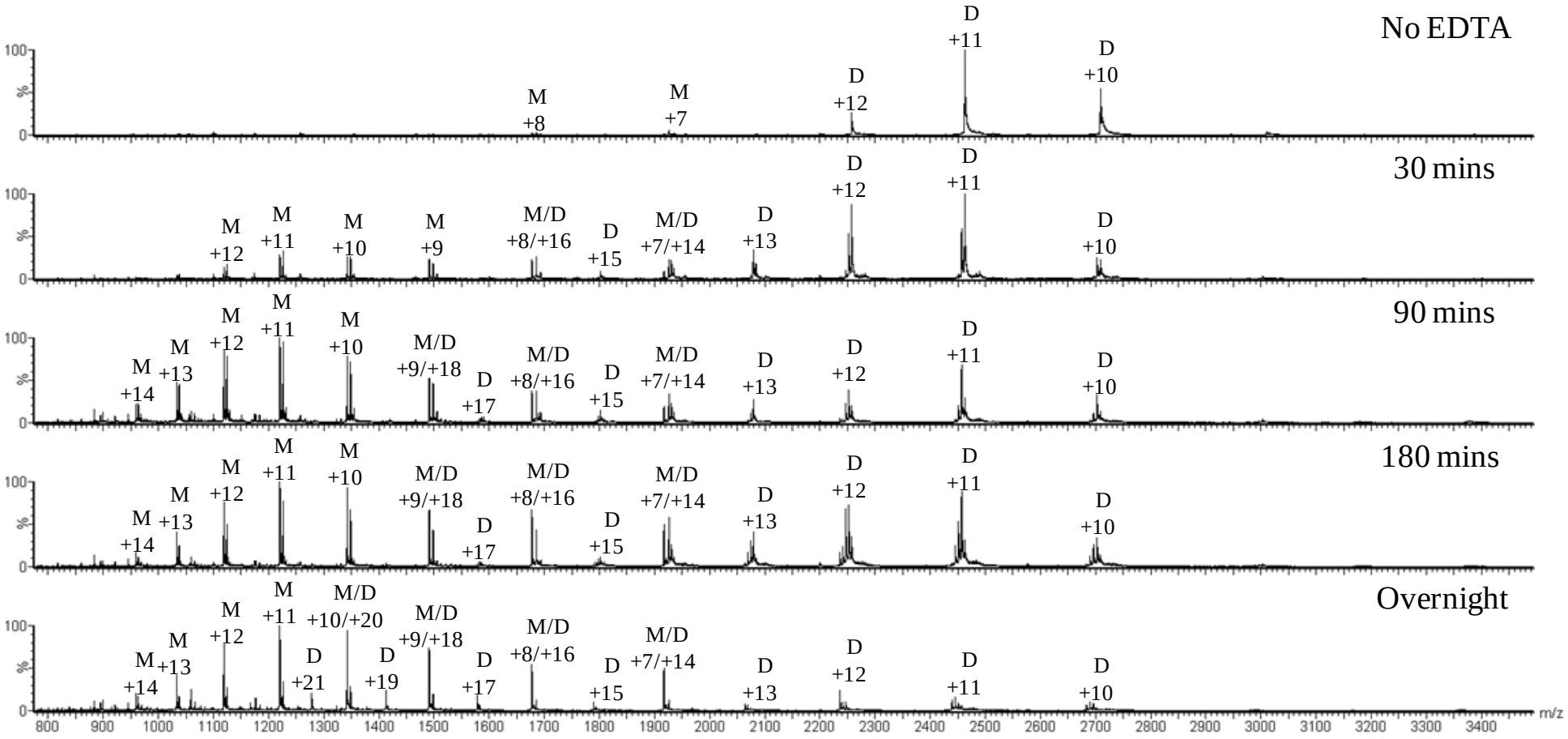


Figure 3.12: ESI mass spectra of 0.8 mM SmtB in 10 mM ammonium bicarbonate pH 7.4 (A) incubated with 1.6 mM EDTA and sampled after 30 mins, 90 mins, 180 mins and overnight.

Initially the relative percentage of monomeric species was determined as approximately 15 %. This is slightly higher than the percentage thought to exist in the solution phase at these concentrations. The shift in equilibrium, towards the monomeric species began within the first 22 minutes and can be clearly observed at the 30 minute time point. Not only can a higher relative abundance of the +7 and +8 monomeric charge states be seen, but there is also the appearance of highly charged monomeric species, Figure 3.12. After 180 minutes the percentage of monomeric species had increased to almost 50 %. The shift in equilibrium continued as the sample was incubated overnight, with the final mass spectrum showing that the SmtB is now predominately monomeric (67 %), with only a few peaks corresponding to dimer, most of which were still associated with some zinc ions.

The loss of zinc ions with time, from dimeric SmtB, can be visualised more easily by focusing on a single dimeric charge state, Figure 3.13. The +11 charge state was chosen as it was the only exclusively dimeric peak and was present in high abundance throughout the time course and at time point zero. Focusing on an odd dimeric charge state prevents monomeric interference affecting the abundances of certain species, especially when the abundance of monomeric species increases at the latter end of the time course. Figure 3.13A shows the +11 species observed at specific points during the time course. Data were converted into Figure 3.10B by calculating the relative abundance of each +11 species in relation to the total abundance of all +11 dimeric SmtB species at each time point throughout the time course. This plot provides a representation of the speciation of the +11 species with time, however, different dimeric charge states, in particular the highly charged +17, +19 and +21 species may have a slightly different speciation. These charge states were therefore investigated separately.

The loss of zinc ions from dimeric $SmtB_2Zn_4$ began almost straight away; resulting in the appearance of a small amount of $SmtB_2Zn_3$ within seven minutes, Figure 3.13B. Zinc ion removal from the fully loaded species continued slowly so that after 134 minutes approximately two thirds of the initial $SmtB_2Zn_4$ had been converted to $SmtB_2Zn_3$, which became the most prominent dimeric species. At this time point approximately 15 % of the charge state intensity distribution could be attributed to $SmtB_2Zn_2$ with a very small proportion existing as $SmtB_2Zn$ or $SmtB_2$.

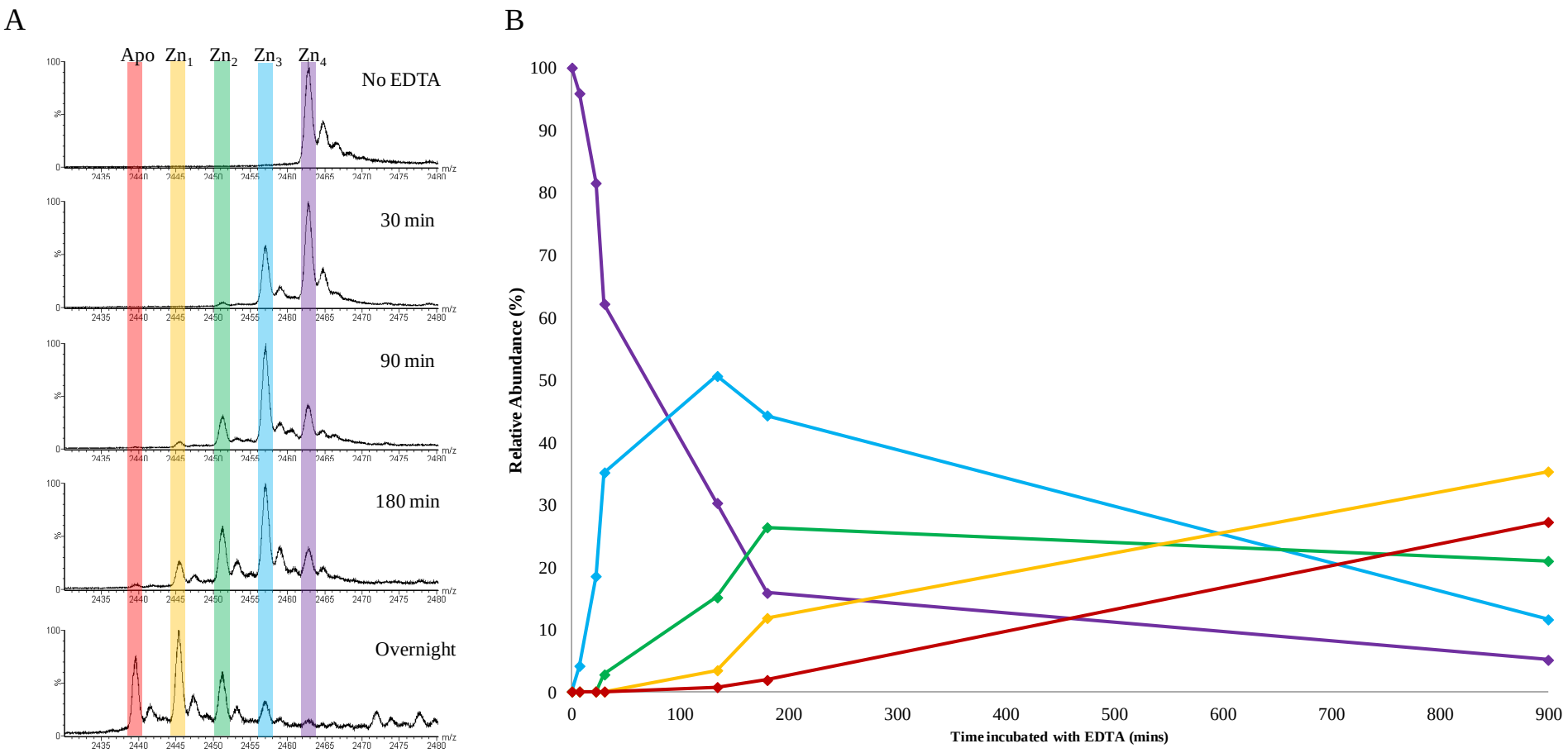


Figure 3.13: The speciation of the +11 charge state dimeric SmtB throughout an EDTA time course.
A close up of the + 11 charge states at different time points (A) accompanied by a plot of relative abundance verses incubation time (B). The species are coloured coded as follows Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red =Apo.

After incubation for 180 minutes the reaction had progressed so that all eight potential species of SmtB could be clearly visualised, Figure 3.14 and Table 3.2.

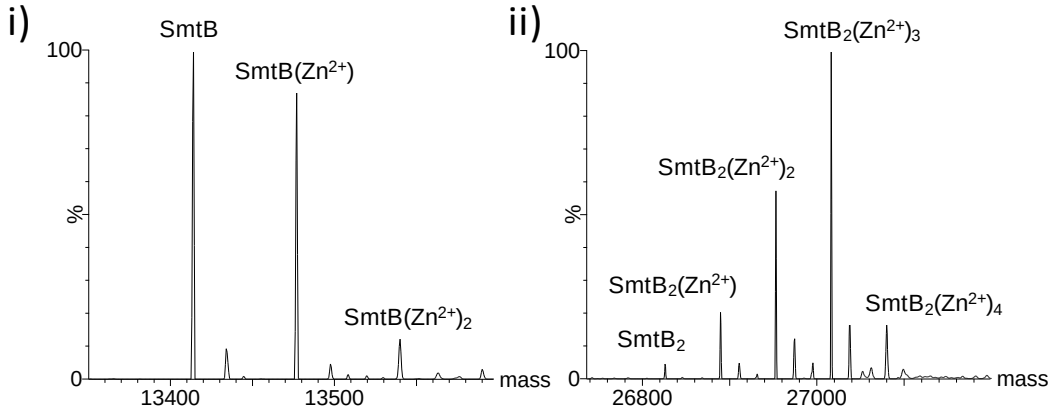


Figure 3.14: Deconvoluted spectra of predominately i) monomeric and ii) dimeric species SmtB after incubation with EDTA for 180 minutes, revealing the different zinc ion: protein stoichiometries.

A deconvolution of the dimeric species after the sample had incubated for 180 minutes produced a distribution of dimeric species in good agreement with the proposed speciation of the +11 charge state. Following the speciation of this charge state can therefore present an accurate representation of the zinc ion removal of the folded dimeric species present within the sample. The observed masses, produced from deconvoluting the raw data, corresponded well with the predicted masses, although a slightly larger mass difference was observed for these samples than for those obtained in the absence of EDTA, especially for the dimeric species.

The highly charged +17 and +19 unfolded species were observed to have a different speciation compared to the deconvoluted spectrum. These two charge states, although very low in abundance, were predominantly composed of apo-$SmtB_2$, $SmtB_2Zn$ and $SmtB_2Zn_2$ with no obvious signals from higher zinc ion stoichiometries. This suggests that the removal of zinc ions causes an unfolding of the dimeric structure, promoting the species with less zinc ions bound to exist at higher charge states.

Table 3.2: The mass differences of all observed species were within ± 0.7 Da of the expected masses of SmtB after the incubation with EDTA for 3 hours.

Monomeric SmtB	Expected mass (Da)	Observed mass (Da)	Mass difference
Zn_2	13540.1	13540.0	-0.1
Zn	13476.7	13477.0	0.3
Apo	13413.3	13414.0	0.7

Dimeric SmtB	Expected mass (Da)	Observed mass (Da)	Mass difference
Zn_4	27080.2	27079.5	-0.7
Zn_3	27016.8	27016.5	-0.3
Zn_2	26953.4	26953.0	-0.4
Zn	26890.0	26889.5	-0.5
Apo	26826.6	26826.0	-0.6

The identified metal ion binding sites within the SmtB dimer are known to have a high affinity for zinc ions (VanZile et al. 2002b), but there is little knowledge surrounding the kinetics of either of the sites. As is the case in all known metal sensor proteins (Guerra and Giedroc 2012), both sets of Zn binding sites are located close to the surface of the protein. Due to the high solvent exposure of the metal ligating residues, these sites are predicted to be ideal for ligand exchange reactions (VanZile et al. 2002b). The data produced from the EDTA time course described here, indicated that the kinetics of zinc ion removal from these sites is slow, with no obvious difference between the α5 and α3N sites. This kinetic stability may be the result of secondary interactions formed between the metal ligating residues and other protein residues (Namuswe and Berg 2012). The hydrogen bond network set up between the non-bonding imidazole NH group of H117, within the α5 site, and L83' on the αR helix (Eicken et al. 2003), may well act to decrease the rate of zinc ion dissociation from this particular site. A high Zn affinity coupled with a slow rate of Zn ion removal would be of potential benefit to the cyanobacterial Zn homeostatic system. The attributes displayed by SmtB could prevent the re-formation of the apo-SmtB:DNA complex, allowing for the constant expression of SmtA under zinc stress conditions. Only when an appropriate level of Zn has been reinstated in the cytosol would apo-SmtB be produced and the homeostatic system reset. The on-rate of Zn binding to SmtB, in particular the α5 sensing sites, is however, predicted to be much quicker. This would ensure a faster response to an increase in cytosolic Zn

concentration. To date no zinc sensing protein's association kinetics have been studied (Guerra and Giedroc 2012). The kinetics of the ArsR/SmtB family member CadC, measured by stopped-flow absorbance spectroscopy has been revealed to be biphasic (Busenlehner and Giedroc 2006). An initial fast [CadC] dependent step was suggested to correlate to lead binding to the α3N site followed by a [CadC] independent rate, thought to represent the proteins conformation change. The rate of CadC metal dissociation was determined to be extremely slow (Busenlehner and Giedroc 2006), similar to the proposed dissociation rate for SmtB. More kinetic experiments would have to be carried out to accurately determine SmtB's, zinc on and off rates. It is possible that, if SmtB does act like CadC, it is not only a zinc sensor, but a zinc sequestration protein.

3.3.4.3 EDTA titration mass spectrometry

To investigate the speciation of SmtB after incubation with different molar equivalents (mol equiv.) of EDTA, an EDTA titration was conducted and characterised utilising mass spectrometry, Figure 3.15. In the absence of EDTA a mass spectrum of SmtB, after four hours incubation at room temperature, retained a tight dimeric charge state envelope spanning the +9 to +13 $SmtB_2Zn_4$ charge states, Figure 3.15. This is indicative of a natively folded protein conformation (Chowdhury et al. 1990) (Loo et al. 1991) (Light-Wahl et al. 1994), over the duration and conditions of the experiment. This sample was subsequently used as a negative control. Longer EDTA incubation periods would have ensured that metal binding equilibrium was fully achieved, but it may be that the time period required to ensure equilibrium, would result in extensive protein oxidation (Section 3.3.4.4). The oxidation of the six cysteines present within the SmtB dimer may well affect the conformations sampled by SmtB and therefore a compromise of four hours incubation was employed. This was deemed long enough, from the time course experiment, to achieve the formation of apo-SmtB but limit sample oxidation.

Consistent with the EDTA time course, if EDTA titration is followed across the full scan range, a change in monomer: dimer equilibrium can be observed. The tight charge state distribution, observed in the control sample was largely maintained up to the addition of 2 mol equiv. of EDTA, in respect to bound zinc ions, although there was a slight broadening of the envelope, which may be indicative of a small

conformation change. Additional peaks also appeared within the spectrum; upon deconvolution the mass of this species was determined as 26754.5 Da. After a study of the primary structure of SmtB and with the knowledge that the N-terminal regions are highly flexible, it is likely that these low intensity peaks correspond to dimeric SmtB with a loss of three N-terminal amino acids (TKP), Figure 3.15A. The addition of 3 mol equiv. of EDTA resulted in a more dramatic and noticeable shift in speciation. Peaks emerge within the low *m/z* range, which have been assigned to be monomeric peaks and/or unfolded highly charged dimeric peaks. The relative abundance of these peaks were larger after SmtB was incubated with 4 mol equiv. of EDTA, Figure 3.15A.

To visualise the loss of zinc ions, from the SmtB dimer the +11 charge state was once again followed more closely. As zinc ions were chelated by EDTA, peaks with lower *m/z* values than the fully loaded species appear, until all five peaks, from apo to fully loaded dimer are visualised, Figure 3.15B. Interestingly the major shift, in monomeric:dimeric equilibrium, observed in the full scan, Figure 3.15A, corresponded to the addition of 3 mol equiv. of EDTA and the appearance of distinctive apo-$SmtB_2$ and $SmtB_2Zn_1$ peaks, Figure 3.15B. This observation suggests that these two dimeric species are more susceptible to unfolding and dissociation than the higher zinc loaded species, perhaps suggesting a role for at least these final two zinc ions in the maintenance of the SmtB dimer.

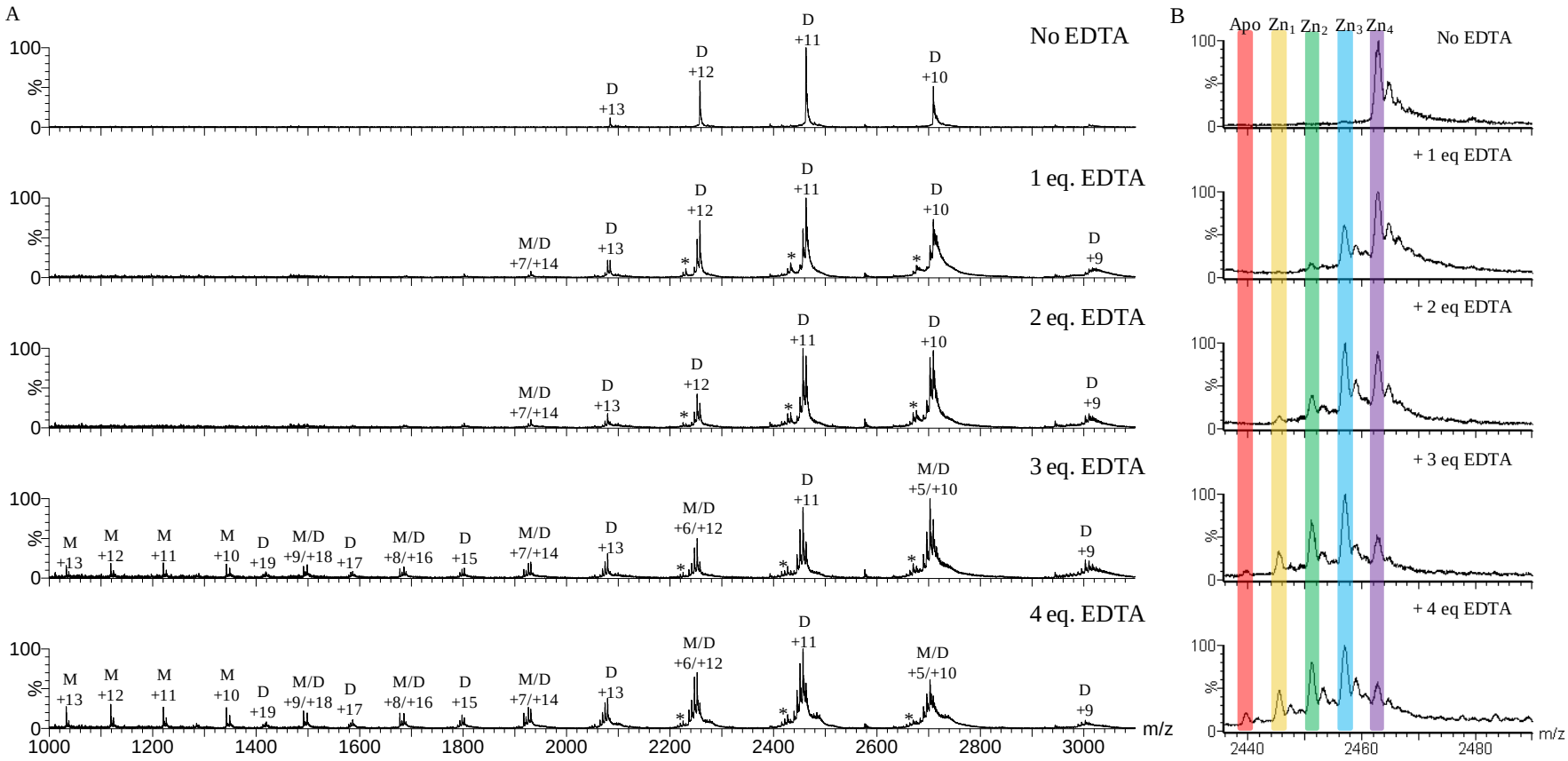


Figure 3.15: (A) Mass spectra obtained, using a cone voltage of 40 V, from experiments in which 0.8 mM SmtB was incubated, for four hours, with the indicated concentrations of EDTA.
SmtB peaks resulting from a 3 residue N-terminal cleavage are highlighted with *. (B) An enlargement of the + 11 dimeric species, displaying the speciation of dimeric SmtB under each EDTA concentration.

Dimeric instability, resulting in the production of monomeric species and highly charged dimeric species could occur either in the solution and/or in the gas phase. The partial unfolding, or increased flexibility, of the dimeric conformation, due to the loss of zinc ions in the solution phase, is likely to expose previously buried basic amino acid residues. The presence of these residues could result in SmtB acquiring more protons during the electrospray process, resulting in the formation of higher dimeric charge states (Chowdhury et al. 1990). The dimeric association constant of SmtB is known to be lower in the absence of zinc ions (Kar et al. 1997). This reduction in dimeric interactions could promote the appearance of monomeric species in the solution phase but may also have an impact on the species observed in the gas phase, as weaker interactions between monomers will increase the occurrence of mass spectrometry-based dissociation. If this process were to occur in the gas phase the partitioning of charges has been observed to be not symmetrical (Dodds et al. 2011), leading to a wide distribution of monomeric charge states. A combination of these processes could explain the observed shift towards higher charged species at lower *m/z*.

The α3N sites, within SmtB, have been calculated to have a 20 fold higher affinity for zinc ions than the α5 sites (VanZile et al. 2002b). In the event of the reaction reaching equilibrium the presence of 2 mol equiv. of EDTA should ensure the removal of Zn ions from the α5 sites whilst the α3N sites should remain occupied. If this were the case, a build up of the $SmtB_2Zn_2$ species should clearly occur, whilst the production of $SmtB_2Zn$ should only occur if the sample contains more than 2 mol equiv. of EDTA. The observed loss of zinc ions from the dimeric +11 charge state, in the presence of 2 mol equiv. of EDTA, showed no build up of the $SmtB_2Zn_2$ species and a clear, albeit low abundant, $SmtB_2Zn$ species. The lack of a predominant $SmtB_2Zn_2$ species was surprising, since this would correlate to species observed in the majority of the recent literature (VanZile et al. 2000; VanZile et al. 2002a). These observations therefore suggest that the binding affinities of the two sites may be closer than previously determined. If this were the case it would explain why the zinc ion in the $SmtB_2Zn$ X-ray crystal structure was found bound to a supposedly lower affinity α5 site, instead of an α3N site (Eicken et al. 2003).

Gas phase stability experiments

The EDTA titration was conducted at two different cone voltages, 40 V and 60 V, in order to investigate gas phase stability. Samples were also studied using a cone voltage of 20 V. This caused a significant loss of sensitivity and also allowed unspecific binding of ions to occur. Due to the unspecific binding observed, which prevented accurate zinc speciation of SmtB to be determined, these results are not discussed in more detail.

The use of higher cone voltage is known to cause dissociation of species by in-source collisionally induced dissociation (CID) (Loo et al. 1988). At higher cone voltages the spectra produced from the EDTA titration samples looked very similar, until the addition of 3 and 4 mol equiv. of EDTA. In these spectra a shift in the equilibrium towards the monomeric species occurred consistent with the samples investigated at 40 V, but the charge states corresponding to monomeric species were more intense, Figure 3.16. These results suggest that dimeric SmtB, when stripped of zinc ions, has a lower stability in the gas phase resulting in the production of more monomeric species when higher energies are used.

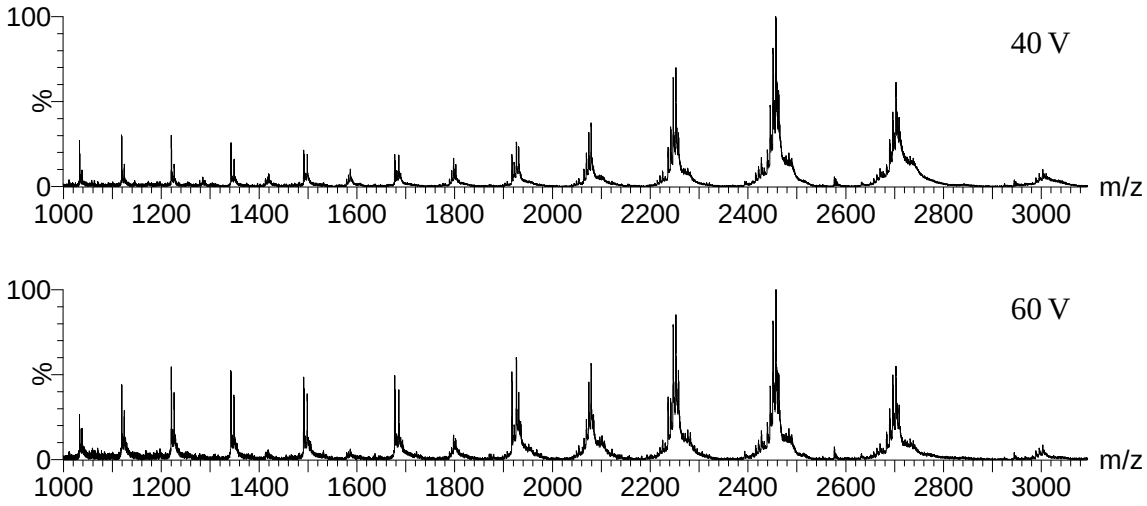


Figure 3.16: Mass spectra of 0.8 mM SmtB incubated with 1.6 mM EDTA for 4 hours. The sample cone voltage is indicated within each spectrum.

To investigate the stability of dimeric SmtB species to in-source CID the relative abundance of each +11 species, throughout the EDTA titration, were calculated at

both experimental cone voltages, Figure 3.17. The appearance of SmtB species with fewer zinc ions bound was consistent between the two data sets, Figure 3.17. This suggests that zinc ions are being removed in the solution phase and not due to any collisions experienced in the gas phase. The relative abundances of all 5 SmtB species between the two data sets were also similar, even when incubated with 3 or 4 mol. equiv. of EDTA. These results suggest that there is no single species of folded SmtB more susceptible to dissociation in the gas phase.

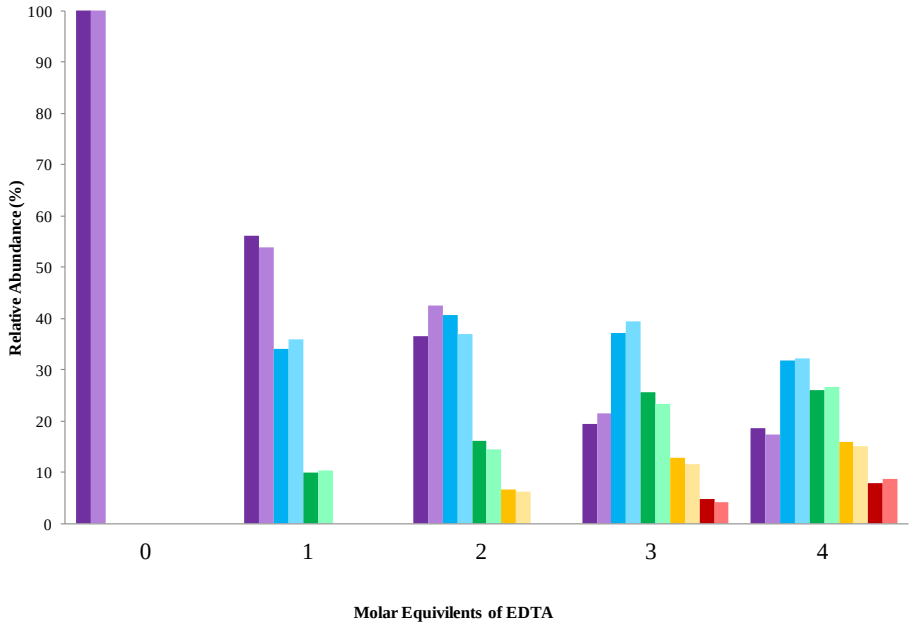


Figure 3.17: The speciation of dimeric SmtB after four hours of incubation with different molar equivalents (0-4) of EDTA in respect to zinc binding sites at cone voltage 40 V (Dark colours) and 60 V (light colours).
The dimeric species are coloured coded: Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red =Apo.

It is therefore more plausible that the highly charged monomeric species, observed at an increased abundance at high cone voltages, are due to the dissociation of the highly charged +17 and +19 dimeric species. These charge states contain a higher proportion of zinc depleted species which appear to be more labile at high energies. In samples exposed to more than 2 mol equiv. of EDTA, removal of ions from the α3N sites would occur. In the absence of these, higher affinity α3N site zinc ions, SmtB is likely to be more flexible, especially in the N-terminal regions. This increased flexibility, combined with exposure to high cone voltages, may cause unfolding of the protein to occur, allowing the dimeric species to pick up more

positive charges. Increased charge repulsion combined with the higher energies and inherent gas phase instability of SmtB could promote further unfolding, dimer dissociation and unsymmetrical charge distribution, leading to a higher proportion of highly charged monomers.

3.3.4.4 Apo-SmtB

The kinetics of zinc ion removal from SmtB utilising the chelator EDTA proceeds slowly. In both the EDTA time course and titration experiments only a small proportion of the sample was completely depleted of zinc ions, resulting in complex spectra which were difficult to interpret. To visualise a pure apo-SmtB sample the zinc ions bound during expression of the protein had to be removed via dialysis against EDTA for up to 48 hours. This procedure was carried out under anaerobic conditions to prevent the oxidation of the six free cysteine residues, within the SmtB dimer. The sample was then further purified to remove the majority of EDTA adducts. The resulting sample was then studied using native mass spectrometry and ion mobility mass spectrometry. These techniques indicated that the apo-SmtB sample was largely monomeric and partially unfolded, Figure 3.18 and Figure 3.19. Higher buffer strengths were employed to determine whether these conditions would stabilise the dimeric apo-SmtB but under every condition the monomeric species remained dominant.

Three main charge envelopes are present within the mass spectrum of apo-SmtB, Figure 3.18. The two most intense charge state distributions, which overlap, correspond to monomeric species, whilst the low intensity charge states are due to dimeric species. The presence of two monomeric charge state envelopes suggests that two different conformations of monomeric apo-SmtB are likely to exist in solution. The highly charged monomeric species are likely to be due to an unfolded monomeric species. The loss of interactions holding together a protein's tertiary structure often results in a more extended conformation, which exposes more basic sites to protonation during ESI (Chowdhury et al. 1990).

The monomeric species $[SmtB+8H]^{8+}$, $[SmtB+7H]^{7+}$ and $[SmtB+6H]^{6+}$are likely to be, at least partially, derived from either folded monomer present in solution or

folded monomer dissociated from the dimeric species. These charge states were also observed in mass spectra of zinc loaded SmtB and may be the result of the monomer/dimer equilibrium in solution. These results are in good agreement with the data produced during the EDTA time course and titration, which also show a larger proportion of monomers and higher charged dimers for apo-SmtB.

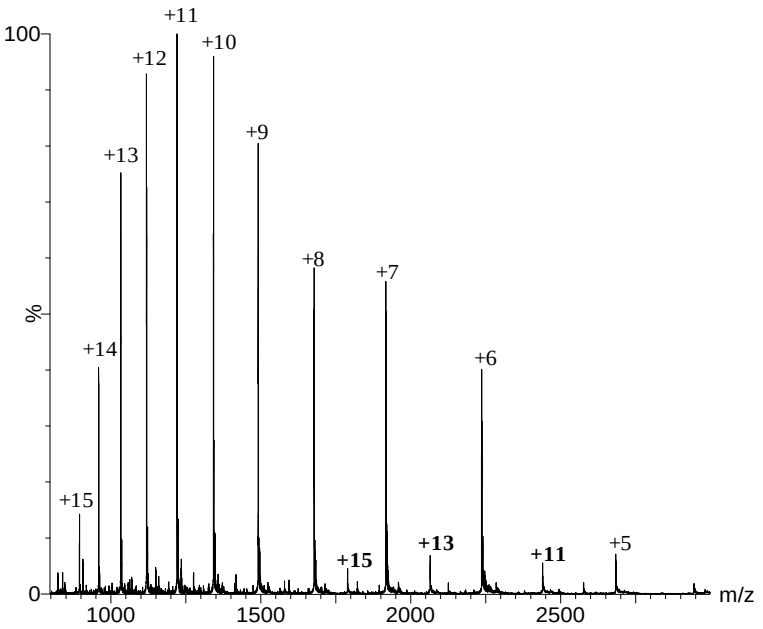


Figure 3.18: Mass spectrum of apo-SmtB in 10 mM ammonium bicarbonate pH 7.4 Bold type highlights charge states corresponding to dimer whilst monomeric species are labelled in normal type.

Apo-SmtB: ion mobility mass spectrometry

Estimated collisional cross-sections (CCS) calculated from ion mobility data, also suggested the presence of a number of conformational families within apo-SmtB, Figure 3.19. These results provide further information regarding apo-SmtB, including the presence of multiple conformations within single *m/z* peaks, such as the monomeric $[SmtB+7H]^{7+}$ and $[SmtB+8H]^{8+}$ peaks. These results suggest that the peaks contain both a compact and a more extended species which are likely to correlate to both folded and unfolded conformations respectively. The theoretical CCS of monomeric apo-SmtB was determined by removing one of the two protomers from the apo-SmtB dimeric X-ray crystal structure. The EHSS calculations determined the CCS as 1607 Å², which potentially separates the folded conformations from the unfolded ones.

A small amount of apo-SmtB dimer was also seen ($[SmtB_2+15H]^{15+}$ to $[SmtB_2+11H]^{11+}$) in the mass spectrum, Figure 3.18. This charge state distribution is wider and more highly charged than that of the zinc loaded SmtB dimer ($[SmtB_2+4Zn+5H]^{13+}$ to $[SmtB_2+4Zn+2H]^{10+}$). This suggests that without zinc ions the SmtB dimer has a more extended conformation, allowing more basic sites to be exposed and protonated.

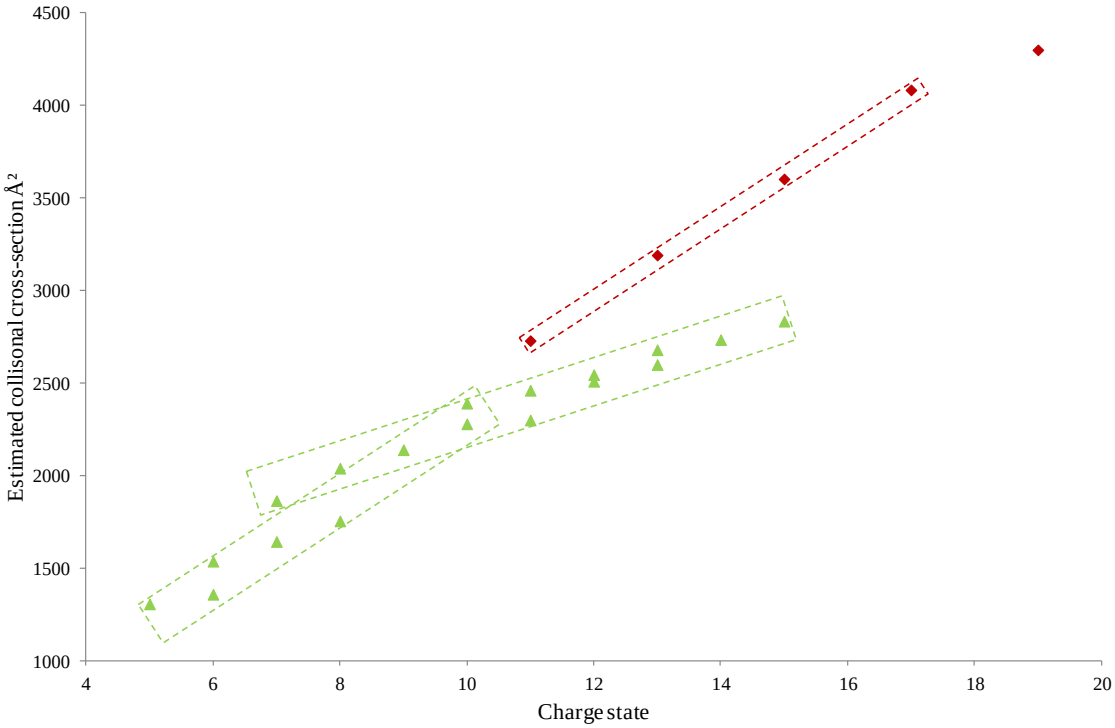


Figure 3.19: Estimated collisional cross-section plot of an apo-SmtB sample. Dimeric (red) and monomeric (green) species are shown along with dashed boxes indicating possible cross-sectional families.

A comparison of the CCS's from a single charge state reveals the increased conformation of the apo-SmtB dimer, Figure 3.20. The +11 charge state for apo-SmtB had a CCS 59 Å² larger than its zinc loaded counterpart, at 2729 Å² and 2656 Å² respectively. The difference between the apo and zinc loaded dimeric +13 charge state conformations however is significantly larger, 251 Å². This implies that the apo-SmtB dimer is a flexible structure which is more susceptible, in comparison with its zinc bound counterpart, to unfolding to reduce Coulomb repulsion. This repulsion is caused by the addition of protons in the gas phase and is comparable to acid induced unfolding in the solution phase (Jarrold 1999). The presence of the four bound zinc ions within SmtB may constrain the protein structure, allowing structural

expansion but limiting unfolding in response to Coulomb repulsion. This form of conformational constraint has been observed previously when disulphide bonds are present within the protein structure (Shelimov et al. 1997), but there is also evidence that ligands, such as a heme group or metal ion, may act in a similar fashion (Scarff et al. 2008) (Faull et al. 2010).

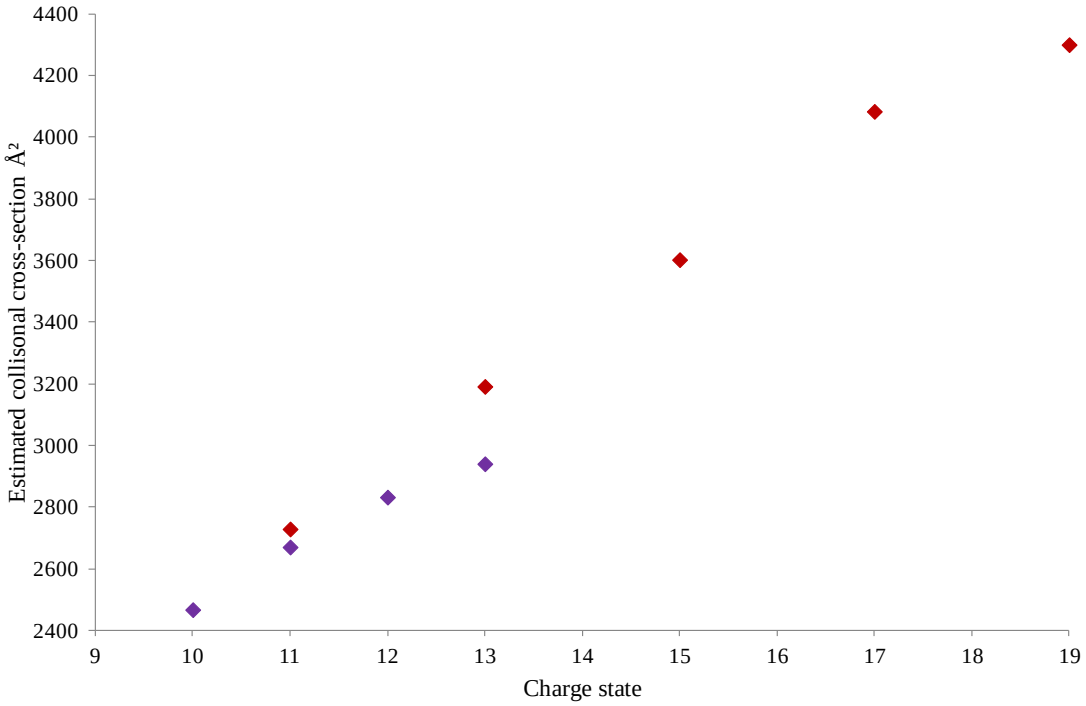


Figure 3.20: Estimated collisional cross-section plots comparing apo (red) and zinc loaded (purple) dimeric SmtB samples.

Sedimentation equilibrium results (Kar et al. 1997) indicate that upon metal depletion the equilibrium, between SmtB dimer and monomer, is pushed towards the monomeric species. Under the conditions outlined in their study (0.1 M NaCl, pH 7.4, 22 °C) the proportion of dimeric SmtB changed, from 97 % to 81 %, after zinc ion removal (Kar et al. 1997). Using the dissociation constants calculated from this work (apo K_d = 3.1 x 10^{-6} M and Zn_4 = 8.0 x 10^{-8} M) and the concentrations used within this study, an estimation of the proportion of monomer and dimer can be calculated. This reveals that the percentage of dimeric SmtB should fall, upon zinc ion removal, from 92 % to around 51 %. Deconvolution of the apo-SmtB spectra obtained revealed the relative percentage of gas phase dimer was between 12-20 %, Figure 3.22. The loss in solution phase stability and associated shift in equilibrium

does not fully explain the apo-SmtB mass spectrum observed, Figure 3.18, which may suggest a loss in stability in the gas phase.

The X-ray crystal structure of apo-SmtB reveals that the dimeric interface is made from a hydrophobic core comprising six α-helices, three from each monomer, Figure 3.21. The interactions between the two monomer units are therefore predominately hydrophobic, with a total of 111 hydrophobic contacts and eight hydrogen bonds identified (Cook et al. 1998).

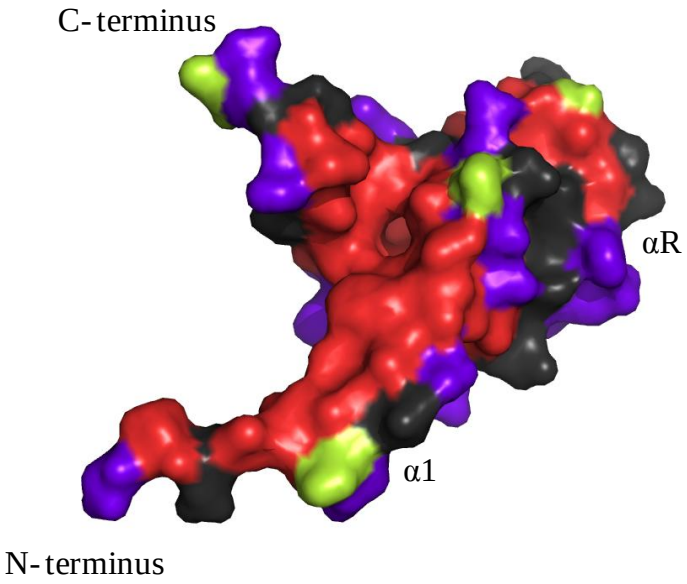


Figure 3.21: A monomer of SmtB showing the potential dimeric interface.
Created from the apo-$SmtB_2$ crystal structure (1R1T). Amino acids have been colour coded so that red corresponds to hydrophobic residues, purple = charged residues, black = polar residues.

In the presence of water this large number of hydrophobic interactions effectively stabilises the dimeric structure, with the help of the hydrogen bonds. In the gas phase, however, the contributions from different forms of non-covalent interactions change. Results suggest that electrostatic interactions and hydrogen bonds are intensified whilst hydrophobic interactions are lost (Robinson et al. 1996). The apo-SmtB dimer is therefore only held together by eight hydrogen bonds in the gas phase, providing an explanation for its reduced stability and lower observed abundance.

The stability provided by the bound zinc ion in the solution phase is significantly more important in the gas phase. When zinc ions are bound they bridge the monomers over four separate sites, effectively doubling the stabilising interactions conserved in the gas phase. This suggests why zinc bound dimers can be observed at high abundance. Upon dialysis against EDTA in aerobic conditions a higher percentage of dimeric species were observed, corresponding to approximately 59 % of the total mass spectrum, Figure 3.22.

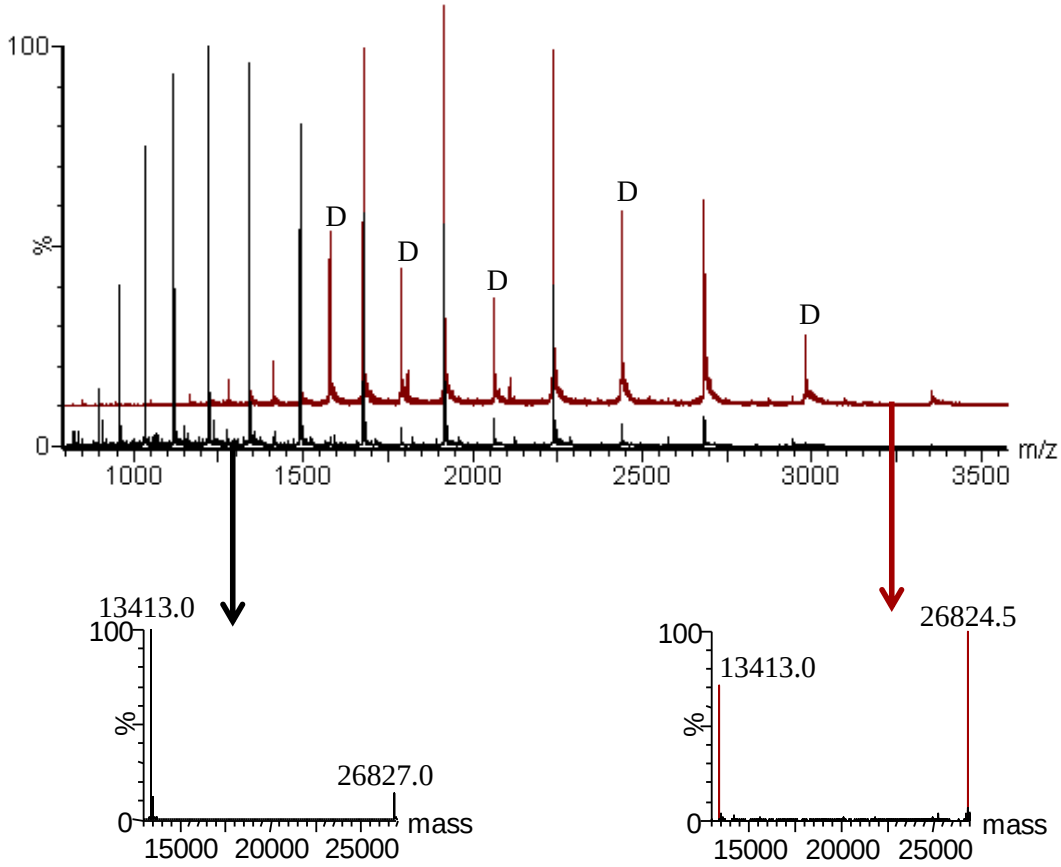


Figure 3.22: A comparison of the apo-SmtB mass spectra produced after dialysis in aerobic (red) and anaerobic (black) conditions. Arrows point to deconvoluted spectra showing the masses of the monomeric and dimeric species.

The mass of the apo-SmtB dimer as determined from deconvolution was 26824.5 ± 0.5 Da. This is 2.1 Da lower than the expected value, which could suggest the oxidation of at least two Cys residues, forming a disulphide bond. These results, combined with the higher proportion of dimeric species, suggest that a disulphide bond may have been formed between C14 and C61'/ C14' and C61, within the α3N site. Since these residues are on different monomers this covalent bond could

stabilise the dimeric structure in the gas phase, allowing a higher proportion of dimeric species to be observed, Figure 3.22. A dimer stabilising disulphide bond has previously been observed between C14 and C61 in an air oxidised C121S SmtB mutant (VanZile et al. 2002b).

3.3.4.5 EDTA titration: ion mobility mass spectrometry

The difference in CCS between $SmtB_2Zn_4$ and apo-$SmtB_2$ for the +11 charge state indicated an expansion of the proteins conformation upon zinc ion removal. To investigate the possibility of subtle conformational changes in dimeric SmtB further ion mobility mass spectrometry (IMMS) experiments were conducted on each zinc ion stoichiometry produced throughout the EDTA titration experiment. A number of different conformational changes had the potential to be visualised. The loss of the first lower affinity zinc ions, from the α5 sites, should reveal any conformational change involved in zinc sensing; whilst the loss of the final two, higher affinity α3N, zinc ions could reveal either a further sensory conformational change or a more extensive conformational change, due to dimeric unfolding. ATD's from the +11 dimeric charge state were extracted for two reasons. Firstly the charge state had to be odd, to prevent any possible monomeric interference. Secondly the lowest charge state, at a sufficient abundance, was chosen as in many cases the lower charge states have been shown to provide the most biologically relevant protein conformation (Jarrold 2000). It should be noted that the lowest charge state does not always provide the most relevant conformation. A number of studies have shown that in certain proteins and protein complexes low charge states correspond to collapsed, biologically irrelevant forms (Hall et al. 2012) (Hopper and Oldham 2009).

At lower cone voltages all of the extracted ATD's had a single peak, Figure 3.23, representing one resolvable protein conformation, at these experimental conditions.

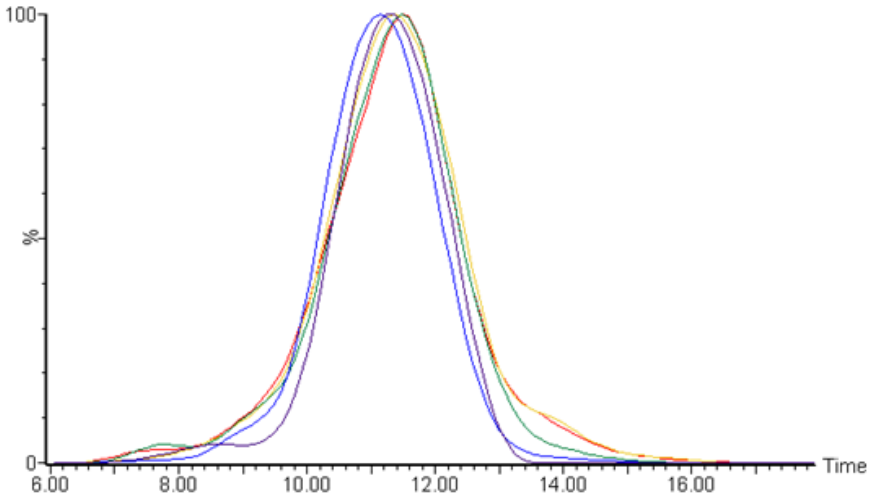


Figure 3.23: ATD's of the +11 dimeric SmtB at 40 V. The colours represent the different stoichiometries of zinc ions to protein after the incubation with EDTA. Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red =Apo.

The ATD of the $SmtB_2Zn_4$ species corresponded to an average estimated collisional cross-section (CSS) of 2649 $Å^2$. This CCS is larger than would be expected from a protein with a molecular weight of 27 kDa. The CCS of Bovine carbonic anhydrase II (BCA) (29 kDa) for example was estimated as only 2004 $Å^2$, Table 3.3.

Table 3.3: Estimated collisional cross sections of proteins which bracket the molecular weight of SmtB. [a](Bush et al. 2010), [b](Hopper and Oldham 2009), [c] (Leary et al. 2009).

Protein	Molecular weight (Da)	Estimated Collisional Cross Section
[a]Cytochrome c	12 k	1490 (+6)
[b]Lysozyme	14 k	`1487 (+8)
[c]BCA	29 k	2004 (+9)
[a]β-lactoglobulin	37 k	3230 (+11)
[a]Transthyretin	56 k	3840 (+14)
[a]Avidin	64 k	4250 (+15)

BCA is a monomeric protein, wheres SmtB is an elongated dimer and therefore its CCS woud be expected to be larger than a globular monomer of the same weight.

Species with fewer bound zinc ions had slightly longer arrival times and larger averaged estimated CCSs; $SmtB_2Zn_3$ (2659 Å²), $SmtB_2Zn_2$ (2675 Å²), $SmtB_2Zn$ (2695 Å²) and $SmtB_2$ (2703 Å²). There was an observed increase of approximately 29 Å² corresponding to the loss of zinc ions from the lower affinity sensory α5 sites and then a further increase of 28 Å² upon the loss of zinc ions from the higher affinity α3N sites, Figure 3.24. Together these gave a combined expansion of 57 Å² between $SmtB_2Zn_4$ and $SmtB_2$, which agrees well with the data produced from the dialysed samples.

The general trend upon zinc ion release from either set of sites was an expansion of the CCS. The total increase, however, in the estimated CCS was small (~2.1 %), although this was larger than the maximum error observed (~1.2%), the expansion associated with the loss of zinc ions from individual sites is probably not large enough to be deemed biologically significant.

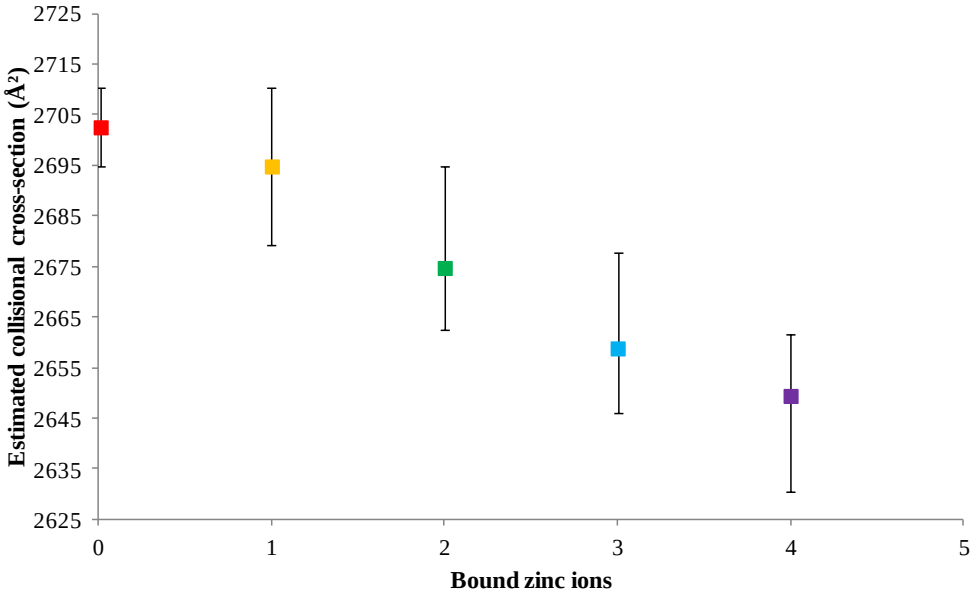


Figure 3.24: Average estimated collisional cross-sections of the +11 dimeric SmtB. The colours represent the different stoichiometries of zinc ions to protein after the incubation with EDTA. Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red =Apo.

The errors associated with the estimated CCSs were calculated by observing the CCS of each species under all five experimental conditions. They are likely to be the

result of each species occupying a family of conformations. These conformations are very similar and are therefore not resolved in the ATD, leading to the appearance of one observable peak. If the SmtB conformations are in dynamic equilibrium, slight changes in this equilibrium could cause shifts in the ATD maxima to be observed, leading to higher associated errors.

There are four published X-ray crystal structures for the dimeric species of SmtB available in the PDB. None of these structures are complete, with all the structures having residues missing from the N-terminal portions of both monomers (Eicken et al. 2003). Theoretical CCS were calculated from three of the four structures available, after the addition of missing hydrogen atoms. These calculated CCS values were then used in a comparison with the experimental data obtained from IMS experiments. The hard sphere scattering (EHSS) model, calculated the theoretical cross-section of $SmtB_2$ (1R1T), $SmtB_2Zn$ (1R23) and $SmtB_2Zn_2$ (1R22) to be 2343 Å², 2316 Å² and 2184Å² respectively. All these structures differ slightly in the number of amino acids present. Each structure is missing at least 41 aa, (1R1T - 45 aa, 1R23 - 41 aa, 1R22 - 52 aa), Figure 3.26.

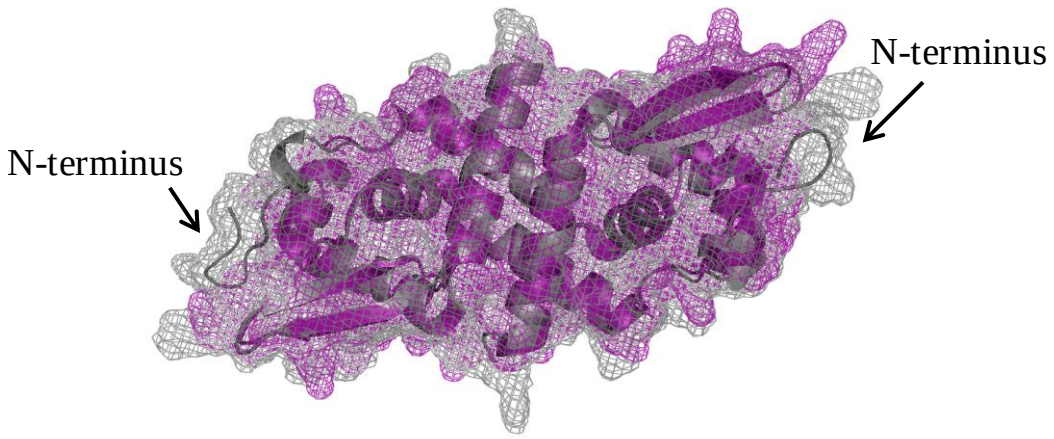


Figure 3.25: An overlay of the 1R23 (pink) crystal structure and proposed model structure (grey), highlighting the effect of additional N-terminal residues.

The extension of the 1R23 crystal structure to produce the model for $SmtB_2Zn_4$ visualises the contribution of these missing amino acids to the CCS. The addition of

19 aa, predominately in the N-terminal regions of this structure, extended the polypeptide chain to form the α3N sites, Figure 3.25 and 3.26.

The CCS of this structure was determined as 2520 Å^2, 205 Å^2 larger than the corresponding 1R23 CCS. The model is still missing 2 x 11 amino acids of the N-terminal regions. This data re-enforces that the CCS of SmtB is larger than expected from its molecular weight and goes some way to explain why the estimated CCS produced from experimental data are significantly larger (> 350 Å^2) than those determined from theoretical calculations.

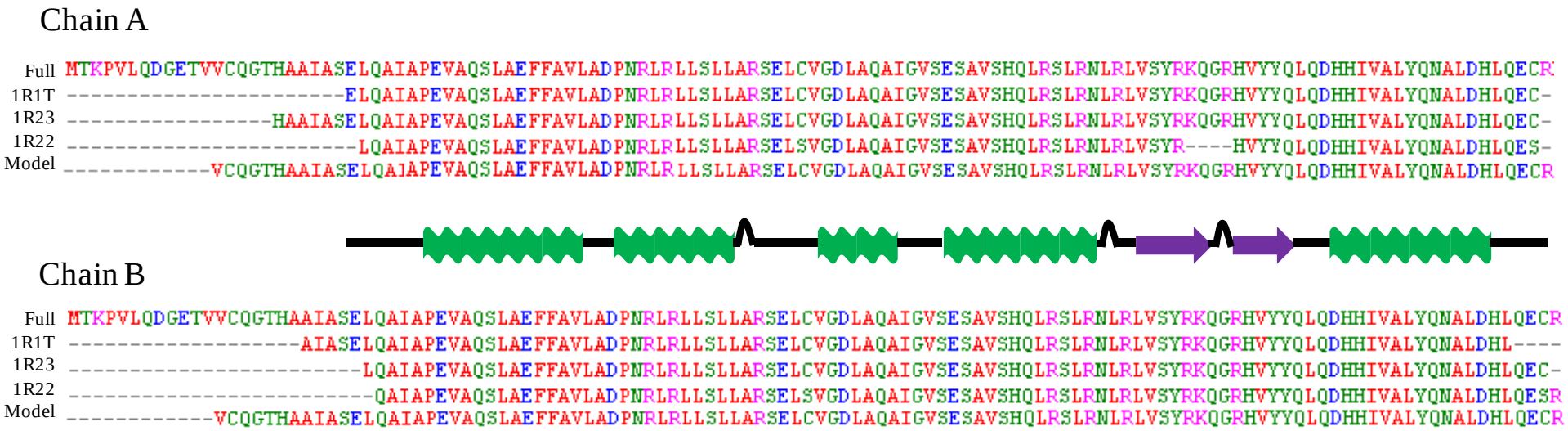


Figure 3.26: Multiple sequence alignments of the full SmtB dimer sequence, the constructs used in x-ray crystallography studies and the model produced from the extension of the 1R23 sequence.
Multiple sequence alignment was created by ClusterW2 (Goujon et al. 2010). Sites involved in metal ion binding are boxed in black and grouped into α3N and α5 sites. Residues are colour coded; small hydrophobic amino acids (red), acidic (blue), basic (pink) and other (green). In between the alignment of the two chains is a schematic of the proposed secondary structural elements from the 1R1T structure. This highlights the regions of secondary structure lost from the 1R22 construct.

Between the CCS calculated from the 1R22 and 1R1T crystal structures an increase in CCS of 159 Å² (6.8 %) was obtained. This expansion could correspond to the quaternary conformational change associated with the loss of the zinc ions from the α5 sites. The loss of presumably the same zinc ions, measured from the ion mobility experiments, lead to a calculated expansion of 29 Å² (1.1 %), over five times smaller than the predicted expansion.

A change in quaternary structure which induces an expansion of the structure is observed in the crystal structures; however it is unlikely to be as large as the difference in the crystal structures CCSs suggests. A part of the discrepancy between the expected and observed values is thought to be down to the 1R22 crystal structure having seven fewer amino acids than the 1R1T structure. Not only does 1R22 have fewer amino acids, which would cause it to have a smaller CCS, but the majority of these amino acids are missing from areas which are likely to greatly influence the CCS i.e., the N-terminus and protruding beta strand, Figure 3.27.

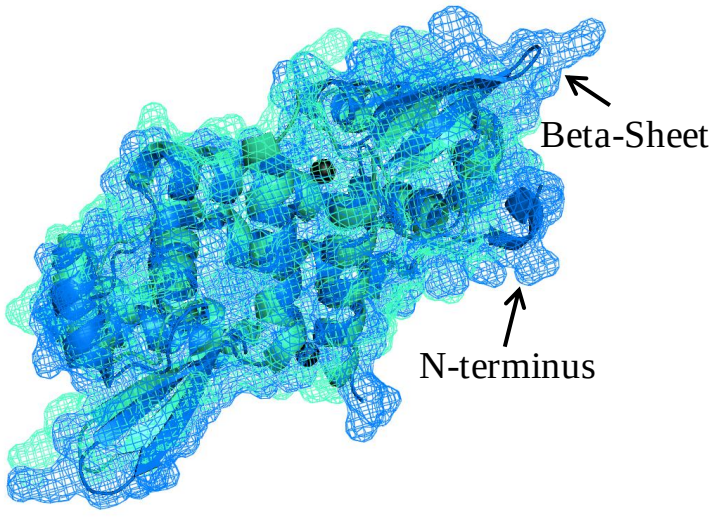


Figure 3.27: An overlay of the $SmtB_2Zn_2$ 1R22 (green) and apo-$SmtB_2$ 1R1T (blue) crystal structures, highlighting structural differences.

The re-addition of these residues would increase the CCS of 1R22, and therefore decrease the changes observed upon zinc removal. The high prevalence of incomplete crystal structures, particularly in mutli-complex assemblies, has prompted the combined use of computational methods, such as course-grained molecular modelling and homology modelling to complete structures (Politis et al.

2010). In this instance this approach could be used to complete the 1R22 structure to the same extent as the 1R1T, allowing the difference in CSS caused solely by the conformational change to be calculated and compared to the difference in CCS determined by IMS. The production of complete models of SmtB, in various zinc binding stoichiometries, may be possible. There is doubt however over whether models calculated using these techniques would be accurate enough to detect the subtle conformational changes produced by zinc ion binding (Atmanene et al. 2012).

The second conformational change, observed by IMMS, of 28 Å^2 is thought to relate to the expansion caused by the loss of the zinc ions from the α3N sites. None of the published X-ray crystal structures contain zinc ions in their α3N sites. In fact some structures have Cys to Ser mutations to prevent metal ion binding to the higher affinity α3N sites (Eicken et al. 2003). Any conformational change associated with this site has therefore never been studied at a molecular level. The compaction observed by sedimentation velocity data upon zinc ion binding to dimeric SmtB may however include the conformational change associated with zinc binding to the α3N site (Kar et al. 1997). The exact contribution made by the α3N site towards the compaction, however, can not be separated from the contribution made by the α5 site. The α3N site has been shown not to be necessary for zinc ion sensing (Turner et al. 1996). It has been proposed, however, that upon SmtB:DNA dissociation and zinc transfer to the α3N sites a further conformational change could occur, to prevent SmtB rebinding to the DNA (VanZile et al. 2002a). It is more likely that the conformational change observed, by ion mobility, upon removal of the final zinc ions from SmtB can simply be attributed to an increased flexibility within the N-terminal arms. The N-terminal region is always partially, and often completely, missing from X-ray structures, indicating that in the absence of zinc ions there is a high disorder in this region. Crystallography or NMR studies of SmtB with occupied α3N sites could reveal if the occupation of these sites reduces the flexibility in these N-terminal regions.

Experimental evidence suggests that SmtB exists as a flexible protein, likely to adopt a number of similar conformations, in both the apo and zinc bound forms. This may be due to the highly flexible β-wing domains or its inherent quaternary structural flexibility (Eicken et al. 2003). In the presence of bound zinc ions regions of the

protein, such as the metal binding helices and N-terminal arms are likely to become more stabilised. A hydrogen bond network, connecting the α5 site and the αR helix, crosses the dimeric interface in the $SmtB_2Zn_2$ C61S/C121S mutant. This network is absent in apo-SmtB (Eicken et al. 2003) and not studied in $SmtB_2Zn_4$. These stabilisation effects may cause $SmtB_2Zn_4$ to predominately exist in one of its more compact quaternary conformations. The removal of zinc ions destabilises these regions, increasing flexibility and allowing $SmtB_2$ to exist more often in its slightly extended conformation. At the resolution achieved by the Synapt G2 this higher proportion of the extended conformation pushes the ATD maxima to a slightly longer drift time, which produces a larger, more expanded CCS.

To investigate the stability of the different zinc stoichiometries a higher cone voltage was implemented. At higher cone voltages there were at least two resolved conformations seen within the ATD, Figure 3.28.

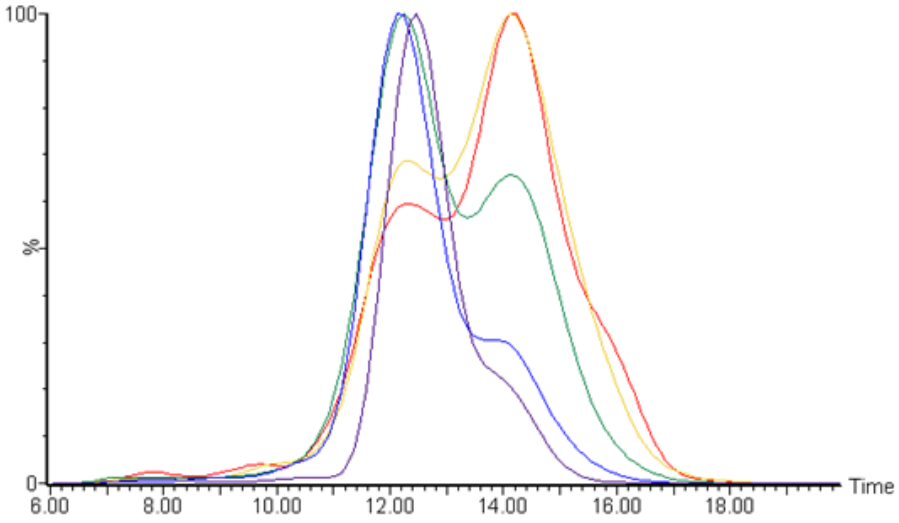


Figure 3.28: ATD's of the +11 dimeric SmtB at 60 CV. The colours represent the different stoichiometries of zinc ions to protein after the incubation with EDTA. Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red =Apo.

From the extracted ATD, observed in energy elevated conditions, the $SmtB_2Zn_4$ $[M+11H]^{11+}$ charge state was calculated to have collisional cross-sections of 2883Å² and 3114Å². Both the CCS calculated from experiments carried out at 60 V were higher than those determined at 40 V. This suggests that exposing the ions to higher

energies causes an expansion of the proteins conformation (2883 Å²), and perhaps some unfolding in the N-terminal regions (3114 Å²), which are presumed to be flexible (Eicken et al. 2003). As the zinc ions were removed from the dimer and the zinc depleted species were observed, there was a gradual increase in the intensity of the more elongated species, so that the $SmtB_2Zn$ and apo species had a higher proportion of elongated species in comparison to compact ones. The largest shifts were observed between the ATD's extracted from 2457 *m/z* and 2451 *m/z*. These correspond to the loss of a zinc ion from the species $SmtB_2Zn_3$ and $SmtB_2Zn_2$. This indicates that SmtB species with high zinc ion stoichiometries are less prone to unfolding at the same energy as zinc depleted species highlighting their involvement in dimeric stabilisation in the gas phase. The presence of bound ligands on proteins has previously been shown to provide resistance against gas phase protein unfolding during collisionally induced dissociation IMMS experiments (Hopper and Oldham 2009).

A zinc induced stabilisation effect was proposed by Kar *et al*, who observed that in the presence of zinc ions the association constant for monomer-dimer increased by two orders of magnitude (Kar et al. 1997). It has however, never been confirmed whether all bound zinc ions confer stability or whether the α3N sites act as stabilisation sites. The calculated zinc ion binding affinities for both sites are large enough for the sites to have a structural function (VanZile et al. 2002b).

In the experimental conditions used here the experimentally determined conformations of $SmtB_2$ and $SmtB_2Zn_4$ were similar, although a slight expansion of the structure did seem to occur upon zinc ion removal. A compaction was observed elsewhere upon the addition of zinc ions (Kar et al. 1997) particularly to the α5 sites (Eicken et al. 2003). The conformational change associated with the binding of zinc to the SmtB:DNA complex has, however, never been studied. The structure of apo-$SmtB_2$ determined from X-ray crystallography, appears to exist in a flat conformation. This conformation should not provide a good interaction with DNA (Arunkumar et al. 2009). To assist in forming a DNA interaction it is thought that apo-SmtB maintains a high flexibility within its quaternary structure. A flat conformation was also seen in the crystal structure of apo-CrzA, an homologous protein to SmtB, but a more compact closed form was observed, by NMR and MD

simulations, within the CzrA:DNA complex (Arunkumar et al. 2009) (Chakravorty et al. 2012). This observation suggests that the conformation adopted by apo-SmtB, not in a complex with DNA, could be significantly different to its DNA bound form. In order to measure the CCS of the $SmtB_2$:DNA complex this would first have to be transferred successfully into the gas phase (Chapter Four).

The titration and dialysis results presented here indicate that four zinc ions can bind to SmtB simultaneously, and that their removal causes a transition from an almost completely dimeric equilibrium to a predominantly monomeric one. The loss of the first two zinc ions, presumed to be from the α5 sites due to their lower affinity, slightly broadened the initial tight charge state envelope of the fully loaded SmtB species. This expansion of the charge state envelope to include more highly charged dimers suggested a small conformational change to the protein, and may represent the conformational change associated with zinc sensing. The comparison of arrival time distributions from the most biologically relevant +11 dimeric charge states revealed that this conformational change was shown to be a very slight expansion. No truly comparable expansion in CCS could be determined from structural coordinates due to the incomplete nature of all the dimeric SmtB X-ray crystal structures. The loss of the final two zinc ions, this time from the higher affinity α3N sites, caused a larger shift towards highly charged dimers and monomeric species. The percentage of monomer in the absence of Zn has been shown to increase, due to a reduction in the calculated dimeric association constant, but not to the same extent as observed in the mass spectrum. The hydrophobic interactions which predominantly maintain the apo-SmtB dimer in the solution phase are likely to be lost in the gas phase. This significant reduction of non covalent interactions between the protomer interfaces would cause a high percentage of dimeric unfolding and dissociation, indicating that zinc ions are required for dimeric stability in the gas phase. Increasing the energy under which the experiments were conducted also revealed that occupation of the zinc binding sites strengthened the dimeric structure and reduced SmtB's susceptibility to gas phase unfolding.

3.3.4.6 Tandem mass spectrometry

Collisionally induced dissociation

To investigate the potential dissociation of SmtB dimers into monomeric species and to observe if higher collision energies could result in zinc ion removal, CID experiments were carried out on SmtB, Figure 3.29.

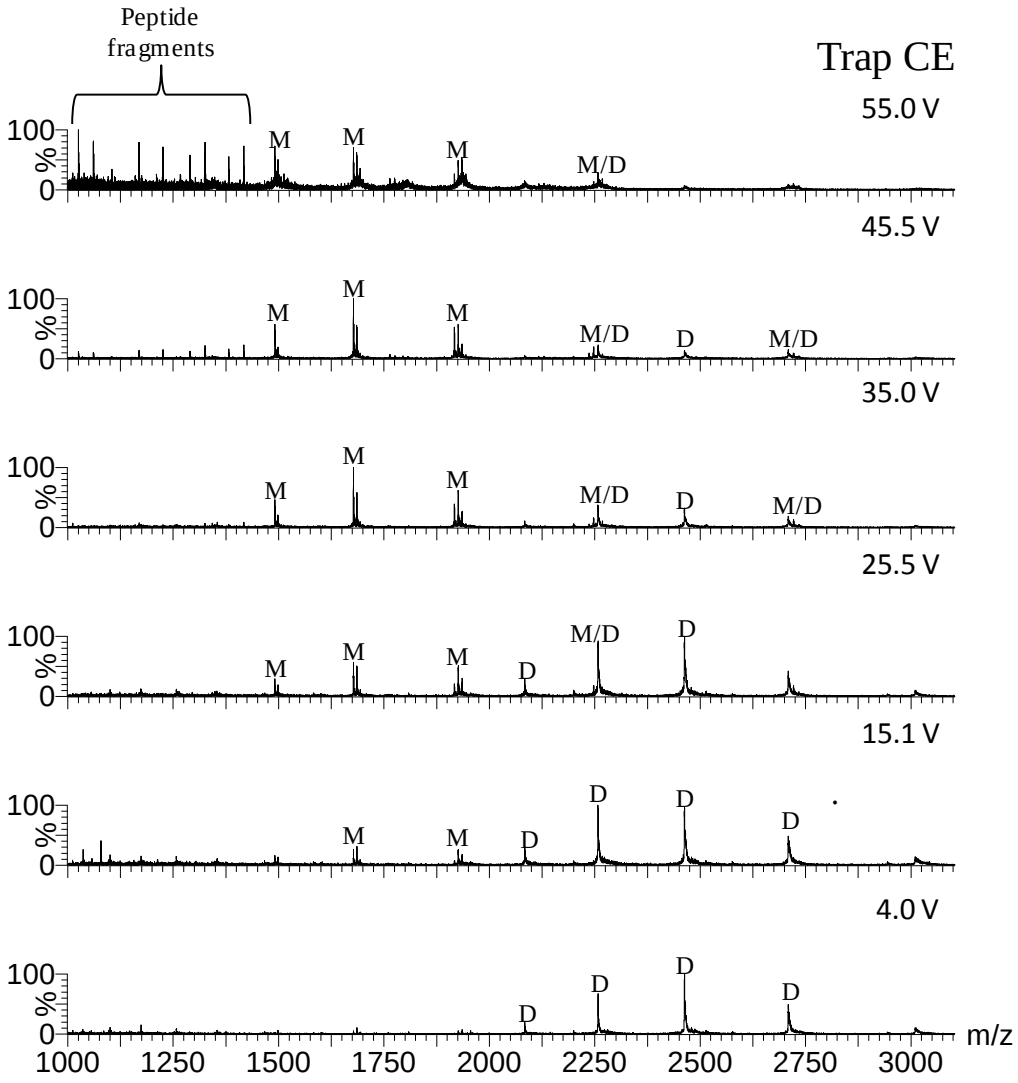


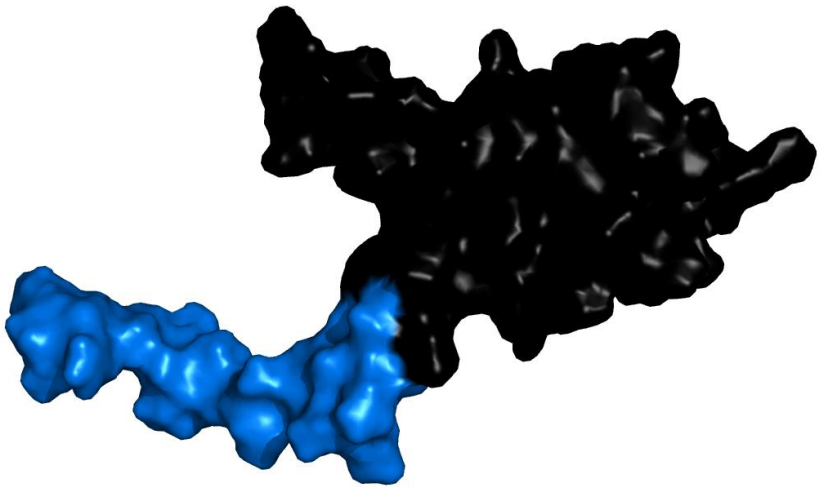
Figure 3.29: CID product ion spectra from SmtB sample in 10 mM ammonium bicarbonate
Monomeric (M) and dimeric (D) charge states are labelled.

The initial sample was largely found to be dimeric but did contain a small amount of monomer. The observed speciation is suggested to resemble the equilibrium found in

solution. A gradual increase in trap collision energy (CE) from 4 V to 35 V resulted in the dissociation of the homodimeric complex into monomeric species. These were predominately the monomeric charge states $[M+9H]^{9+}$, $[M+8H]^{8+}$ and $[M+7H]^{7+}$ with smaller contributions from $[M+6H]^{6+}$ and $[M+5H]^{5+}$. The absence of highly charged monomeric species confirmed that those species observed in the EDTA time course, EDTA and pH titration were either the result of unfolded monomers in solution or the gas phase dissociation of highly charged dimeric species. Above a trap CE of 35 V dimeric species continued to break down into monomer but peaks suspected to correspond to peptide fragments started to become more prominent, Figure 3.29.

The higher intensity fragment ion peaks were deconvoluted and submitted to the FindPept tool found on ExPASy (Gattiker et al. 2002; Gasteiger E. 2005). This tool can identify peptides formed by unspecific cleavage but has also been used in conjunction with CID experiments, in particular to identify y ions (Armirotti et al. 2009). The majority of fragments entered into FindPept were identified as y ions belonging to SmtB. Unidentified ions were then presumed to be b or a ions and were converted to y ions by the addition of water (18.01 Da) or formic acid (46.00 Da) respectively. These assumptions resulted in the identification of 14 out of 15 masses submitted to FindPept. The final mass was identified as an x ion and converted into a y ion by the removal of CO followed by the addition of H_2(25.98 Da), Table 3.4. The resulting 21 possible peptide fragments identified covered the majority of the amino acid sequence (94 %). Just over 38 % of the fragments are derived from the first 37 residues of the N-terminal region of the protein. The first 25 residues of SmtB are known to be disordered / flexible and are often missing from X-ray structures. Here the formation of monomeric species, from dimeric SmtB, exposes this N-terminal region increasing the likelihood of fragmentation in this region, Figure 3.30. The high sequence coverage obtained indicates that many of the fragment ions observed could contain partial binding sites, Table 3.3, although none were identified still associated with zinc ions.

A



B

MTKPVLQDGETVVCQGTHAAIASELQAIAPEVAQSLAEFFA
VLADPNRLRLLSLLARSELCVGDLAQAIGVSESAVSHQLR
SLRNLRLVSYRKQGRHVYYQLQDHHIVALYQNALDHLQECR

Figure 3.30: Monomeric SmtB structure (A) and sequence (B)
The areas highlighted blue show the N-terminal region of the monomeric protein which contributed to approximately 38 % of the fragments obtained in CID, note that residues 2-11 are missing from the structure. Areas of the sequence not covered by fragment peptides are underlined, the N-terminal methionine (highlighted in bold) is cleaved during bacterial expression and the residues making up the proposed zinc binding sites are highlighted in red.

Even at the highest collision energy employed (55 V), where monomeric ions themselves began to fragment, removal of zinc ions was not observed from the remaining homodimeric species. The product ion peaks which proceed the fully loaded $[M+12H]^{12+}$ SmtB peaks, appear to show possible zinc ion removal at higher trap CE voltages. These peaks were only observed in even charge state peak clusters, Figure 3.29. These peaks have therefore been attributed to lower charged, zinc depleted, monomeric peaks.

Table 3.4: The peptides produced from CID and the position of those peptides within the SmtB sequence.
Residues involved in zinc ion binding are highlighted in red. All non-y type ions were converted to y-type ions as described elsewhere (Armirotti et al. 2009) and are indicated by a *.

Observed Mass $[M+H]^+$ (Da)	Predicted Mass $[M+H]^+$ (Da)	Δ Mass (Da)	Peptide	Position
2067.1*	2067.54	-0.046	TKPVLQDGETVVCQGTHAAI(A)	2-21
2120.1	2120.091	-0.009	(A)SELQAIAPEVAQSLAEFFAV(L)	23-42
2207.1	2207.122	0.012	(V)VCQGTHAAIASELQAIAPEVAQ(S)	13-34
2294.1	2294.144	0.044	(V)VCQGTHAAIASELQAIAPEVAQS(L)	13-35
2294.1	2294.221	0.121	(V)GDLAQAIGVSESAVSHQLRS LR(N)	63-84
2336.2	2336.285	0.084	(L)RNLRLVSYRKQGRHVYYQ(L)	85-102
2392.2*	2392.317	0.116	(L)AQAIGVSESAVSHQLRSLRN LR(L)	66-87
2449.3	2449.369	0.068	(S)LRNLRLVSYRKQGRHVYYQ(L)	83-101
2449.3	2449.369	0.068	(L)RNLRLVSYRKQGRHVYYQL(Q)	84-102
2506.3	2506.297	-0.003	(T)VVCQGTHAAIASELQAIAPEVAQSL(A)	12-36
2577.3	2577.263	-0.036	(R)HVYYQLQDHHIVALYQNALDH(L)	97-117
2577.3	2577.334	0.034	(T)VVCQGTHAAIASELQAIAPE VAQSLA(E)	12-37
2577.3	2577.427	0.127	(L)RNLRLVSYRKQGRHVYYQLQ(D)	84-103
2603.4	2603.360	-0.039	(A)IASELQAIAPEVAQSLAEFF AVLAD(P)	21-45
2666.41*	2666.345	-0.064	TKPVLQDGETVVCQGTHAAI ASELQA(I)	2-27
2666.41*	2666.359	-0.051	(K)QGRHVYYQLQDHHIVALYQNAL(D)	94-115
2713.31*	2713.492	0.181	(R)LRLLSLLARSELCVGDLAQA IGVSES(A)	49-74
2779.51*	2779.429	-0.080	TKPVLQDGETVVCQGTHAAIASELQAI(A)	2-28
2779.51*	2779.581	0.071	(S)HQLRSLRNLRLVSYRKQGRH VY(Y)	78-99
2821.51*	2821.404	-0.106	(V)LQDGETVVCQGTHAAIASEL QAIAPEVA(Q)	6-33
2832.5	2832.580	0.080	(G)DLAQAIGVSESAVSHQLRSL RNLRLV(S)	64-89

The use of higher collision energies did seem to favour the formation of apo-SmtB monomers, Figure 3.29. The formation of apo-SmtB monomers could either result from zinc ions being removed directly from the monomeric species SmtBZn and $SmtBZn_2$, or result from the direct production of apo-SmtB from holo-dimeric dissociation. The appearance of lower charged monomeric species $[M+5H]^{5+}$ with up to four zinc ions bound suggested that the direct formation of apo-SmtB can proceed, but requires the use of higher energies.

The extraction of the ATD's from the dimeric $[M+11H]^{11+}$ SmtB peak showed that exposure to higher collision energies changed the conformation of SmtB, Figure 3.31. The equilibrium between the two partially resolved conformations seen at the lower collision energy (4 V) was pushed toward the extended conformation, as the collision energy was increased. This shift in equilibrium was observed until the $[M+11H]^{11+}$ charge state existed solely as the larger of the two conformations.

There is an observed relationship between the proportion of extended conformation present in the ATD, Figure 3.31, and the percentage of monomeric species observed in the mass spectrum, Figure 3.29. As the proportion of extended conformation increases the percentage of observable monomeric forms also increases. The more extended conformation in the ATD is therefore proposed to belong to a partially unfolded SmtB dimeric species, which may be the step before dimeric dissociation. A species which occupied a very similar conformational space was observed in partially loaded or apo-dimeric SmtB species, utilising a cone voltage of 60 V and a collision energy of 4 V, Figure 3.25.

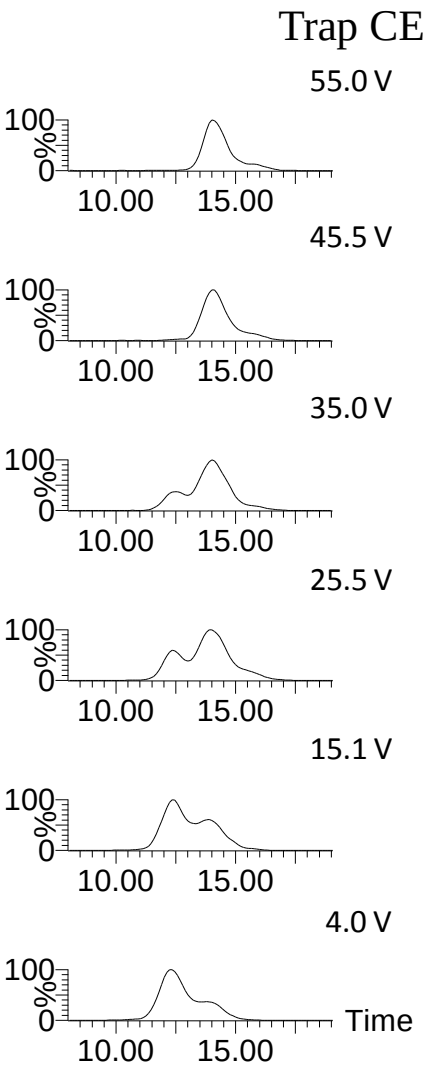


Figure 3.31: ATD's of $[M+11H]^{11+}$ SmtB upon an increase in Trap CE during CID experiments.

Electron Transfer Dissociation

To investigate whether the location of the four zinc binding sites, found within purified recombinant SmtB, could be determined, a top down ETD experiment was performed. Determining the metal binding sites of proteins by MS/MS has proved difficult without modification (Kaltashov et al. 2006). There have been some instances however, where CID has been successfully used to determine metal binding regions. Metal ions have been shown to be retained after protein fragmentation on an 80 residue Cu binding protein (Erales et al. 2009) and supercharged carbonic anhydrase (Yin and Loo 2011).

Dimeric ETD

Here the +13 $SmtB_2Zn_4$ charge state was chosen to be fragmented, by ETD, Figure 3.32. Since this charge state was solely dimeric and highly charged it was chosen as the most likely charge state to produce fragment ions retaining a zinc ion. The native dimeric SmtB species did not fragment well using ETD and the sequence coverage achieved was minimal. The fragments produced were identified as belonging to the N-terminal region of SmtB. These regions are likely to be readily accessible to the flouranthene radical anions, whilst cleavage in the remainder of SmtB was more difficult due to the presence of secondary and tertiary protein structure. This is similar to the observation that large proteins tend to only fragment at the ends of the polypeptide chain due to the increased presence of intermolecular interaction after solvent removal from denaturing conditions (Han et al. 2006).

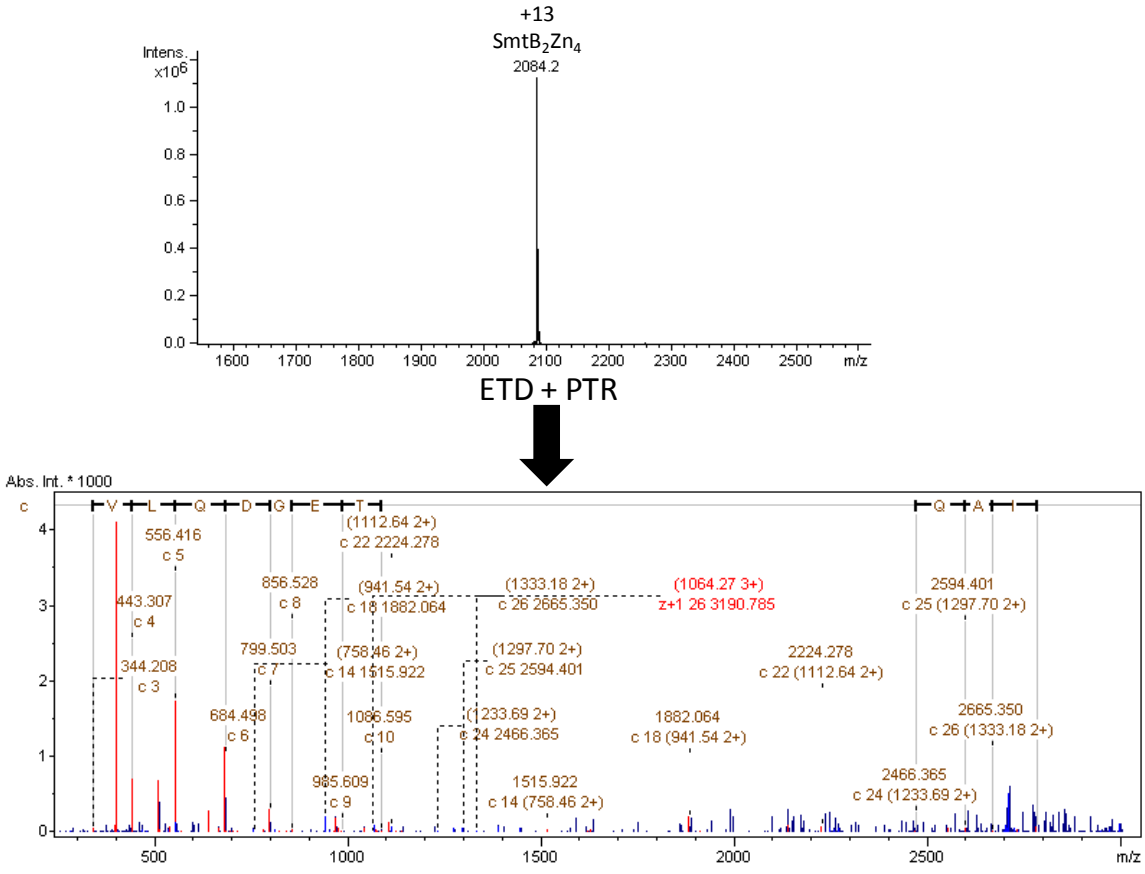


Figure 3.32: Isolation of the +13 $SmtB_2Zn_4$ species in the trap, followed by electron transfer dissociation and proton transfer reaction yielding a product ion spectrum.

Of the c ion fragments produced containing the partial α3N site (residues C14 and H18) none were found to retain a single zinc ion. ETD fragmentation experiments were also carried out using the +12 dimeric zinc bound charge state, since ETD has been shown to be charge state dependant (Rozman and Gaskell 2011). These experiments were also unsuccessful at producing a wide range of fragment ions, covering the proposed zinc binding regions. Poor sequence coverage has previously been encountered by Loo *et al.* in their attempt to locate the zinc binding sites within native carbonic anhydrase using CAD. It was not until they supercharged the protein using 0.5 % m-NBA that they achieved sequence coverage over the proposed zinc binding region and a 192 aa. peptide fragment retaining a zinc ion (Yin and Loo 2011). Attempts would therefore be undertaken to supercharge SmtB dimer in any future work, since the addition of further charges enhances both the efficiency of ETD and the production of metal bound fragments in CAD/ETC (Yin and Loo 2011).

Monomeric ETD

The presence of highly charged monomeric species in the mass spectrum suggests the presence of unfolded monomeric species in addition to more folded dimers. ETD experiments on these highly charged and more flexible species may yield higher protein sequence coverage. The monomeric species only contain partial binding sites and the production of fragment ions retaining a bound zinc would be less likely than from a complete binding site. The heightened strength of electrostatic interactions in the gas phase, relative to the solution phase, may assist in the formation of zinc bound fragments.

The +15 monomeric SmtB charge state was chosen, in its apo form and with two zinc ions bound, 895.2 *m/z* and 903.7 *m/z* respectively. These charge states were isolated in the trap and then exposed to ETD and PTR. The monomeric species, both in the presence and absence of bound zinc ions did fragment more efficiently than the dimeric species, due to their higher charge and more unfolded nature. The predominant species observed were still charge reduced monomers produced from the PTR reaction on non-fragmented monomers, Figure 3.33.

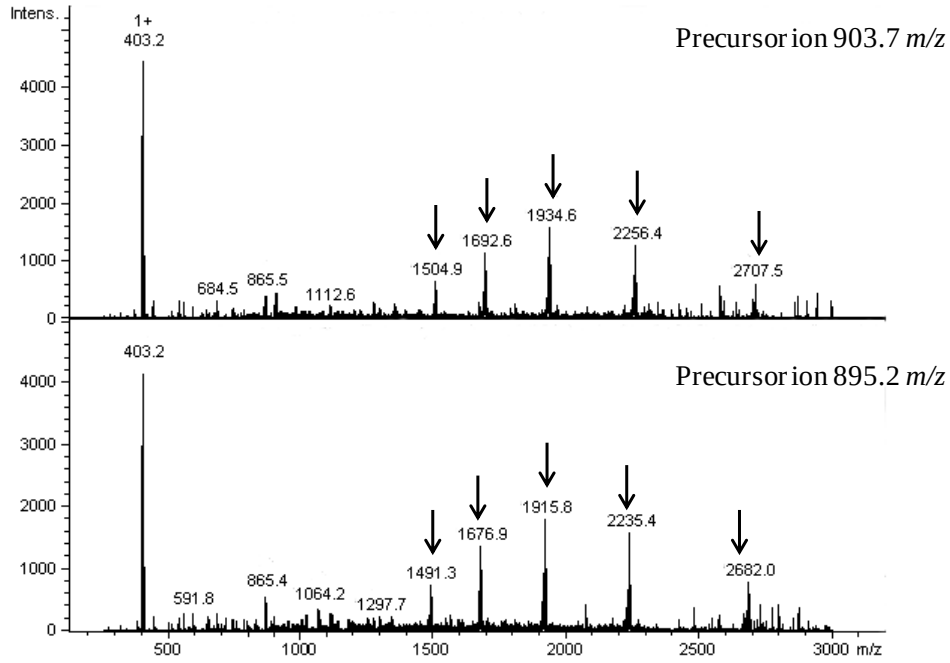


Figure 3.33: Tandem mass spectra after ETD and PTR reactions on the SmtBZn2 and SmtB +15 charge states, precursor ions were 903.7 and 895.2 *m/z* respectively. Arrows highlight the charge reduced +5 to +9 monomeric species, present in both spectra.

In general the spectra produced from the zinc bound and apo precursor ions were similar. On closer inspection the monomeric species with no zinc ions bound produced more fragments than its zinc bound counterpart. This implies that the zinc ion may be involved in the fragmentation process. The binding of metal ions, including copper and zinc have been shown to reduce the fragmentation efficiency of CID (Erales et al. 2009) (Loo et al. 1994) and ECD (van der Burgt et al. 2009).

The apo precursor charge state produced N-terminal C ion fragments, highlighting the flexibility of the first 30 residues in the SmtB monomer. In addition z+1 ions from the C-terminus were also produced with cleavage often occurring in the beta-sheet or the α5 helix, Figure 3.34. The fragmentation observed within the beta-sheet region reduces significantly in the presence of bound zinc ions, Figure 3.34. The beta sheet region of the homologous protein CzrA is thought to be more flexible in the absence of zinc and these results may indicate that the binding of zinc ions to SmtB could stabilise these areas in a similar fashion.

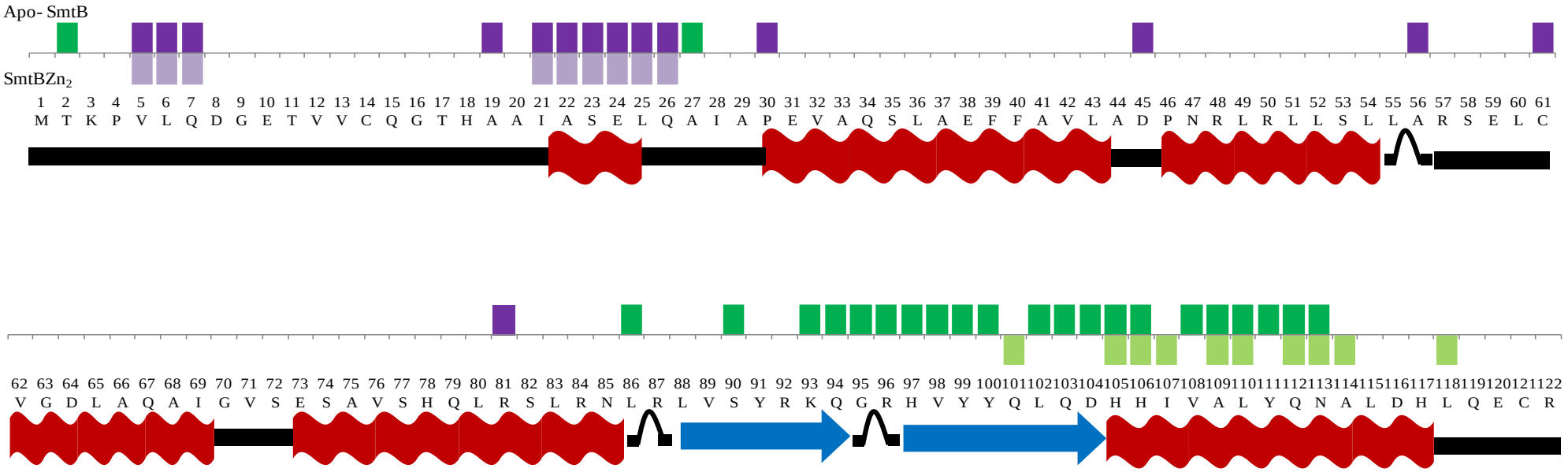


Figure 3.34: A comparison of ETD fragments between apo-SmtB and $SmtBZn_2$ +15 monomers. The positions of observed c (purple) and z+1 (green) fragments within the monomeric sequence (residues 1-122) are displayed. Secondary structural elements of the monomer are highlighted, including unstructured (black line), turns (black curves), alpha helices (red waves) and beta sheets (blue arrows).

The increased fragmentation observed in both monomers within the C-terminal α5 helix, Figure 3.34 compared to the same region within the dimer, Figure 3.32 possibly relates to the level of access achieved by the ETD reagent. In a folded dimer the α5 helices form part of the dimeric interface and are therefore highly involved in tertiary structure, resulting in low access and few to no fragments. In the absence of the second protomer, i.e. in a monomeric structure, the structural hindrance is reduced and so fragments may be observed.

The use of BioTools to search for modifications, in this case the addition of zinc ions, to the SmtB monomer was unsuccessful with the probability score calculated either the same or lower than that for the apo-SmtB sequence, Table 3.5. This indicates that there is a higher probability that the fragments obtained from ETD were apo and not zinc bound. The lack of sequence coverage however is a major factor with all the scores obtained being low.

Table 3.5: Probability scores produced in BioTools after zinc ion modification to known and predicated zinc binding residues.

Modification Residue	Binding site	Score
None	-	88
C14	α3N	24
C18	α3N	24
C61	α3N	88
D64	α3N	88
D104	α5	88
H106	α5	84
H117	α5	72
E120	α5	71

Due to the low sequence coverage obtained by computer fitting the search for metal bound fragments was also conducted manually looking for the addition of one or two zinc ions to each fragment known to contain a partial zinc ion binding site. Of the 17 identified fragments containing a zinc ligating residue four possible zinc bound

fragments were identified. These were $Z+1_{22}+Zn$, $Z+1_{10}+Zn$, $Z+1_9+Zn$, and $C_{25}+Zn$, Figure 3.35 and Figure 3.36.

The observation of zinc bound to z+1 fragments, which make up the C-terminus of monomeric SmtB, although initially surprising, agreed well with the fact that zinc was bound to the α5 zinc binding site. X-ray crystallography studies have identified the zinc binding ligands as H117, E120 from one monomer and H106 and D104 from the opposing monomer (Eicken et al. 2003). All the identified zinc bound fragments contain either all four of these residues, such as $Z+1_{22}$ or at least a partial α5 zinc binding site, $Z+1_9$ and $Z+1_{10}$. These results therefore agree well with the previous X-ray data and suggest that the isolated $SmtBZn_2$ ions did in fact have zinc ions bound in this region.

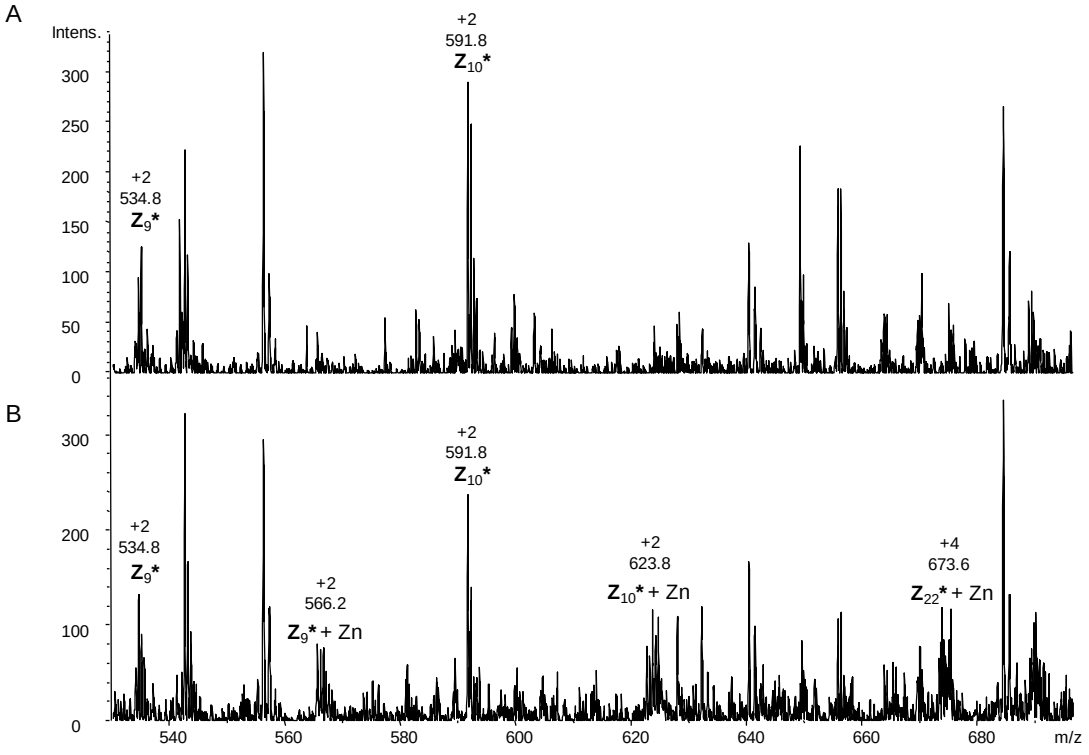


Figure 3.35: ETD/PTR spectrum of the +15 charge state of A) apo-SmtB 895.3 *m/z* and B) $SmtBZn_2$ 903.8 *m/z*. Labels indicate apo or zinc loaded Z* (z+1) fragments.

In these experiments although a number of fragments of the N-terminus were identified only a single c ion was observed to have a possible zinc ion attached,

Figure 3.36. In this instance the C61 and D64 residues area are buried too deeply within the proteins secondary and tertiary structure to be able to obtain sufficient fragmentation containing this region. The observed C_{25} fragment does however contain a partial α3N site with both the C14 and H18 residues. A C_{25}+Zn fragment is observed in the $SmtBZn_2$ spectrum, allowing the possibility of the involvement of C14 and His18 in zinc ion ligation. The signal to noise ratio of the C_{25}+Zn peak is however low and longer acquisition times or further instrument parameter optimisation would be required to confirm its existence.

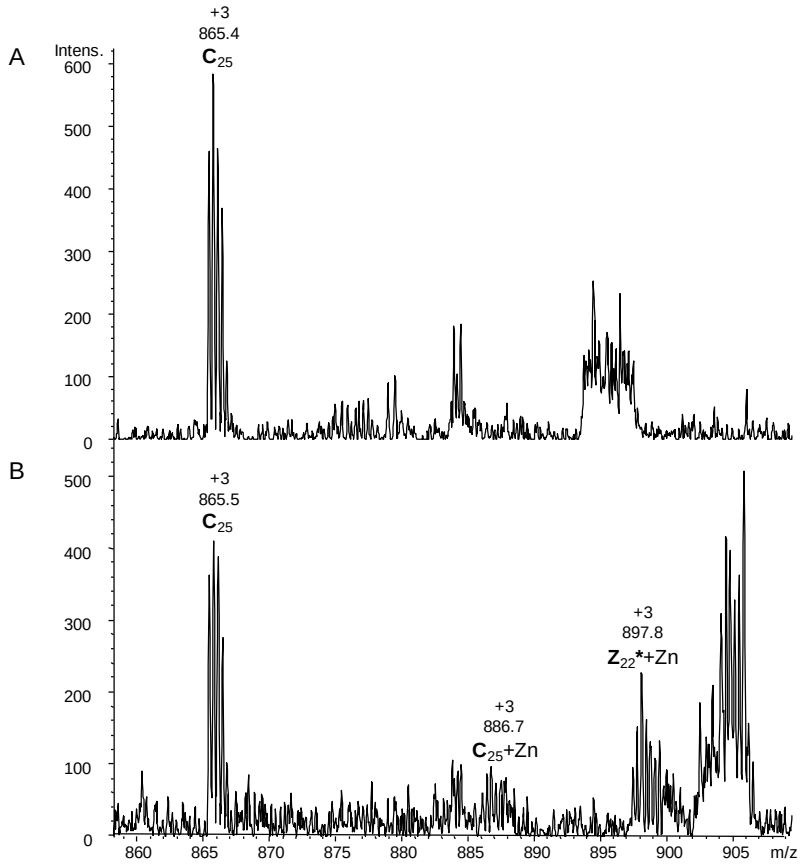


Figure 3.36: ETD/PTR spectrum of the +15 charge state of A) apo-SmtB 895.3 *m/z* and B) $SmtBZn_2$ 903.8 *m/z*.
Labels indicate apo and zinc loaded Z* (z+1) or C fragments.

The natural abundance of zinc isotopes; ^{64}Zn (48.6%), ^{66}Zn (27.9%), ^{67}Zn (4.1%), ^{68}Zn (18.8%) and ^{70}Zn (0.6%), can give zinc containing peptides an observably distinct isotopic distribution. These distrubtions can therefore be used to validate the presence of zinc in small proteins (Wortmann et al. 2005) or peptides produced by tandem MS (Keltner et al. 2010). Comparisons between the observed isotopic

distributions of the four possible SmtB zinc bound fragments and simulations of these isotopic distributions would assist in the positive identification of these fragments. The maximum resolution of the AmaZon speed ETD ion trap (20,000) is high enough to distinguish between these species although some of the peaks have low signal to noise, in particular the C25+Zn fragment, which may complicate assignments.

These results not only highlight the potential involvement of particular regions for zinc ion binding but may provide, after further experiments, additional evidence suggesting that the α3N site can be simultaneously occupied, along with the α5 site, when SmtB is expressed in the presence of zinc ions in the bacterial cytosol. The observed occupation of the α5 sites in free SmtB is not consistent with published literature (VanZile et al. 2002b).

3.3.5 Re-addition of zinc ions to apo-SmtB

In order to investigate whether the addition of zinc ions to the $SmtB_2$:DNA complex disrupts the preformed complex (VanZile et al. 2002a) or causes a conformational change (Kar et al. 2001) the addition of zinc ions to apo-SmtB was investigated.

The mass spectrum of apo-SmtB, in 10 mM ammonium acetate, produced from dialysis against EDTA under anaerobic conditions is comparable to that of the samples prepared in ammonium bicarbonate. A low percentage of dimeric apo-SmtB species is observed along with a higher intensity of folded and unfolded monomeric species, Figure 3.37A. Upon deconvolution the observed masses of both monomeric (13414 Da) and dimeric (26827 Da) compare well will their expected masses, 13413.3 Da and 26826.6 Da respectively, Figure 3.38.

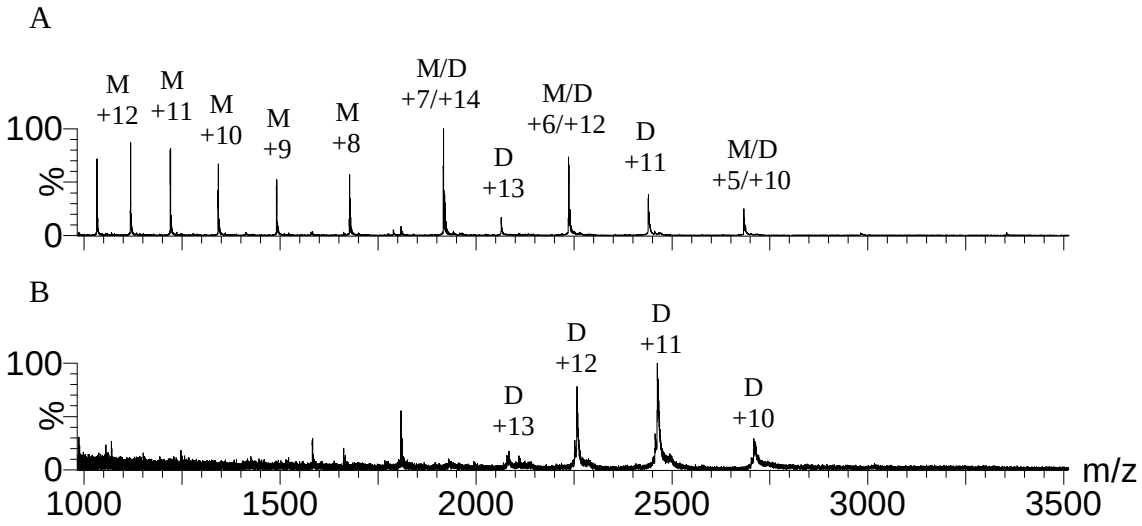


Figure 3.37: Mass spectrum of A) apo-SmtB in 10 mM ammonium acetate pH 7.4 and B) apo-SmtB incubated for 90 minutes with excess zinc acetate.
Charge states corresponding to dimer, monomer are labelled

Previous work involving apo-SmtB has highlighted the difficulties associated with the re-addition of zinc ions. Oxidation of the six Cys residues can lead to cross linking of the dimers. If oxidation occurs at the α3N site, the formation of a disulphide bond would prevent the re-addition of zinc ions to the protein. The mass difference between the expected and observed masses, of the apo species, is within the normal experimental error observed in these experiments. This indicates oxidation of the Cys residues has not occurred to a significant extent. To help prevent oxidation of a protein sample reducing agents, such as DTT, are often added. DTT, in addition to any remaining EDTA, documented to be notoriously difficult to remove from SmtB (VanZile et al. 2000), will compete for the chelation of zinc ions (Krezel et al. 2001). The deconvoluted data, Figure 3.38, indicates that sufficient steps were undertaken to remove excess EDTA, which can lead to protein:EDTA adducts (+ 292 Da). It does not however prove that all the EDTA has been taken out of solution. The addition of a stoichiometric concentration of zinc ions, therefore, is still likely to under populate the SmtB binding sites. To overcome this problem an excess of zinc ions could be added, allowing complete occupancy of the zinc binding sites.

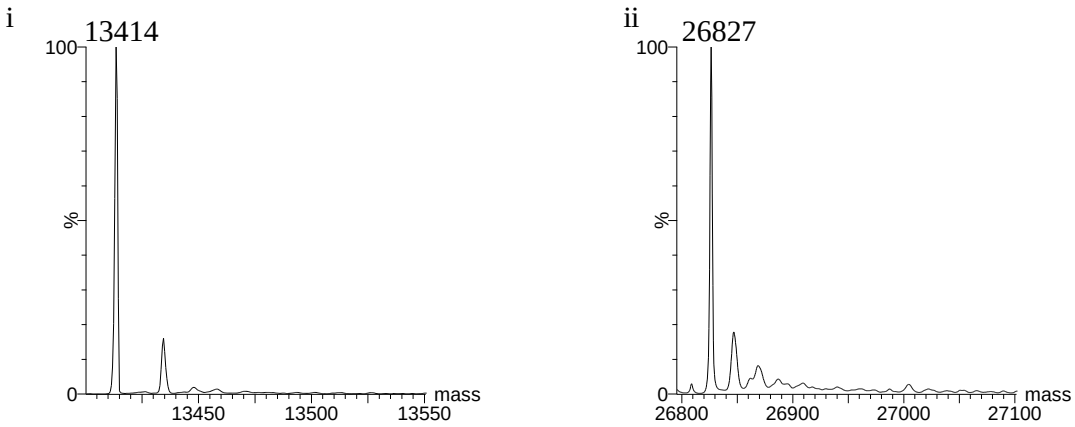


Figure 3.38: Deconvoluted spectra of i) monomeric and ii) dimeric apo-SmtB
The masses of the main peaks are labelled in Daltons, whereas subsequent peaks correspond to Na^+ adducts.

The addition of a ~5 fold excess of zinc ions followed by a 90 minute incubation period resulted in a completely different mass spectrum from the apo-SmtB mass spectrum, Figure 3.37B. The charge state envelope within this spectrum resembles the $SmtB_2Zn_4$ spectrum and shows tight distribution of dimeric SmtB species, with no obvious peaks corresponding to monomers. Deconvolution of this data, Figure 3.39, reveals that the majority of species within the spectrum are $SmtB_2Zn_4$ (27080.1 Da) and $SmtB_2Zn_3$ (27016.7 Da), with a small amount of $SmtB_2Zn_2$ (27953.4 Da). The observed masses correspond well with the expected masses, except in the case of $SmtB_2Zn_2$. In this case there is a mass difference of -2.4 Da, which may indicate that some sample oxidation has occurred. The ability to re-add up to four zinc ions to apo-SmtB suggests that, not only is $SmtB_2Zn_4$ created in the bacterial cytosol, but it can also be created in vitro. The absence of a $SmtB_2Zn_5$ species suggests that the binding is specific and the mass spectrum is comparable to the solution phase and not an artefact of the electrospray process.

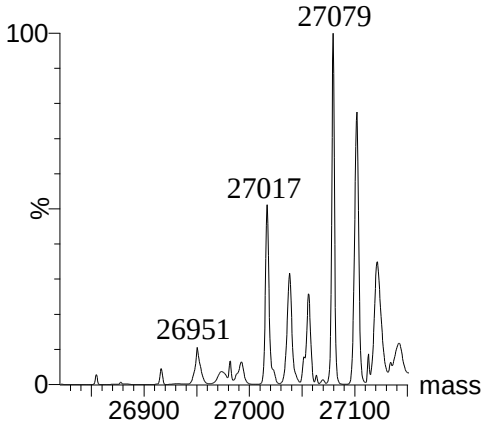


Figure 3.39: Deconvoluted spectra of apo-SmtB incubated with zinc acetate.
The masses of the main peaks are labelled in daltons, whereas subsequent peaks correspond to salt adducts.

The addition of an excess of zinc ions at low protein concentrations has been documented to result in SmtB precipitation (VanZile et al. 2000). In this case, although no obvious particulates were formed in solution, upon Zn^{2+} addition, the signal to noise intensity observed was lower. This may be due to the zinc acetate itself as the addition of zinc ions has previously been shown to reduce signal intensity, at zinc concentrations as low as ~ 50 µM, even within peptide samples where no zinc binding occurs (Wortmann et al. 2005). Protein precipitation can however not be completely ruled out. A lower observed protein concentration, in this instance, does not affect the result. If the addition of zinc ions to the SmtB:DNA complex was to be further investigated efforts would need to be undertaken to ensure an excess of Zn^{2+} was not used. This is important since if the protein precipitates out upon Zn^{2+} addition it will appear experimentally as if the SmtB:DNA complex has degraded.

3.4 Conclusions

Electrospray mass spectrometry has long been considered a robust method for determining protein:ligand stoichiometry within protein complexes. The work presented here has highlighted the use of native MS in determining metal ion:protein stoichiometries, especially if the protein in question exists at equilibrium between states with varying metal binding capacities. The implementation of ion mobility mass spectrometry experiments allows the conformation of proteins to be probed at different metal binding stoichiometries providing complementary data to other structural techniques, in particular X-ray crystallography and NMR.

The SmtB protein expressed in media supplemented with added zinc ions has been shown to exist predominantly as $SmtB_2Zn_4$. This was the only observed dimeric species, however monomeric species corresponding to apo-SmtB, SmtBZn and $SmtBZn_2$ were also observed in low abundance, consistent with solution phase data. ICP-OES experiments supported the gas phase measurements which argue that all four sites in SmtB have the capacity to bind zinc ions simultaneously. This stoichiometry was also confirmed in vitro upon the addition of zinc ions to apo-SmtB samples. SmtB expressed in the presence of added cadmium ions was not found in a cadmium bound species but instead as $SmtB_2Zn_4$ at much lower yields, which not only emphasises the strong zinc binding ability of SmtB, but suggests that zinc may have a role in SmtB stability.

The removal of zinc ions proceeded slowly during EDTA time course and titration experiments, however all eight species of SmtB were observed in both instances. The equilibrium between monomer and dimer was also noted to change upon zinc ion removal, with more monomeric species observed. This transition towards monomeric species was consistent with a reduction in the dimer association constant previously observed. This transition was shown to occur at a greater degree than was observed in the solution phase. Samples exposed to larger cone voltages and so higher energies were shown to dissociate into monomer more readily. This appears to be the result of partially unfolded dimeric species dissociating and not due to a

specific zinc depleted species of folded SmtB being more unstable than higher loaded species.

In order to investigate the apo-SmtB species more clearly the production of a pure apo-SmtB sample was required. This was carried out by performing dialysis, under anaerobic conditions, against EDTA on the stable $SmtB_2Zn_4$ protein. The resulting protein sample was highly monomeric, with only a small percentage belonging to dimeric apo-SmtB. This dramatic shift in equilibrium towards monomeric species was proposed to be caused by the loss of hydrophobic interactions in the gas phase, the predominant non-covalent interaction found to exist at the dimeric interface. Dialysis performed in an oxygenated environment produced a higher proportion of dimeric apo-SmtB but this appeared to be due to the formation of stabilising inter-monomer disulphide bonds.

The EDTA course revealed that the kinetics of zinc ion removal proceeded slowly. To investigate the thermodynamic stability of the two sets of zinc binding sites an EDTA titration was conducted. The subsequent data showed that the removal of zinc ions occurred in a sequential fashion and produced no evidence to support the idea of cooperativity or a particularly stable SmtB species, as had been previously thought. This work indicated that the two separate sets of binding sites, α5 and 3αN, may have closer zinc binding affinities than determined previously. A pH titration confirmed that SmtB has a strong zinc binding affinity, higher than all published values for SmtA. The observation of a build up of $SmtB_2Zn_2$ during pH titrations showed that, although both sets of sites have a high affinity, their sensitivity to pH is slightly different, which may be a reflection of a difference in absolute Zn affinity.

Rotationally averaged collisional cross sections calculated from ion mobility experiments revealed that SmtB is a flexible protein likely to adopt a number of similar conformations in both its apo and zinc bound form. More extended conformations were observed in the absence of bound zinc ions, however the conformational change observed was extremely small compared to the changes calculated from existing crystal structures. Exposing these zinc depleted species to higher energies was observed to destabilise the dimeric structure, creating an extended conformation. This could also be achieved by exposing $SmtB_2Zn_4$ to higher

collision energies during CID experiments. The species observed are proposed to represent partially unfolded SmtB which may be in the process of dimer unfolding and dissociation.

Experiments designed to locate the zinc binding sites were carried out using tandem mass spectrometry, utilising two different methodologies, CID and ETD. CID produced fragments which contained partial zinc ion binding sites but none were identified with zinc ions bound. ETD spectra of dimeric SmtB produced little fragmentation with low sequence coverage. Highly charged monomers were also studied. Apo-SmtB was found to fragment to a higher extent in flexible regions of the structure than $SmtBZn_2$ indicating zinc ions either perturb the ETD fragmentation process or that the binding of zinc to SmtB stabilises previously flexible regions. Manual assignments identified zinc bound z+1 fragments from the C-terminus, supporting X-ray crystallography evidence that a zinc binding site is located in this region. A potential zinc bound c fragment was also observed which after further work could confirm the involvement of the N-terminus in zinc binding.

Chapter Four: The interactions of SmtB with DNA

4.1 Introduction

4.1.1 The binding of SmtB to DNA

The interaction of SmtB with DNA has been the subject of a number of studies over the last two decades (Morby et al. 1993; Erbe et al. 1995; Turner et al. 1996; Kar et al. 2001; VanZile et al. 2002a). The DNA recognition helices within SmtB have been identified, along with two separate SmtB binding sites within the operator-promoter sequence, (refer to Chapter One for details). No high resolution structure has been obtained, either by NMR or X-ray crystallography of any SmtB: DNA complex. The mechanism of SmtB binding has therefore not been fully characterised, although a number of models have been proposed.

Elucidation of the first crystal structure of apo-SmtB (Cook et al. 1998), combined with the use of methylation interference assays and bioinformatics, led to a greater understanding of possible interactions between $SmtB_2$ and DNA. Cook *et al.* superimposed the predicted binding helices of SmtB upon the known DNA binding helices of different H-T-H proteins. The use of hepatocyte nuclear factor 3 (HNF-3), as a template, presented the most convincing model in which both ends of SmtB bound within two consecutive major grooves of the DNA. This binding event was predicted to impose a 30 ° bend in the DNA and to involve the SmtB residues Cys 61 and His 97 (Cook et al. 1998).

A combination of sedimentation velocity data, interference assays (Kar et al. 2001) and fluorescence anisotropy (VanZile et al. 2002a) have since updated the initial model. The current model postulates that both the S2/S1 and S4/S3 sites have the ability to bind up to two SmtB dimers. It remains unclear, however, whether a SmtB tetramer is formed on the DNA or whether two SmtB dimers bind on either side of the DNA, Figure 4.1A. Kar *et al.* based their sequential binding model upon the formation of a tetramer at the S2/S1 site. This tetramer forms a bridge to the S3 site, which is only displaced after another SmtB dimer binds to the S4 site (Kar et al. 2001). Although apo-SmtB is known to form weak tetramers in solution (Kar et al. 1997), the experimental binding affinities of SmtB dimers, determined by fluorescence anisotropy, were found to increase with DNA length. This evidence

supported the proposed complex architecture of SmtB dimers binding on either side of DNA, Figure 4.1B, since the addition of base pairs would increase the interaction surface (VanZile et al. 2002a).

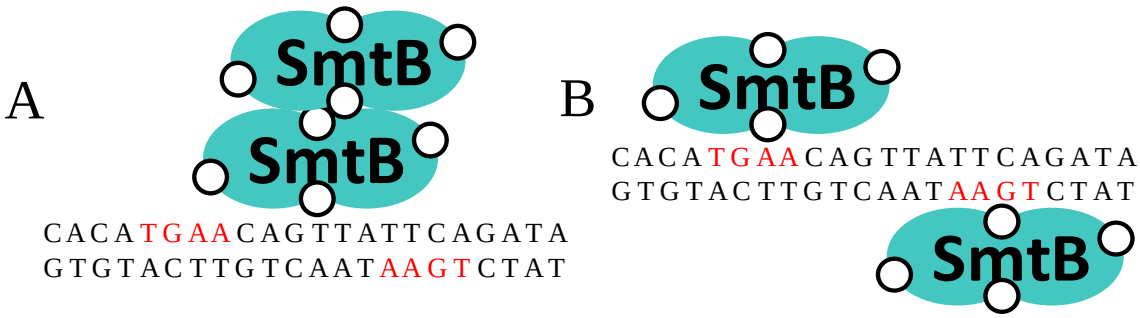


Figure 4.1: Schematic displaying the two possible SmtB: DNA complexes
A SmtB tetramer bound to the DNA (A) or two SmtB dimers on either side of the DNA (B). The 5'- TGAA motifs are shaded in red.

In this model the SmtB dimers were proposed to be centred on a 5'- TGAA motif, Figure 4.1B, in contrast with each end interacting with the sequence, Figure 4.1A. This off-centred binding has also been documented for other ArsR/SmtB members. One dimer of CadC was found to bind to its operator-promotor sequence in this fashion, centred over a 5'-TCAA motif (Endo and Silver 1995). A further dimer has more recently been discovered to bind under low NaCl conditions (Busenlehner et al. 2002). The sensing and DNA binding of CadC was found to be different to that of SmtB. Since CadC is an α3N sensor (Busenlehner et al. 2002) instead of an α5 sensor, its DNA binding 12-2-12 repeat would have a different core motif (Busenlehner et al. 2003). This implies that a different mode of DNA binding may be involved.

CzrA, a homologous protein to SmtB, has more recently been studied in a complex with a 28 bp duplex DNA using NMR (Arunkumar et al. 2009). Unlike CadC, CzrA is an α5 sensor which senses zinc ions. This allowed a more accurate comparison to be made between the CzrA:DNA and SmtB:DNA complexes. Like SmtB, two CzrA dimers (note CzrA = ZntR), were found to have the ability to bind to a single 12-2-12 inverted repeat at a high affinity (Busenlehner et al. 2003; Lee et al. 2006).

Within the NMR CzrA:DNA complex, only a 1:1 CzrA dimer:DNA stoichiometry could be achieved and the bound DNA structure remained unknown. The model produced, suggested that the interaction of the αR helices with the 5'- TGAA motifs, Figure 4.2, bent the DNA by 35 ° (Arunkumar et al. 2009). This bend in the DNA corresponded well to the original SmtB:DNA binding model proposed by Cook *et al.* (Cook et al. 1998).

Figure 4.2: Schematic displaying the CzrA: DNA complex
The 5'- TGAA motifs, thought to interact with the αR helix, are shaded in red.

Limited information regarding the CzrA:DNA interface was obtained, and due to sequence conservation, was thought to be applicable to other α5 sensors, such as SmtB. This work revealed the involvement of CzrA residues, with bracketed SmtB equivalents; Q53 (E73), V42 (V62), H58 (H78), S54 (S74) and S57 (S77) (Arunkumar et al. 2009). These residues lie predominantly within the αR helix although V62 revealed a possible contact with the α3 helix.

4.1.2 The dissociation of SmtB from DNA

An increase in zinc ion concentration is thought to lead to zinc ions binding to SmtB, within the SmtB:DNA complex, changing its conformation and leading to complex dissociation. Experiments have shown that zinc-bound SmtB has a reduced affinity for DNA (Morby et al. 1993; Erbe et al. 1995). It was not until 2001 that a SmtB: DNA complex was exposed to zinc ions, in order to study the dissociation process. In this study no dissociation was observed by DNase I footprinting or gel shift analysis, although a widening of the footprint was observed, suggesting a possible conformational change. (Kar et al. 2001). This led to the conclusion that the conformational change caused by Zn^{2+} binding could allow RNA polymerase to access the operator-promoter region without SmtB displacement. The inducement of a conformational change, without DNA displacement, was however not consistent

with research carried out on other members of the SmtB/ArsR family (Busenlehner et al. 2002). It has however been recognised in the Mer metalloregulatory family (Ma et al. 2009). SmtB:DNA complex dissociation, upon zinc ion addition, was later observed by fluorescence anisotropy (VanZile et al. 2002a). The addition of 1 molar equivalent of zinc ions (in respect to monomer concentration) was found to cause a reduction in anisotropy. These results, combined with the determined equilibrium association constants for the formation of apo and Zn-bound SmtB:DNA complexes, allowed population analysis to be conducted. This revealed a shift in equilibrium, upon zinc ion addition, from a mixture of $4SmtB_2$:DNA/$3SmtB_2$:DNA to a mix of $2SmtB_2$:DNA/$SmtB_2$:DNA. Even though only ~ 10 % free DNA was produced in this reaction, the addition of zinc ions clearly showed the dissociation of the multimeric complex.

Aims

Although the interaction of SmtB with its DNA binding sequence has been the subject of a number of studies the metalloregulator and its DNA binding site have not been studied by mass spectrometry or ion mobility mass spectrometry. No bacterial metalloregualtor's interaction with DNA has been investigated utilizing these approaches and only recently has the TWIMS technique been used to investigate conformational changes within protein:DNA complexes (Atmanene et al. 2010). The application of mass spectrometry related techniques, including MS/MS and ion mobility, could permit further study of the complexes composition and shape. This would provide complementary structural data and perhaps answer some of the remaining questions. In order to conduct these experiments, which are still relatively novel in biological mass spectrometry, the complex had to be first created and transported successfully into the gas phase. This chapter outlines a study of the SmtB:DNA complex in the gas phase, including analysis of oligonucleotide components.

4.2 Materials and Methods

4.2.1 Chemicals and Materials

All chemicals were obtained from Sigma Aldrich or Fisher Scientific unless stated otherwise. All solvents used in MS were of an analytical grade.

Single stranded oligonucleotides corresponding to the two complimentary strands of the S2/S1 binding site, Oligo1 (CAC ATG AAC AGT TAT TCA GAT A) and Oligo2 (TAT CTG AAT AAC TGT TCA TGT G) were purchased from Invitrogen. The lyophilised samples were re-dissolved in MilliQ water creating a 100 µM stock, suitable for freezing. The samples had undergone RF-HPLC and further desalting was unnecessary.

4.2.2 Oligonucleotide ion mobility mass spectrometry

Re-dissolved lyophilised samples were diluted into 10 mM ammonium bicarbonate pH 7.8 to a concentration of 5 µM. Samples were introduced into the source region via negative mode electrospray from nanoflow probe tips. Instrumentation variables were optimised to ensure a stable spray under near native conditions. Experiments were carried out with a source temperature of 80 °C, capillary voltage of 1.2 kV and a cone voltage of 30 V.

4.2.3 Oligonucleotide annealing

An aliquot of each oligonucleotide was buffer exchanged in 200 mM ammonium bicarbonate and concentrated with Amicon centrifugal filters, 3 kDa cut-off (Millipore). To produce the dsDNA, thought to interact with SmtB, a 1: 1 molar ratio was incubated at 95 °C for 5 minutes and then left to slowly cool to room temperature over a number of hours. The resulting sample was frozen for storage.

4.2.4 DNA mass spectrometry

Annealed samples were diluted to a concentration of 10 µM in 10 mM ammonium bicarbonate pH 7.8 before negative mode ESI. A low concentration (<100 µM) was used to minimise unspecific interactions. The sample was sprayed from fused silica nanospray needles in negative ion mode. Conditions were optimised to ensure high sensitivity whilst keeping the dsDNA intact. Experiments were carried out over a

mass range of 1000-3000 *m/z*, with a source temperature of 80 °C, capillary voltage of 1.2 kV and a cone voltage of 50 V.

4.2.5 Interaction of SmtB and dsDNA

In order to investigate the interaction of SmtB with the S2/S1 site, different molar ratios of apo-SmtB and dsDNA were incubated for two hours, under anaerobic conditions. A control of the fully zinc bound SmtB was also incubated with the dsDNA to ensure that specific binding was taking place.

4.2.6 Mass spectrometry of the SmtB DNA mixture

Mass spectrometry experiments were carried out on the SmtB dsDNA mixtures and the control in positive ESI mode. The mixtures were diluted to a combined concentration of 20 μM using the solvent 10 mM ammonium bicarbonate pH 7.4. Attempts to collect data in ammonium acetate and at different concentrations of ammonium bicarbonate were unsuccessful, as was the use of negative mode ESI.

Nano-ESI was employed, and the majority of conditions were retained from the DNA mass spectrometry experiment outlined earlier. A higher backing pressure of 4.5 mbar was utilised along with a cone voltage of 100 V. Measurements were taken over two different mass to charge ranges, 1500-5000 *m/z* and 3000-5000 *m/z*. This allowed the components of the mixture within the sample to be observed and then allowed sensitivity for the complex to be optimised.

4.3 Results and discussion

4.3.1 Oligonucleotides and dsDNA

A 22mer oligonucleotide containing the S2/S1 site was chosen to measure the interaction of SmtB with DNA in the gas phase. Although this system contained more than the optimal MS length of 16 bp (Beck 2011), the extra bases flanking the S2/S1 binding site were necessary to ensure SmtB binding (VanZile et al. 2002a). The use of the complete operator-promoter sequence was deemed not to be practical, due to its size (100 bp) since the level of cation adducts would have led to a significant loss in sensitivity (Beck 2011). The S4/S3 site was not initially studied, since previous work on the interaction of SmtB with the operator promoter region revealed that SmtB had a higher affinity to the S2/S1 site (Turner et al. 1996). This higher affinity would ensure lower sample concentrations could be used during the method development stages. In a biological context, the S2/S1 binding site was also an ideal candidate for the initial study of SmtB:DNA complex, since the binding of SmtB to this site was thought to control the transcription of SmtA (VanZile et al. 2002b).

4.3.2 ESI-MS of ssDNA and dsDNA

To check the molecular weight of the oligonucleotides, before annealing them to create dsDNA, near native ESI-MS was conducted in negative ion mode, Figure 4.3. The oligonucleotides all produced similar spectra, each displaying a dominant charge series between $[M-4H]^{4-}$ and $[M-6H]^{6-}$, with little or no fragmentation. In the gas phase, there is the potential for each phosphate group to be deprotonated and to provide the oligonucleotide with a negative charge. The charge states observed, Figure 4.3, were substantially less charged than the oligonucleotide's theoretical maximum charge. This reduction in charge was proposed to be due to the 10 mM ammonium bicarbonate spraying solution employed. Previous work has indicated that the presence of ammonium bicarbonate/acetate can effect oligonucleotide charge state distributions in a concentration dependent manner (Guo et al. 2005; Touboul and Zenobi 2009). It is thought that ammonium ions, within these buffers, donate protons to the oligonucleotides nitrogenous bases upon entering the gas phase (Guo et al. 2005), therefore reducing the negative charge observed.

After deconvolution onto a true mass scale, both oligonucleotide masses, 6726.5±0.5 and 6739.5±0.5 Da, were in good agreement with their predicted masses of 6726.5 and 6739.5 Da. The smaller peaks present at higher masses in the deconvolution were calculated to be in agreement with the presence of sodium adducts, Figure 4.3.

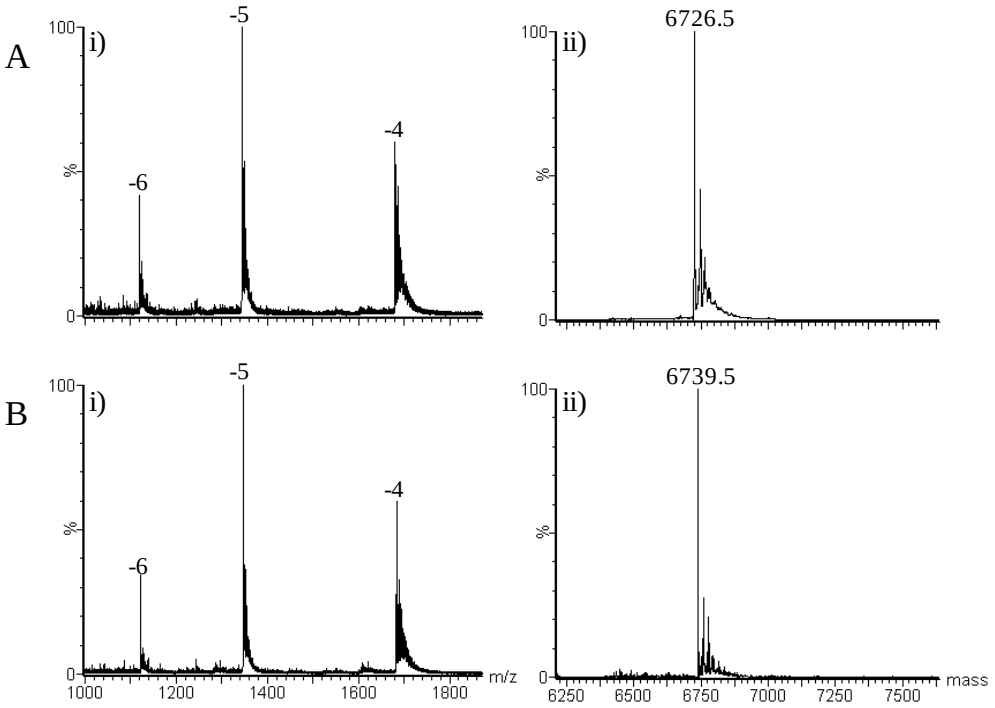


Figure 4.3: i) Negative ESI mass spectra of A) 5 µM Oligonucleotide one and B) 5µM Oligonucleotide two. ii) shows the deconvolution of the mass spectra.

Once the complementary oligonucleotides had been annealed, to produce dsDNA, containing the S2/S1 SmtB binding site, negative mode ESI-MS experiments were conducted on the oligonucleotide mixture, Figure 4.4. From the mass spectrum it can be seen that there are two different, overlapping, charge state envelopes. One was found to correspond to un-annealed ssDNA, including charge states $[M\text{-}6H]^{6-}$ to $[M\text{-}3H]^{3-}$ for both Oligo 1 and Oligo 2. The other charge state envelope corresponded to dsDNA, made up of the charge states $[M\text{-}8H]^{8-}$ to $[M\text{-}6H]^{6-}$. The overlap of charge states relating to ssDNA and dsDNA is likely to be more common than the overlap of monomeric and dimeric proteins. As the deprotonation sites within DNA are thought to be the backbone phosphate groups, (Moradian et al. 2002), the formation of a dsDNA helix is unlikely to shield the potential deprotonation sites. In proteins the residues with a potential to become charged are often found to be distributed

evenly across the proteins surface. A percentage of these, on each monomer, will therefore be covered up upon dimerisation.

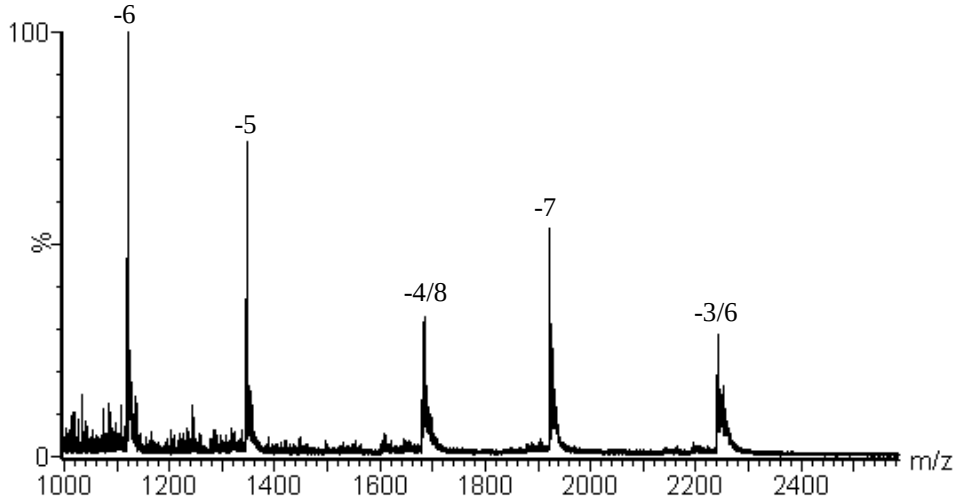


Figure 4.4: A typical ESI mass spectrum of dsDNA in negative mode, both unbound oligonucleotides and dsDNA with labelled charged states are shown.

The presence of ssDNA oligonucleotides in the annealed sample, Figure 4.4, was unsurprising, since the annealing process between two complementary nucleotides is not 100 % efficient, and no attempt was taken to purify the mixture. The annealing efficiency and stability of dsDNA can be substantially decreased when exposed to low ionic strength. High ionic strengths (100 mM +) were therefore used whenever possible, and diluted prior to MS (Gupta et al. 2001). Here the proportion of C and G bases, within the dsDNA, was found to be low (31 %); which implied a low melting temperature and stability. The breakdown of some dsDNA, into single strands is therefore likely, although attempts were taken to keep this to a minimum.

When compared to the spectra obtained from the oligonucleotide samples, Figure 4.3, a higher proportion of fragmentation in the low *m/z* range in the dsDNA sample was observed. This is thought to be due to the fragmentation of the ssDNA as the optimum cone voltage for dsDNA was slightly higher than that for ssDNA.

The oligonucleotides used to create the dsDNA had a mass difference of 13 Da and so there is no difficulty assigning *m/z* peaks for the dsDNA. These range from $[M\text{-}6H]^{6-}$ to $[M\text{-}8H]^{8-}$. If the sequences had been self-complementary, and therefore

possessed the same mass, only odd charge states associated with the duplex could have been confidently assigned.

The mass spectrum of the dsDNA charge state envelope was studied in greater detail to determine if any non-specific association was occurring between the un-annealed ssDNA forming homoduplexes. Low intensity peaks were observed bracketing some of the identified dsDNA peaks, Figure 4.5.

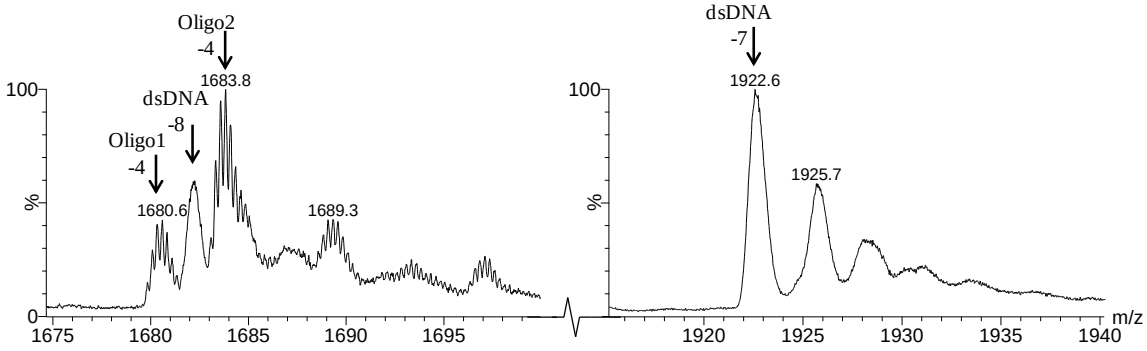


Figure 4.5: Enlarged peaks from the dsDNA charge envelope revealing the presence of Oligo1 and Oligo2, sandwiching the even dsDNA peaks only.

The higher resolution observed on these sandwiching peaks when compared to the known dsDNA peaks, indicated that these peaks were probably due to the presence of lower mass species. This assumption, combined with the observation that the peaks only occurred in conjunction with the $[M-6H]^{6-}$ and $[M-8H]^{8-}$ dsDNA peaks and not the $[M-7H]^{7-}$ dsDNA peak, led to the conclusion that these peaks were probably derived from $[M-4H]^{4-}$ and$[M-3H]^{3-}$ ssDNA oligonucleotides.

4.3.3 Ion mobility mass spectrometry of ssDNA and dsDNA

Ion mobility mass spectrometry experiments were performed on ssDNA oligonucleotides and dsDNA, to investigate their gas phase conformation. The arrival time distributions (ATD) for each charge state were extracted and compared. For both the ssDNA oligonucleotides only a single ATD was observed for each charge state, revealing the presence, under the experimental conditions employed, of a single resolvable conformation. As the number of charges increased from -4 to -6, if the conformations were of a similar collisional cross-section (CCS), a decrease in arrival time would be observed. This is due to the fact that an ion's mobility through

the IMS cell depends on mass, charge and CCS. This trend is observed, for both oligonucleotides upon the addition of one charge. The addition of further charges does not decrease the arrival time, instead a small increase is observed. This suggests that the presence of six negative charges may provide sufficient repulsion to elongate the conformation of the ssDNA, in agreement with studies elsewhere (Hoaglund et al. 1997).

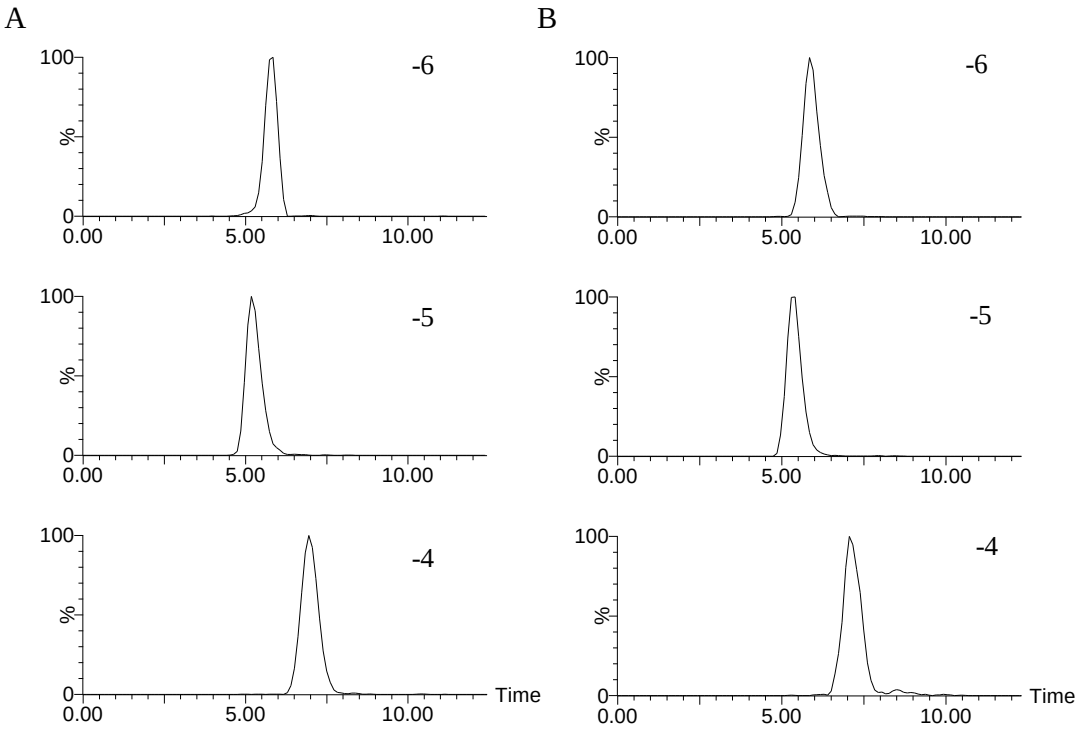


Figure 4.6: Arrival time distributions of the $[M-4H]^{4-}$, $[M-5H]^{5-}$ and $[M-6H]^{6-}$ ions of Oligonucleotide 1 (A) and M= Oligonucleotide 2 (B).

Similar conclusions may be drawn from an analysis of the ATD's extracted from dsDNA peaks. The addition of a charge affects the arrival time to a lesser extent due to the mass being approximately double the mass of either of the two oligonucleotides. In this instance the addition of 8 charges was found to be responsible for the formation of the extended conformation.

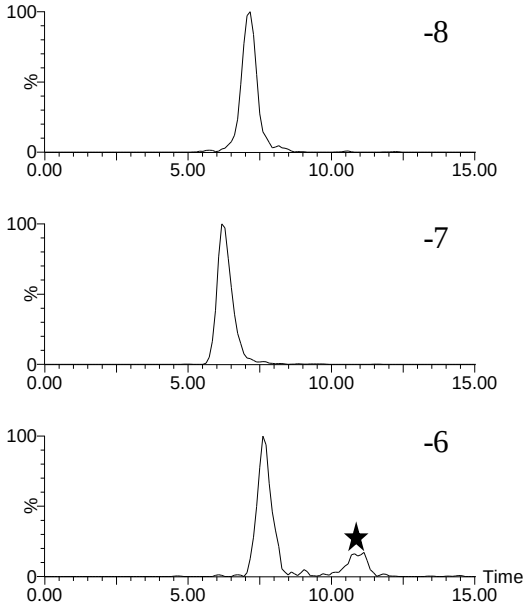


Figure 4.7: Arrival time distributions of the $[M-6H]^{6-}$, $[M-7H]^{7-}$ and $[M-8H]^{8-}$ ions of dsDNA, and the star corresponds to interference from ssDNA.

Since the annealed dsDNA forming the S2/S1 site was over 18 bps in length (22 bps), and was rich in A and T residues (69 %) it is likely to have maintained a B-form helix within the gas phase (Baker and Bowers 2007). In order to confirm the conservation of the dsDNA solution structure, further experiments involving molecular dynamics calculations followed by the calculation of a theoretical collisional cross-section of this species would need to be undertaken. The ATD's could then be used to calculate estimated CCS using either the oligonucleotide calibrant, oligothymidine (Clemmer) and or by using a combination of ssDNA, duplex and quadruplex DNA (Arcella et al. 2012). This would allow a comparison to be made between experimental and theoretical values.

4.3.4 SmtB dsDNA complex

Without zinc ions the dimeric SmtB species is relatively unstable, particularly in the gas phase (Chapter Three), This should not however affect the chances of creating the SmtB:dsDNA complex and transporting it into the gas phase. Apo-SmtB has been found to be relatively stable in solution, under anaerobic conditions, under which it associates with dsDNA. This interaction with DNA should help stabilise the complex.

Different ratios of SmtB and DNA were incubated and then characterised using ESI-MS in positive ion mode, Figure 4.8. At a ratio of 0.5:1 (SmtB:DNA) the abundant ions within the spectrum were found to be derived from dsDNA $[M+ 6H]^{6+}$ and $[M+ 5H]^{5+}$, with a small amount of un-annealed oligonucleotides also being observed. An increased ratio of 2:1 (SmtB: DNA) provides a similar spectrum, although species corresponding to apo-SmtB monomer start to appear. Between 3500-5000 *m/z* peaks also started to appear. They were not very intense and quite broad, making *m/z* assignment difficult. The mixture with a ratio of 4:1 (SmtB:DNA) was observed to be dominated by apo-SmtB monomer peaks, although dsDNA peaks were still visible. The low abundance peaks in the 3500-5000 *m/z* range became more pronounced and assignment of *m/z* values became easier although the peaks were still broad, possibly due to the formation of cation adducts. These peaks could either correspond to higher order DNA complexes, protein complexes or a complex of the two. The absence of the peaks in the 0.5:1 (SmtB:DNA) ratio discounted DNA as a component and highlighted the requirement for apo-SmtB. The *m/z* values predicted from apo-tetrameric SmtB were calculated since this species has been observed by other groups at low abundance. The observed peaks did not match this hypothesis. The peaks were tentatively assigned to a $SmtB_2$:DNA complex with a 1:1 stoichiometry, bound to between nine and twelve protons. The clearer, narrower peaks of the $[M+ 10H]^{10+}$ and the $[M+ 11H]^{11+}$ allowed for more confident assignments, than could be obtained with the $[M+ 9H]^{9+}$ and $[M+ 12H]^{12+}$ peaks.

Although previous studies have not observed monomeric SmtB binding to DNA a cooperative, sequential, monomer binding model has not been ruled out (Busenlehner et al. 2003). Mass spectrometry does not measure the sum of all species present, but indicates the presence of each one individually. The observation of transient intermediates is therefore more likely. The absence of any peaks corresponding to an apo-SmtB monomer:DNA complex, Figure 4.8, provided stronger evidence that dimerisation of SmtB is required before SmtB:DNA complex formation. The need for dimerisation before DNA binding has also been observed in another member of the ArsR/SmtB family, ArsR in *E. coli* (Xu and Rosen 1997).

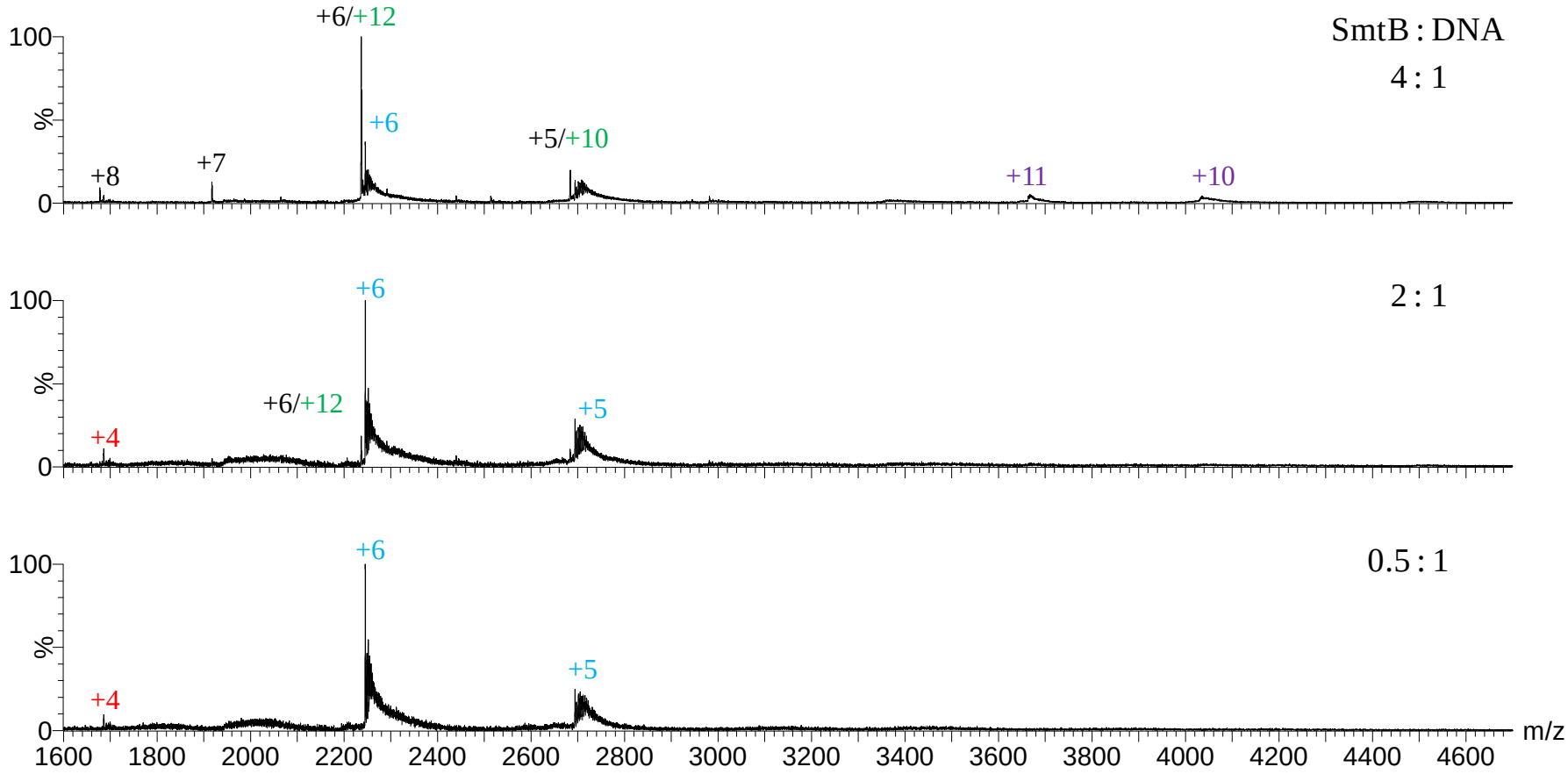


Figure 4.8: Positive ion ESI mass spectra of different ratios of apo-SmtB: dsDNA after incubation at room temperature for two hours. Different species are represented by the following colours, ssDNA (red), dsDNA (blue), SmtB (black) $SmtB_2$ (green) and $SmtB_2$:DNA (purple).

The control mixture containing zinc bound SmtB dimer and dsDNA, which had been incubated under the same conditions revealed no complex association. This control reaction served two purposes. It confirmed that that the fully zinc bound species ($SmtB_2Zn_4$) did not associate with the S2/S1 binding site, Figure 4.9B. This is in agreement with previously published data (Morby et al. 1993; Erbe et al. 1995) where a reduction in DNA affinity was observed on zinc ion binding. The cause of the reduction in DNA affinity has been studied for other members of the ArsR/SmtB family, in particular CzrA (Chakravorty et al. 2012). The electrostatic surface potential was calculated for CzrA in both the high affinity, closed and low affinity open conformation. The protein's DNA binding interface able to be distinguished as a patch of positively charged residues, which are able to interact with the negatively charged DNA molecule. This region was predominately positively charged in the closed conformation, but was disrupted in the open conformation, highlighting the need for electrostatic interactions to form the $CzrA_2$:DNA complex.

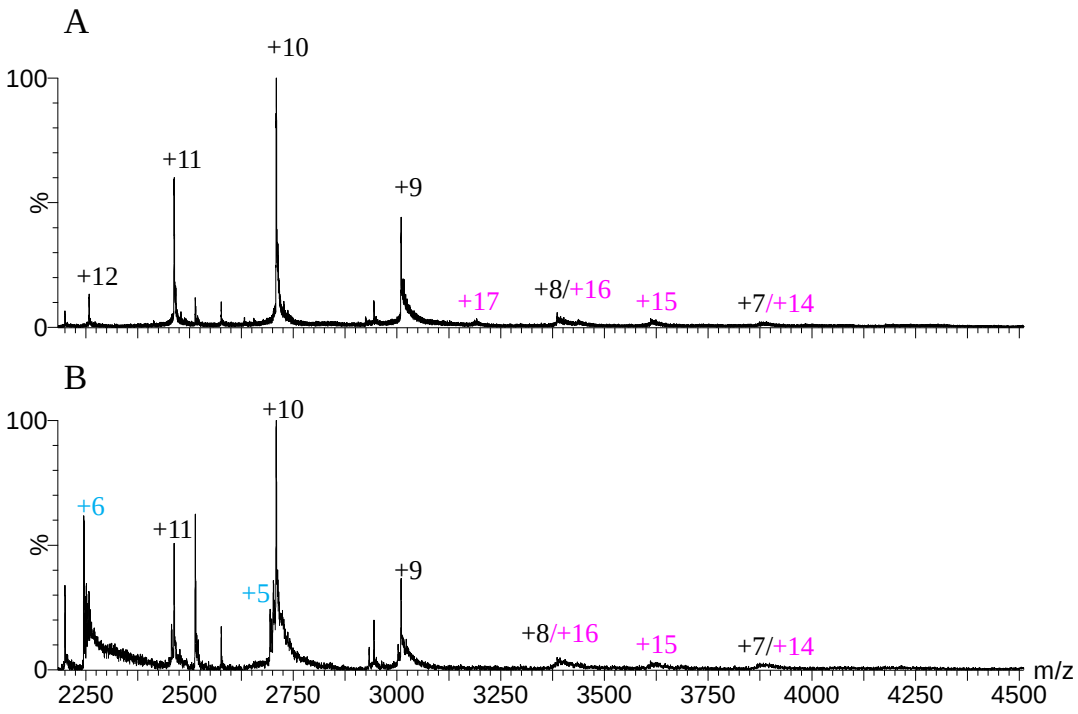


Figure 4.9: Positive ion ESI mass spectra of different ratios of A) $SmtB_2Zn_4$ and B) $SmtB_2Zn_4$ and dsDNA mixture, after incubation at room temperature for two hours. Different species are represented by the following colours, dsDNA (blue), $SmtB_2Zn_4$ (black) and $SmtB_4Zn_8$ (pink).

The absence of an observed complex, in the control sample, also indicated that the binding of apo-SmB to DNA is probably due to specific interactions occurring in the solution phase and is not a by product of the electrospray process. The hypothesis that the interaction is due to specific interactions is supported by the fact that peaks corresponding to apo-SmtB dimer bound to Oligonucleotide one or two were not observed, Figure 4.9, demonstrating that both sides of the DNA are required to form a stable complex.

Once the observation of peaks corresponding to a $SmtB_2$DNA complex had been made, experimental parameters could be optimised including cone voltage and backing pressure. This led to a narrowing of the peaks and an increase in their intensity, Figure 4.10.

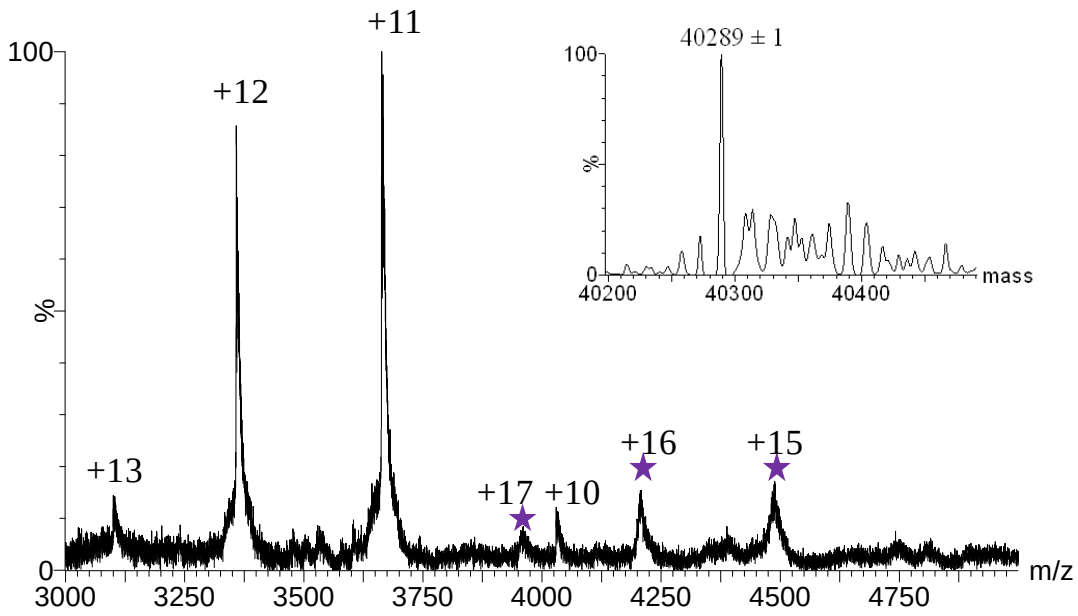


Figure 4.10: ESI mass spectrum of $SmtB_2$:DNA complexes
Two labelled charge state distributions are labelled. The predominant peaks correlate to a $SmtB_2$:DNA complex, whilst possible peaks correlating to a $2SmtB_2$:DNA peak are highlighted with a star. Insert shows the deconvolution of the 1:1 stoichiometry species.

Here two distinct charge state envelopes could be identified, one at a higher intensity than the other. Deconvolution of the mass spectrum resulted in observed masses of

40289 ± 1 Da and 67350 ± 1 Da. The former corresponds well to that calculated for the $SmtB_2$:DNA species (40292.6 Da), whilst the second mass is proposed to relate to the $2SmtB_2$:DNA complex (67119.4 Da). The experimentally observed mass, for the 2:1 complex is 230.6 Da larger than the predicted mass. This is due to the broad low intensity peaks caused by the binding of adducts. Within the optimised spectra there were no peaks observed corresponding to monomeric species bound to the DNA complex or any SmtB species bound to single stranded oligonucleotides.

4.3.5 Dissociation of the SmtB dsDNA complex

The re-addition of zinc ions to an apo-SmtB sample could be successfully achieved in the absence of DNA (Section 3.3.5). It was not determined whether the reduction in signal to noise observed was due to the addition of zinc acetate or protein precipitation. To overcome this problem, when re-adding zinc to the $SmtB_2$:DNA complex either a 1 mol equiv. or 2 mol equiv. of Zn would be added in respect to monomer under anaerobic conditions. This experimental procedure would ensure no Cys oxidation occurred whilst lowering the chance of protein precipitation. The addition of 1 mol equiv. of zinc ions should be sufficient to dissociate SmtB from the O/P sequence, if the zinc sensing α5 are preferentially occupied over the α3N site and the α3N sites do not contribute to SmtB's sensing capabilities. Results also presented indicated that all four sites in the SmtB dimer may be occupied simultaneously in the absence of DNA. It would be interesting to carry out further studies to determine whether 1 or 2 mol equiv. of zinc ions would be required for $SmtB_2$:DNA complex dissociation.

4.4 Conclusions

The use of mass spectrometry methods to characterise nucleic acids and their complexes has advanced significantly over recent years, with the application of TWIMS to complexes implemented in 2010. This research highlights the use of mass spectrometry and ion mobility mass spectrometry to characterise not only metalloregulatory proteins but also their oligonucleotide interaction partners and associated complexes.

Mass spectrometry experiments carried out in negative ESI ion mode allowed the determination of accurate mass for both oligonucleotides and also the annealed double stranded DNA, making up the S2/S1 site, SmtB binding site.

Incubation of apo-SmtB with the S2/S1 site of the *smt* operon resulted in the production of an $SmtB_2$:DNA complex. The stoichiometry was confirmed to be primarily 1:1, but a 2:1 complex was also observed. There was no experimental evidence for zinc bound SmtB forming a DNA complex or the sequential binding of apo-SmtB monomers resulting in the step-wise formation of the complex. These observations were in agreement with the literature to date, which described results obtained by complementary techniques.

Chapter Five: SmtA and its potential interaction with SmtB

5.1 Introduction

5.1.1 Metallothioneins

Metallothioneins (MTs) are metal binding, low molecular weight proteins or peptides, which are rich in cysteine residues. Isolated from a wide range of eukaryotes (Blindauer and Leszczyszyn 2010) and a growing number of prokaryotes (Blindauer 2011), they have been proposed to be involved in storage and detoxification of both essential trace elements (Zn, Cu) and inherently toxic (Cd) metals. Their exact function however has not to date been fully elucidated and they are still of high interest to many research groups. They are ubiquitous within eukaryotes and a classification system has been implemented based on sequence similarity. This involves three separate classes (I-III) (Fowler et al. 1987; Andreini et al. 2008).

5.1.2 SmtA

SmtA was the first prokaryotic MT to be characterised (Olafson et al. 1988). It was initially isolated from *Synechococcus* sp. Amino acid sequence alignment of SmtA to known eukaryotic MT sequences has revealed little sequence homology and consequently it has been categorised as a class II metallothionein. This means that the locations of the cysteines within its primary structure are thought to be only distantly related to the position of the cysteines within equine renal metallothionein (Andreini et al. 2008). Previous work measuring cell growth curves and MT levels, upon exposure to metal ions, showed that upon zinc or cadmium ion addition, a lag in growth was observed. Cellular growth, in these cultures, resumed in line with an increased level of MT (Olafson 1986), highlighting the role of MTs in detoxification and storage. An increase in *smtA* transcripts was later observed upon the exposure of cyanobacterial cells to metal ions, notably zinc and cadmium (Robinson et al. 1990; Huckle et al. 1993). This was later revealed to be due to de-repression of the *smtA* gene by SmtB (Turner et al. 1995).

SmtA from *Synechococcus* PCC7942, expressed in *E .coli*, was determined to be a 55 amino acid protein containing a total of 9 cysteine residues (Shi et al. 1992). Expressed as a fusion protein (GST-SmtA), in metal ion supplemented media, Zn,

Hg, and Cd were all determined to bind to GST-SmtA at higher stoichiometries than to GST protein on its own. Zinc ion binding was maintained even upon the release of SmtA from the fusion protein, confirming its metal ion binding capabilities (Shi et al. 1992). SmtA was determined to have a relatively high affinity for Zn^{2+} in comparison to mammalian MTs, with a recorded pH of half dissociation of 4.10 compared to a value of 4.50 from equine MT (Shi et al. 1992). Using a range of analytical techniques, including ESI-MS and ICP-AES, SmtA was later shown to coordinate four zinc ions per molecule, at a biologically relevant pH (Blindauer et al. 2001). An average zinc binding constant was later determined as 7.9×10^{10} (Blindauer et al. 2007). Unlike the majority of metallothioneins, however, the metal ion cluster was found not solely composed of cysteine residues, instead the cluster was found to use a combination of histidine and cysteine ligands (Daniels et al. 1998). The mutation of His49 and His40, replacing them with Arg, increased the pH of half dissociation of SmtA to 4.62 ± 0.22 and 4.48 ± 0.2 respectively, displaying a reduction in zinc affinity and implying their involvement in Zn coordination (Daniels et al. 1998). The Zn^{2+} ligating residues within the zinc binding cluster were further investigated by NMR spectroscopy, which confirmed the involvement of Cys9, Cys11, Cys14, Cys16, Cys32, Cys36, His40, Cys47, His49, Cys52, Cys54, Figure 5.1.

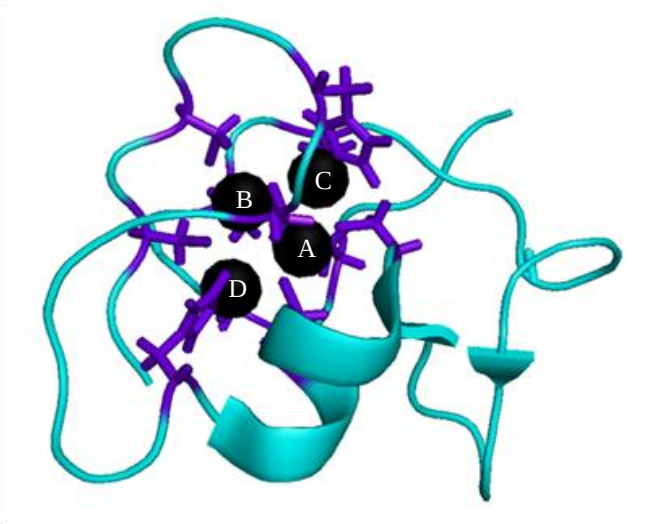


Figure 5.1: NMR structure of SmtA (1JDD)
Zinc ions binding sites are lettered and occupying zinc ions are highlighted in black with their coordinating residues (Cys and His) displayed in purple. The small sections of secondary structure can also be visualised.

SmtA mutants H40C and H49C were created to investigate the importance of the unusual presence of histidine residues instead of the cysteines, which are usually found in MTs (Blindauer et al. 2007). The mutants were still capable of binding four zinc ions. Samples studied using ESI-MS and ICP-AES were found to be a mix of $SmtAZn_4$ and $SmtAZn_3$ species. These results, along with lower average zinc binding constants for the mutants in comparison to wildtype SmtA, demonstrated reduced complex stability.

The NMR solution structure also revealed the presence of a significant secondary structure (Blindauer et al. 2001). A metallothionein's secondary structure usually involves very basic elements such as turns and loops with perhaps a few short helices. SmtA has been found to contain short α-helix and β-sheets, Figure 5.1, which together with Zn site A, form a zinc finger. This is comparable to those found in eukaryotes (Blindauer et al. 2001). Research focusing on the exchange of zinc ions for cadmium ions (Blindauer et al. 2001), along with competition for Zn^{2+} by EDTA (Leszczyszyn et al. 2007) suggested that three zinc ions are readily labile within SmtA's metal binding cluster, whilst the final zinc ion was found to be relatively inactive. The inactive zinc ion forms a constituent part of the zinc finger and is considered therefore to be structural. Upon removal of this zinc ion, SmtA is thought to loose all of its secondary structural elements (Blindauer et al. 2002; Leszczyszyn et al. 2007). The presence of a zinc finger suggests there may be a biomolecular binding partner for SmtA. This could be a protein or DNA, but has not yet been identified.

What is known of the structure of SmtA together with the observed dynamics of its zinc ion binding sites have led to the proposal that, not only does SmtA have the capacity to remove excess zinc ions from the cytosol, but it may also release zinc when required. This, together with the presence of the zinc finger, indicates that SmtA may have the ability to interact with a partner and highlights the fact that SmtA may act as a metal ion chaperone, binding to proteins and transferring their bound zinc. VanZile *et al*. have suggested that the zinc finger may allow the formation of a SmtA:SmtB complex which can then aid in zinc ion removal from SmtB (VanZile et al. 2002a).

Aims

The aims of this research were initially to express and purify SmtA from *E. coli* cultures grown in zinc-supplemented media. The resulting zinc loaded SmtA would then be characterised using ion mobility and mass spectrometry, together with zinc depleted species. The SmtA:SmtB complex hypothesis presented above would then be evaluated.

5.2 Materials and Methods

5.2.1 Expression and purification of SmtA

The expression of the *smtA* gene followed the same procedure, for the expression of *smtB* outlined in Chapter Two. The plasmid used as the DNA source in this transformation of the *E. coli* cells was pMHNR1.1. The purification procedure outlined in Figure 5.2 represents the work flow employed from the bacterial pellet to the purified protein.

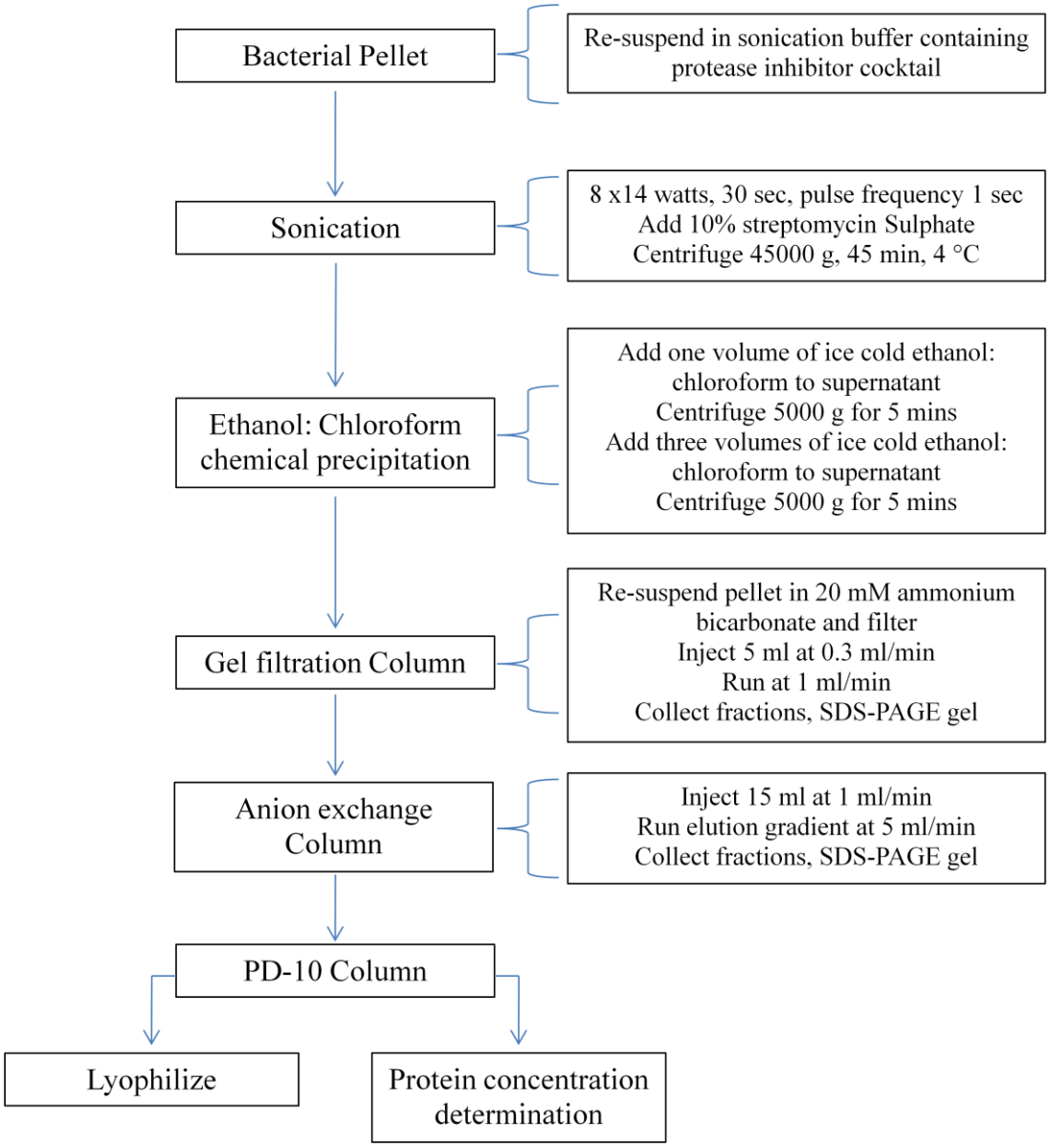


Figure 5.2: Purification workflow for SmtA

To defrosted bacterial cell pellets, 3 mL/g of sonication buffer and 250 µL/g of protease inhibitor cocktail were added to reduce protein degradation. Cells were disrupted by sonication and 0.375 mL/g 10% Streptomycin sulphate was added. Any cell debris produced was pelleted by centrifugation and the resulting volume of the supernatant was measured. One volume equivalence of ethanol: chloroform (100: 8) was added drop wise to the supernatant whilst stirring on ice. The resulting solution was centrifuged and three further volumes of ethanol: chloroform were added to the resulting supernatant, whilst stirring on ice. The next centrifugation step pelleted the smaller proteins within the sample; therefore this pellet was re-suspended in 2 mL/g 20 mM ammonium bicarbonate buffer and filtered through a 45 µM membrane. The filtrate was then made up to 5 mL, with 20 mM ammonium bicarbonate, and loaded into a Superdex 75 HiLoad 16/60 size exclusion column at 0.3 mL/min, from a superloop. The buffer used was 20 mM ammonium bicarbonate buffer, filtered and degassed before use. A SDS-PAGE gel was used to identify the SmtA containing fractions, Figure 5.3. A mass spectrum indicated that the sample was impure, Figure 5.3 and so further purification was required.

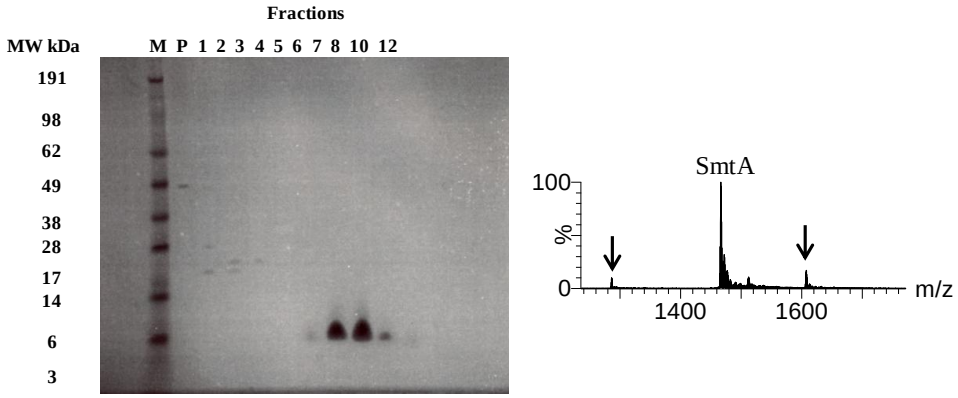


Figure 5.3: SDS-PAGE gel of the fractions obtained from the SEC column
Lane M was loaded with pre-stained molecular weight markers, their weight is indicated to the left of the gel. The subsequent lanes were loaded with 10 µL of sample from the fractions indicated above the lanes. Inset shows a native mass spectrum of SmtA obtained from pooled fractions. Arrows indicate contaminating proteins.

SmtA fractions obtained from SEC chromatography were loaded onto an anion exchange column in 10 mM ammonium bicarbonate and eluted using a 0-100%

gradient of 10 mM ammonium bicarbonate, 1 M NaCl over 20 column volumes. Fractions with an absorbance at A_{220} were run using an SDS-PAGE gel and the resulting SmtA fractions were run through a PD-10 desalting column (sephadex G 25, GE Healthcare). The desalted 3 mL fraction had its protein concentration determined by the Cys assay (Chapter Two) and then was aliquoted and lyophilised for storage at -80 °C.

5.2.2 Sample preparation for ESI mass spectrometry and ion mobility

A lyophilised SmtA sample was re-suspended in 10 mM ammonium bicarbonate, adjusted to pH 7.4 and concentrated using Millipore centrifugal filter units with a 3 kDa cut off. The concentrated sample was washed five times to ensure the removal of excess sodium ions, using the protocol outlined for SmtB (Chapter Three). Spin times were extended due to the use of a lower molecular weight spin filter. Denatured samples were prepared by dilution of the concentrated sample to ~ 10 µM in 30 % acetonitrile, 0.1 % formic acid, whereas native samples were diluted with ammonium bicarbonate pH 7.4.

5.2.3 EDTA time course

The EDTA time course was conducted as outlined in Chapter Three for the SmtB time course. In brief a sample containing 0.8 mM SmtA and 1.6 mM EDTA was left to incubate at room temperature. Aliquots were taken after 7, 22, 30, 60, 90, 134, 180 minutes. Each aliquot was diluted down to 13 µM in 10 mM ammonium bicarbonate, pH 7.4 before introduction into the mass spectrometer.

5.2.4 Mixing SmtB with SmtA

To investigate any possible interaction between SmtB and SmtA, a control reaction was set up. Lyophilised samples of SmtA and SmtB were resuspended in 10 mM ammonium bicarbonate pH 7.4. These samples then underwent buffer exchange in amicon centrifugal devices using a 3 kDa cut off and 10 kDa cut off for SmtA and SmtB respectively. Following a number of desalting steps, up to five in total, the proteins were concentrated into stock solutions. A number of different reaction mixtures were then set up with SmtA:SmtB ratios between 1:1 and 1:60 SmtA:SmtB. Following a short incubation period of 15 minutes the samples were diluted to a total protein concentration of 10-15 µM, in 10 mM ammonium bicarbonate pH 7.4. An

aliquot of these samples was then loaded into fused silica nanospray needles and mass spectrometry measurements were carried out as outlined below.

5.2.5 ESI mass spectrometry and ion mobility analysis

Mass spectrometry experiments were carried out using a Synapt G2 HDMS system (Waters, Milford, USA) utilising a nanoflow electrospray source in conjunction with fused silica nanospray needles. Before data acquisition a calibration using 2 mg/mL CsI, dissolved in 50% aqueous isopropanol, was conducted in positive ion, resolution mode, scanning between 500-3500 *m/z*. Experiments were carried out with 1.2 kV applied to the capillary, a source temperature of 80 °C, with the cone voltage being varied between 20 V, 40 V or 60 V depending on the experiment. The optimum cone voltage for native conformations was determined to be 20 V, whilst 60 V was applied to investigate partial unfolding at elevated energies. MassLynx software V4.1 was used for the acquisition and processing of data. Deconvolution of the high resolution data required the use of the MaxEnt3 algorithm instead of the MaxEnt1 algorithm previously used in conjunction with SmtB measurements.

The following ion mobility parameters were used; wave height 40 V, IMS gas 90 mL/min, wave velocity 800 m/s. ATD's were extracted and visualised, to obtain their arrival time, using MassLynx software V4.1 in conjunction with Driftscope software. In one instance for the $SmtAZn_4$ +3 charge state an unusual isotopic distribution was observed, compared to the predicted model generated by the MassLynx software. Upon extraction of the ATD from the +3 peak two separate arrival times were observed. Back extraction of each arrival time peak, produced a mass spectrum of the ions whose drift times fell within that ATD. This procedure revealed that only one back extracted ATD produced a SmtA isotopic distribution. To produce accurate CCS for SmtA all ion mobility ATD data was calibrated over the mobility range, using a protein with known CCS and arrival time. In this instance sperm whale myoglobin was used in the calibration procedure, (as outlined in Chapter One).

5.2.6 MOBCAL

To produce theoretical CCS's of SmtA the MOBCAL software was used with the PA and EHSS algorithms. The NMR structure 1JJD, available in the Protein Data

Bank, was used to generate a theoretical cross section of the $SmtAZn_4$ species allowing comparisons between experimentally derived rotationally averaged CCS's and calculated CCS's to be made.

5.3 Results and discussion

5.3.1 Mass spectrometry of SmtA

Mass spectrometry measurements were used to ensure the purity of the prepared SmtA. The sample was run under denaturing conditions and the spectrum obtained deconvoluted onto a true mass scale, Figure 5.4. Under denaturing conditions a charge state envelope was produced, which included charge states corresponding to apo-SmtA with between 4 and 8 protons attached, Figure 5.4.

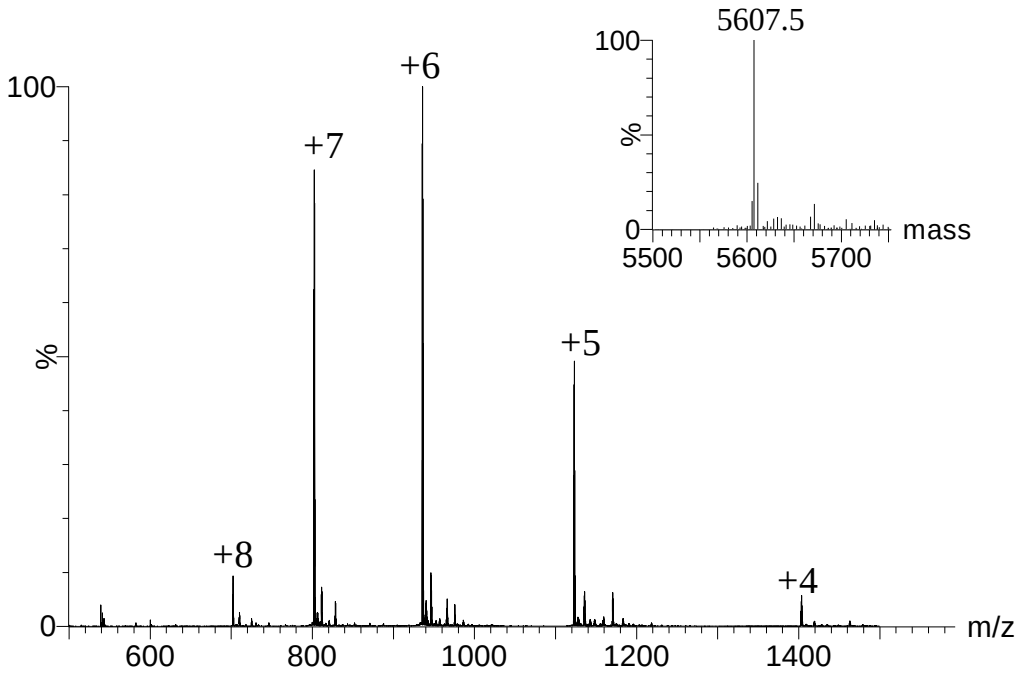


Figure 5.4: ESI mass spectrum of denatured SmtA
SmtA concentration was 10 µM in 30 % acetonitrile, 0.1 % formic acid. Inset shows the neutral deconvoluted mass.

The singly charged deconvoluted mass was calculated by the deisotoping MaxEnt 3 algorithm as 5607.5 Da, with a small adjacent peak of 5671.3. The most abundant mass observed agreed well with the theoretical monoisotopic mass of apo-SmtA with a cleaved N-terminal methionine (5606.34 Da). The small adjacent peak was calculated to be 63.8 Da larger. This was thought to correspond to a small amount of $SmtAZn_1$ species being present due to the high affinity of the final zinc ion binding site which has been suggested in the literature (Blindauer et al. 2001; Blindauer et al. 2007; Leszczyszyn et al. 2007).

5.3.2 Ion mobility and mass spectrometry of native and denatured SmtA

SmtA has a pI of 7.85 meaning that under pH 7.4 solution conditions the protein will be positively charged. A near native mass spectrum was therefore obtained by positive mode ESI and confirmed that the purified SmtA was monomeric with a single tight charge state distrubution, Figure 5.5. A comparison of masses obtained from the native mass spectrum and the denatured mass spectrum confirmed that the zinc ion to protein stoichiometry was $SmtAZn_4$. The mass difference of 253.5 Da observed corresponded to the addition of four zinc ions with the loss of 8 protons, ((4 x 65.39)-8 = 253.56).

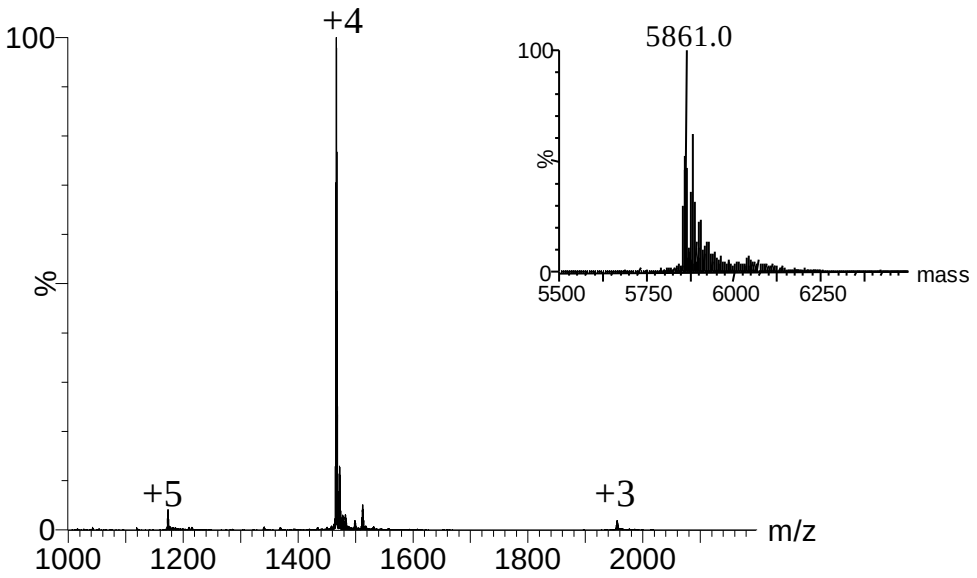


Figure 5.5: ESI mass spectrum of 10 µM SmtA in 10 mM ammonium bicarbonate pH 7.4. The inset shows the deconvoluted spectrum.

SmtA is a small protein and the resolution achieved by the TOF mass analyser (FWHM = 15,000), in resolution mode, is high enough to achieve baseline resolution of the isotopic peaks. This separation allowed for a comparison of the observed $[SmtA+4Zn-4H]^{4+}$ peak with that generated from an isotopic model, using the molecular formula of $SmtAZn_4$, Figure 5.6. The isotopic distribution of the peak generated matched almost exactly with the experimental data, suggesting the presence of intact SmtA with four zinc ions bound.

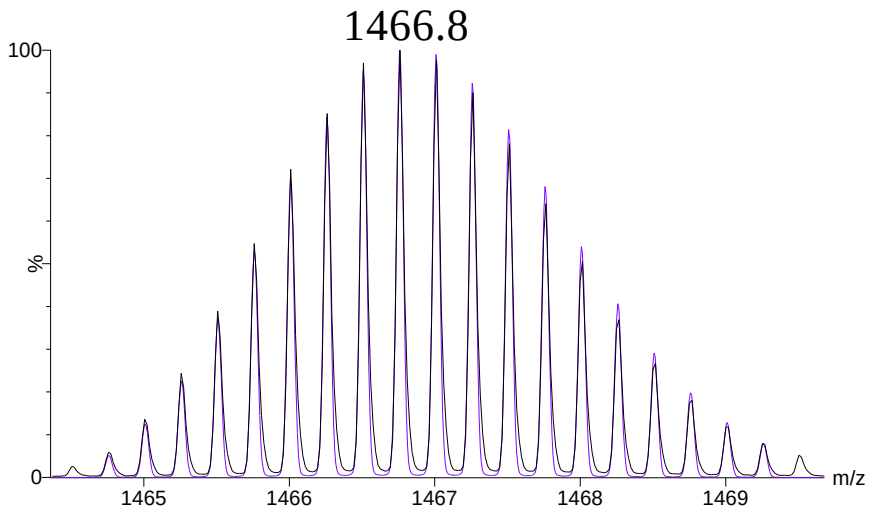


Figure 5.6: An overlay of the $[SmtA+4Zn-4H]^{4+}$ peak (black) and an isotopic model (purple) of the same peak.

From observed arrival time distributions (ATD's) estimated collisional cross-sections were determined using the calibrant sperm whale myoglobin. These were then compared to the theoretical CCS produced by MOBCAL software from the solution phase NMR structure of zinc bound SmtA 1JJD. The NMR structure had three residues (TST) missing from the N-terminus of the sequence and may result in a slightly smaller theoretical CCS, than if the full sequence had been present.

The estimated CCS's 705 Å², 706 Å² and 789 Å² for the zinc bound native SmtA charge states $[SmtA+4Zn-5H]^{3+}$, $[SmtA+4Zn-4H]^{4+}$ and $[SmtA+4Zn-3H]^{5+}$ lay within or just below the two theoretical values determined by the PA and EHSS methods, Figure 5.6. For larger molecules (> 2 kDa) the PA method has been shown to underestimate CCS (Bleiholder et al. 2011; Jurneczko and Barran 2011) whereas for smaller ions such as SmtA, EHSS over estimates the CCS (Jurneczko and Barran 2011). Explanations concerning properties of the different algorithms can be found in Chapter Three. Results for the lower charge states of SmtA, +3 and +4, are consistent with the observation that estimated CCS for monomeric proteins are often smaller than their calculated CCS (Jurneczko and Barran 2011). This is thought to arise from the polar side chains collapsing and self solvating in the gas phase structure (Breuker and McLafferty 2008). The denatured sample, as seen from the

mass spectrum, Figure 5.4, contains more unfolded protein species leading to the production of higher charge states during electrospray. The addition of extra protons to unstructured proteins can often lead to charge repulsion, causing unfolding in the gas phase. This structural unfolding may cause the estimated CCS's to lie above the calculated CCS values, Figure 5.7.

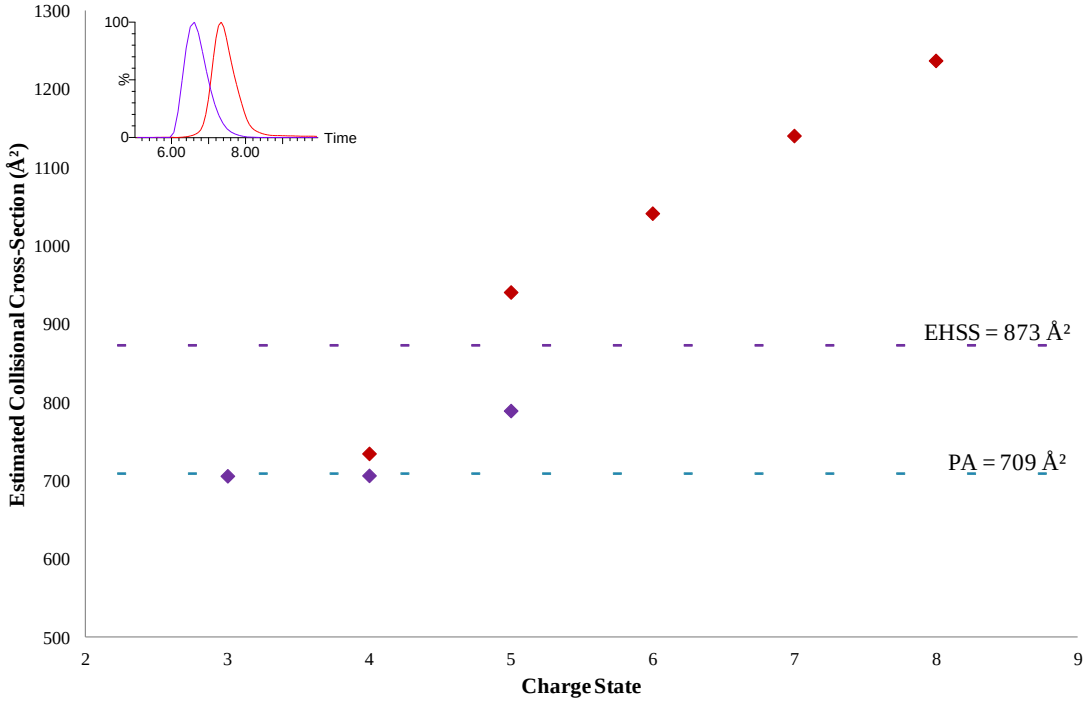


Figure 5.7: A plot of the estimated CCS's of holo-SmtA, native (purple) and denatured (red) apo-SmtA.
The dashed lines indicate the theoretical CCS, determined by MOBCAL software using the PA and EHSS algorithms. The insert shows an example of the ATD's of the +4 SmtA peak.

There is some overlap of charge states observed between the native and denatured samples. The presence of the +4 and +5 charge states within both experiments may allow comparisons to be made. The difference between the CCS's of the +5 charge states is significant, with the denatured sample having a volume ~150 Å² larger than that of the native sample. The difference between values for the +4 charge states is relatively small with the denatured sample larger by ~ 25 Å², Figure 5.7. This may indicate that, although the structure of SmtA is disrupted, the addition of four protons does not give rise to a CCS corresponding to that of the unfolded state. This

result was taken into consideration when designing the ion mobility mass spectrometry zinc depletion experiments carried out on SmtA.

5.3.3 Native mass spectrometry of zinc depleted SmtA species

A combination of NMR and ESI-MS experiments has recently suggested the presence of a $SmtAZn_1$ intermediate during EDTA mediated zinc ion removal from SmtA (Leszczyszyn et al. 2007). To investigate this further, a SmtA:EDTA time course was followed by ESI-MS, under the same concentrations as the SmtB time course, Figure 5.8. Direct kinetic comparisons can not easily be drawn between the published SmtA time course (Leszczyszyn et al. 2007) and Figure 5.8, due to the use of different experimental conditions.

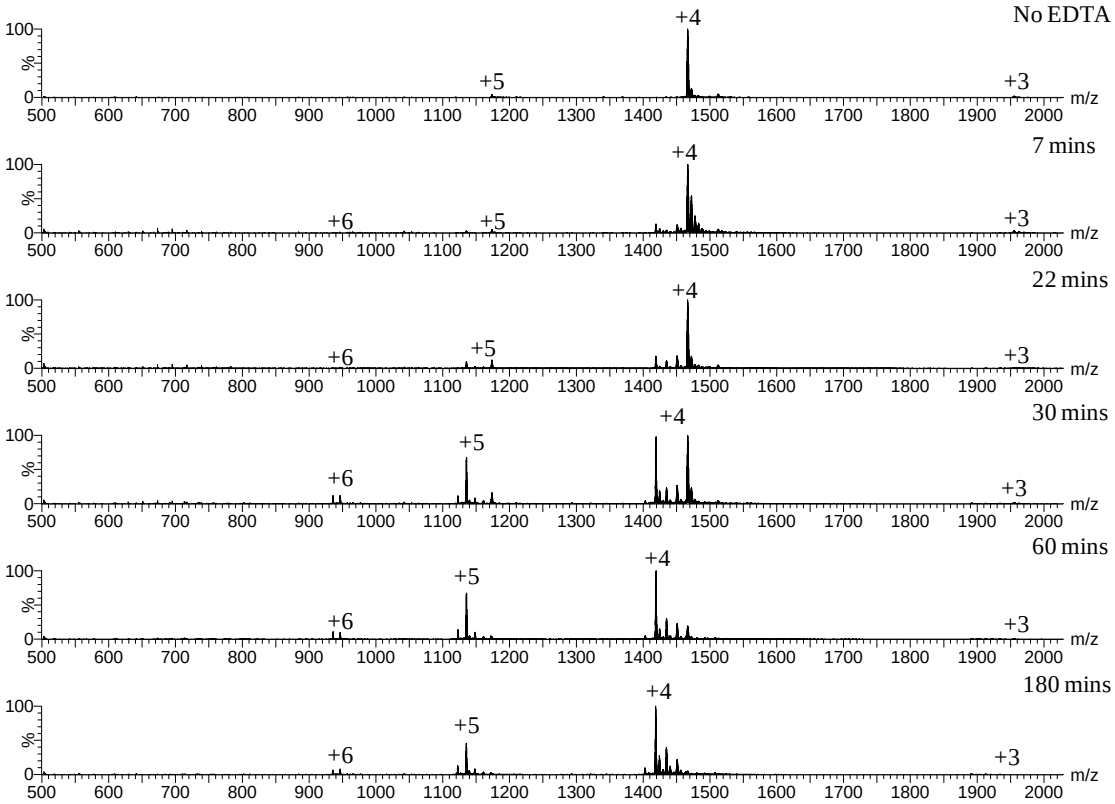


Figure 5.8: ESI mass spectra of 0.4 mM SmtA in 10 mM ammonium bicarbonate pH 7.4 incubated with 1.6 mM EDTA and sampled after the indicated time points.

Trends observed between the published SmtA time course (Leszczyszyn et al. 2007) and the time course presented here can be identified. When these factors are taken into consideration the SmtA time course proceeded as predicted (Leszczyszyn et al.

2007), for fully loaded $SmtAZn_4$. The SmtA was isolated lacking the terminal Met and not with the three residue cleavage (MST) seen previously (Leszczyszyn et al. 2007). In the absence of EDTA a tight charge state distribution was observed with a predominant +4 charge state. The +4 charge state remained the most abundant charge state throughout the time course, although the proportion of +5 species increased and peaks corresponding to the +6 charge state appeared with time.

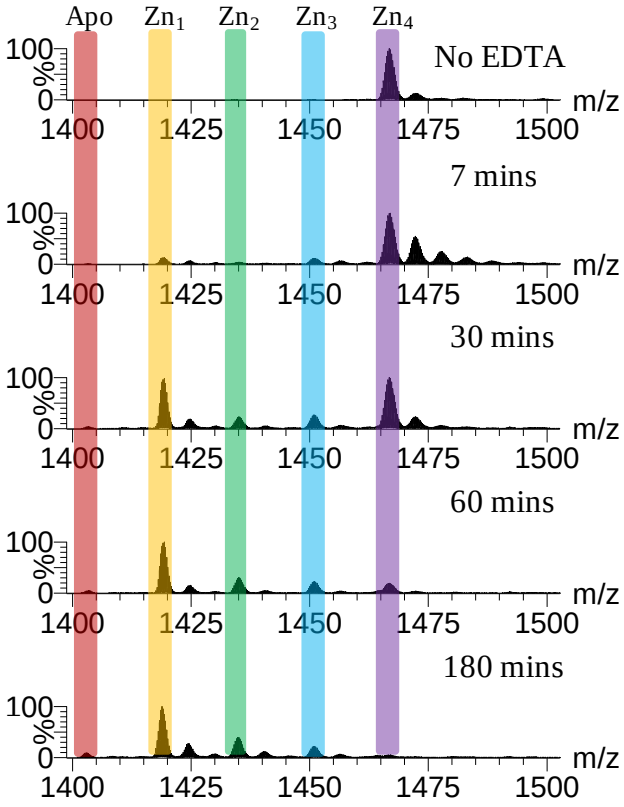


Figure 5.9: A close up of the +4 SmtA charge state after incubation with an equimolar concentration of EDTA, in respect to zinc ions, after 7, 30, 60 and180 minutes.

Unlike SmtB, zinc ion depletion does not occur sequentially from SmtA. The removal of the first zinc ions from site C is rapidly followed by zinc ion removal from sites B and D followed by a build up of the $SmtAZn_1$ species, Figure 5.1 (Leszczyszyn et al. 2007). These distinct phases were observed in the displayed time course, Figure 5.9. Metal depletion by EDTA chelation began almost straight away with a $SmtAZn_3$ and a $SmtAZn_1$ species observed, at low concentrations within 7 minutes. The very low abundance of the $SmtAZn_2$ species observed suggests that

zinc ion removal, from sites B, C and D proceeds rapidly. The gradual build up of $SmtAZn_1$ and slow transition to apo-SmtA was observed and agreed well with the previously published theory that this is an intermediate species (Leszczyszyn et al. 2007).

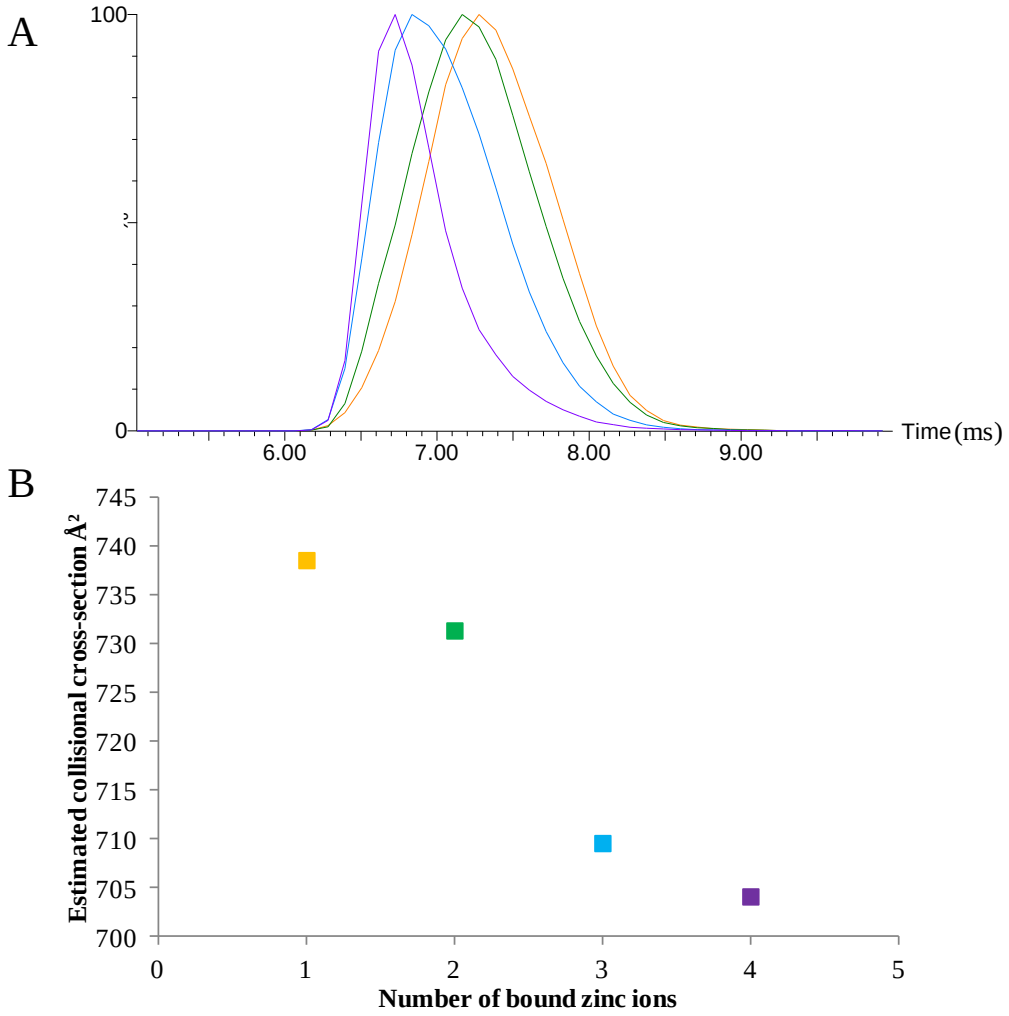
When the full spectra were looked at in more detail it became apparent that the +6 charge state appeared very early on within the time course. In fact a +6 species is detected at low abundance within 7 minutes. The speciation of this charge state was different to the +4 charge states, since the only observable +6 species throughout the time course were due to SmtAZn and apo-SmtA. The +5 charge state, observable from time point zero, was gradually dominated by these two zinc depleted species, Figure 5.8. The speciation of the higher charge states indicated that the loss of three or more zinc ions from SmtA extended the conformation of the protein sufficiently to allow further protonation during electrospray process. Although a slight broadening of the charge state distribution occurred during zinc ion removal, no significant protein cleavage was observed, even when a small amount of apo-SmtA, which has been seen previously (Leszczyszyn et al. 2007), was observed.

5.3.4 Ion mobility mass spectrometry of zinc depleted SmtA species

Removal of zinc ions by EDTA competition allowed the conformations of SmtA species to be investigated. The stability of these conformations was also studied by exposing the ions to higher energies within the mass spectrometer. Since the +4 SmtA charge state was the most prominent charge state observed in all recorded native mass spectra, it was chosen to investigate any changes in conformation associated with the loss of zinc ions from SmtA, Figure 5.10.

Under native like conditions the +4 $SmtAZn_4$ charge state was observed as a single resolvable conformation, with a CCS estimated at 704 $Å^2$. The loss of a single zinc ion from SmtA resulted in a species with a slightly expanded structure as shown by the increase in estimated CCS to 710 $Å^2$. The ATD width observed for the $SmtAZn_3$ species increased substantially, by approximately one third at full width half maximum (FWHM), compared with the ATD extracted from the fully loaded species. The widening of the ATD may suggest the possibility of multiple conformations of a similar CCS, perhaps indicating the presence of a more dynamic

protein. The loss of the next zinc ion, which results in the formation of the $SmtAZn_2$ species increased the CCS of the +4 charge state by around 21 Å². The production of a potentially folded intermediate $SmtAZn_1$ caused a further slight increase to the



CCS to 739 Å². The production of a stable apo-SmtA +4 species did not occur with a high enough intensity for IMMS experiments to be carried out

.

Figure 5.10: ATD's of the different species of the +4 SmtA charge state (A) and corresponding plot of estimated CCS (B).
The colours represent the different stoichiometries of zinc ions to protein after the incubation with EDTA. Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$ and Orange = Zn^{2+}.

The production of apo-SmtA from exposing the protein to denaturing conditions resulted in a observed CCS of 734 Å² for the +4 charge state, slightly smaller than the estimated CCS for the $SmtAZn_1$ intermediate. This does not provide an accurate representative candidate for native apo-SmtA; but it is interesting to note that the

rotationally averaged CCS are similar, perhaps suggesting that $SmtAZn_2$ and SmtAZn exist in a largely unfolded state.

Information obtained from native MS and IMMS experiments can tentatively be used to explain the stepwise mechanism of zinc ion removal from $SmtAZn_4$, Figure 5.11. The mechanism which has been proposed involves four phases (Leszczyszyn et al. 2007). In phase one EDTA is proposed to attack $SmtAZn_4$, at site C, directly leading to the formation of the $SmtAZn_3$ species. The wide ATD obtained from the $SmtAZn_3$ species suggested that this species may be more elongated and dynamic than the $SmtAZn_4$ species. This increase in dynamics may be caused by the removal of the site C zinc ion. The zinc ion, within this site, is ligated by C16, C32, C47 and H49. The latter two ligating residues are thought to be situated on a flexible loop (Blindauer et al. 2007) and the bound zinc ion bridges and may stabilise this loop to the remainder of the structure.

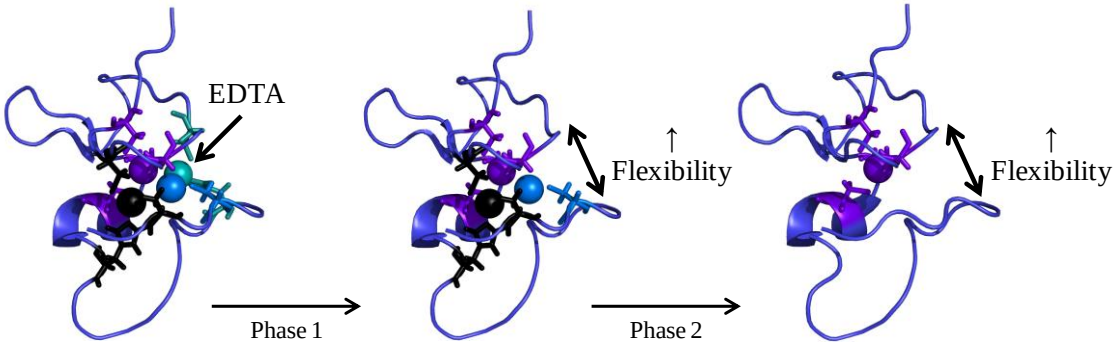


Figure 5.11: Schematic representation of the zinc ion removal model from $SmtAZn_4$.
The four zinc ion sites are colour coordinated; site A is purple, site B is blue, site C is green and site D is black. Zinc ion removal is shown by the deletion of the zinc ion and ligating residues from the structure, however all three structures are 1JDD.

The following step (phase two), Figure 5.11, is thought to involve the removal of the next two zinc ions, at a fast rate. The flexible cleft possibly left by the removal of the site C zinc ion may allow the B and D site to become exposed to EDTA leading to their quick removal. These two zinc ions also bridge the cleft and so the loss of these ions may elongate the structure substantially, as shown by the observed increase in CCS and the appearance of highly charged SmtA species. The removal of the final

zinc ion (phase 3) is thought to remove all solution phase secondary structural elements, whilst phase 4 was suggested to correspond to apo-SmtA degradation, not observed under these conditions.

To investigate the gas phase stability of the different stoichiometric species, ATD's were collected at higher cone energies (cone voltage 60 in comparison to cone voltage 20), Figure 5.12. All of the extracted ATD's contained a single resolvable conformation of similar FWHM. The fully loaded SmtA species were found to possess the most compact structure with an estimated CCS of 744 ± 4 Å².

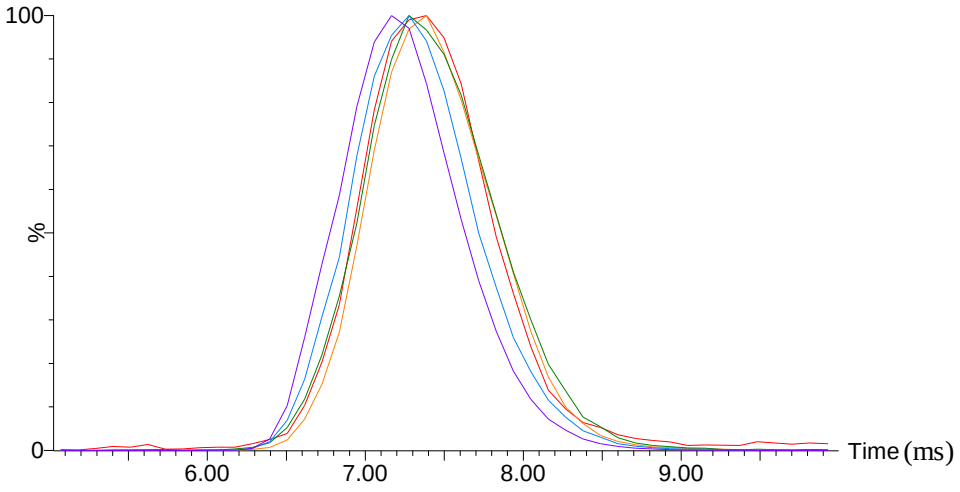


Figure 5.12: ATD's of the different species of the $[M+4H]^{4+}$ SmtA charge state upon exposure to higher energy.
The colours represent the different stoichiometries of zinc ions to protein after the incubation with EDTA. Purple = $4Zn^{2+}$, Blue = $3Zn^{2+}$, Green = $2Zn^{2+}$, Orange = Zn^{2+} and Red = Apo.

The increase of estimated CCS upon zinc ion depletion, until the production of the $SmtAZn_1$ intermediate, was small. An expansion of only 1.1 % to that of the fully loaded species, to a CCS of 752 ± 5Å² was observed. In this experiment a small amount of apo-SmtA was detected and the extracted ATD produced an estimated CCS of 751 ± 5 Å². This was within the experimental error of the calculated CCS of the intermediate species. This was surprising since, upon the loss of the final zinc ion, all secondary structural elements within SmtA have been proposed to be lost. A loss of secondary structure would be likely to result in less structural confinement within the protein, leading to either a more unfolded structure or increased flexibility

when compared with the $SmtAZn_1$ species. This should lead to a wider observed ATD or a more extended estimated CCS, neither of which were clearly observed. The higher energies used for this data set have elongated all the cross sections of SmtA, and further unfolding of the structure does not seem to have occurred.

These results could lead to two possible conclusions, when considering what happens to the SmtA structure upon the removal of the final zinc ion. One possibility is that the increased flexibility and partial unfolding which occurs during zinc ion removal produces a structure that possesses a CCS similar to that of the unfolded protein. This would mean that when the secondary structural elements are lost, upon removal of the final zinc ion, no distinct change in CCS would be observed. A second possibility is that some secondary structural aspects remain in the gas phase, stabilizing the structure and preventing it from unfolding further. This possibility has been previously observed in zinc fingers. A Cys_2His_2 zinc finger peptide has been studied in the gas phase (Berezovskaya et al. 2011). The peptide lost secondary structure in the solution phase upon the removal of its single zinc ion. In the gas phase however apo and holo-peptide produced extremely similar CCS's. Molecular dynamics (MD) simulations revealed that transient helical structures existed within the apo-form in the gas phase stabilising the structure, whilst these interactions were absent in the solution phase apo structure (Berezovskaya et al. 2011).

There is not enough experimental evidence to rule out either of the two of these possibilities. The use of MD simulations, further ion mobility experiments and CID could provide more information and may lead to a definitive answer.

5.3.5 Mixing SmtA and SmtB

A question often posed when studying SmtA and its repressor protein SmtB is whether zinc transfer occurs between them (VanZile et al. 2000; VanZile et al. 2002a). In a biological context the transfer, of zinc ions, from fully zinc loaded SmtB to either apo-SmtA or partially loaded SmtA would seem more logical than transfer in the opposite direction. This process, if faster than apo-*smtB* transcription and translation, could result in an efficient off switch for the homeostatic system to employ. In a biochemical context the transfer, from SmtB to SmtA, seems unlikely to depend on thermodynamics, since the binding constants for zinc ions to SmtB

(VanZile et al. 2002b), along with the pH of half dissociation (Chapter Three) show SmtB to have a higher affinity for zinc ions than that of SmtA (Shi et al. 1992; Blindauer et al. 2007) . The kinetics displayed by SmtA may indicate that, like many aspects of zinc homeostatic systems, kinetics and not thermodynamics govern metal ion transfer (Finney and O'Halloran 2003). In the presence of zinc ions, and de-repression of the smt operator-promoter sequence both SmtA and SmtB proteins are thought to be produced. The *in-vivo* rate of *smtA* transcription, upon zinc ion exposure is known to be fast and sustained, in comparison to that of *groEL* (Ybarra and Webb 1999). The transcription rate of *smtB* remains unknown but one may speculate that the de-repressed *smtA* gene would be transcribed at a faster rate than *smtB* due to the protein product having a proposed zinc ion storage and detoxification function. It is therefore likely that *in-vivo* SmtA:SmtB ratio would lie in favour of SmtA.

Samples containing a mixture of $SmtAZn_4$ and $SmtB_2Zn_4$ were investigated. These experiments were used to gauge if the two proteins under investigation had different MS response factors. The relative stability of the zinc bound versions allowed many different ratios to be investigated, without extensive dialysis, and the potential sample loss involved in conducting dialysis. Assuming that the loss of zinc ions would not dramatically affect the proteins response factor, these results could then be used to design appropriate experiments to study a mixture of any zinc depleted species.

Incubation of zinc loaded SmtA and SmtB, followed by the analysis of the mixture by ESI-MS resulted in the following spectrum, Figure 5.13. The incubation of a 1:1 mixture of SmtB:SmtA resulted in a mass spectrum dominated by SmtA signals, with no observable contribution from SmtB. This is thought to be due to the difference in response factor between the two proteins. A proteins response factor depends on its ability to be ionised, transmitted and detected (Boeri Erba et al. 2011). Variations in ionisation efficiency could be caused by a difference in the percentage of exposed non-polar residues (Cech and Enke 2000) or a difference in charge state (Gabelica et al. 2003). Within a mixture the influence of other components, often refered to as matrix effects (Pan and McLuckey 2003b) can also influence a proteins response factor. The matrix effects caused by small cations (Pan and McLuckey

2003a) and other proteins (Pan and McLuckey 2003b) show that proteins can be more effective than small cations at competing for charges during ionisation, leading to the suppression of other protein signals. In this instance it is perhaps the lower pI of SmtB (6.54) compared with SmtA (7.85) which may lead to the observed difference in response factors, as observed in other studies (Pan and McLuckey 2003b). The spectrum obtained after the incubation of a 6:1 SmtB:SmtA ratio was still dominated by SmtA peaks, however the more abundant $[SmtB_2+4Zn+4H]^{12+}$ and $[SmtB2+4Zn+3H]^{11+}$ peaks become visible upon closer inspection.

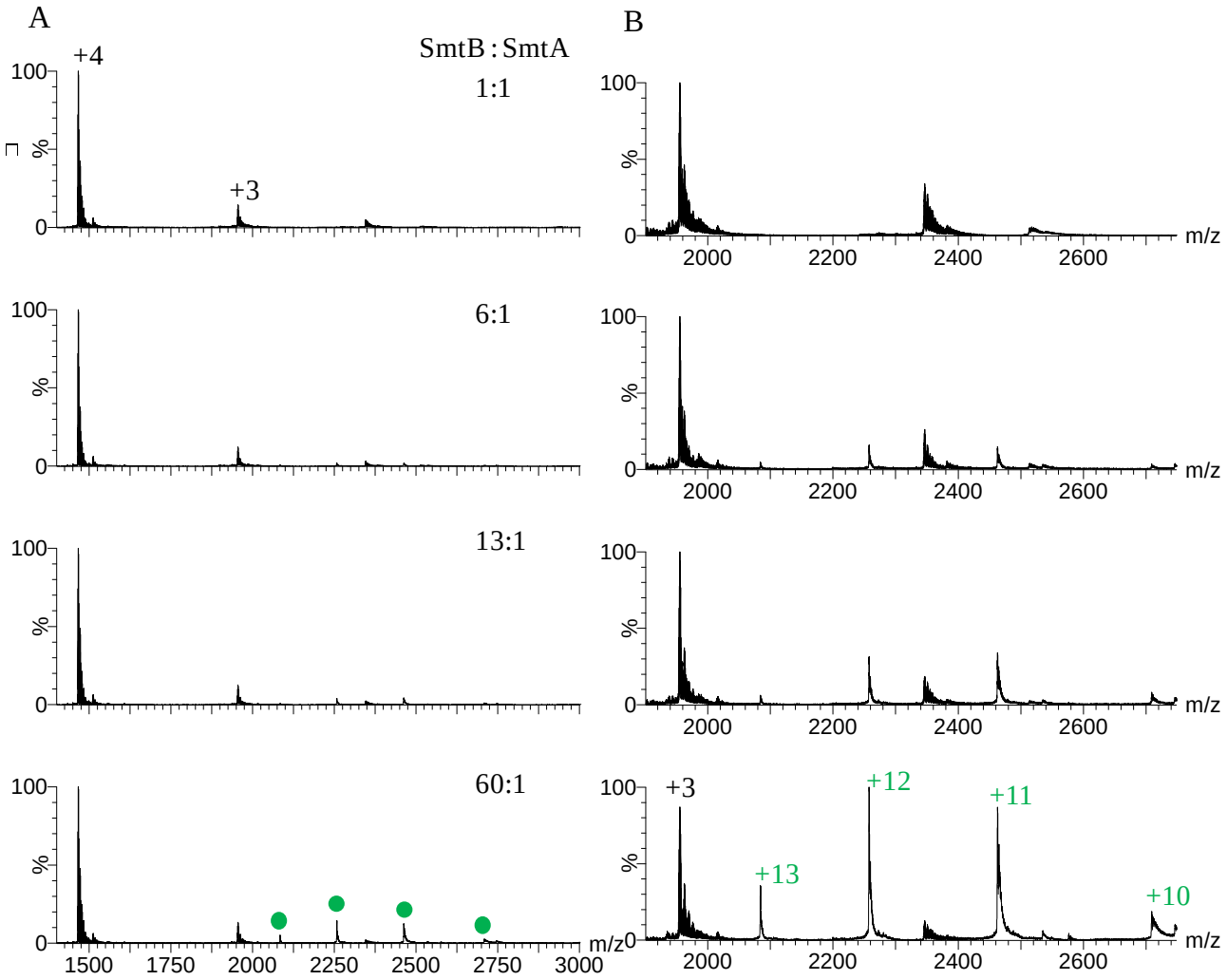


Figure 5.13: ESI-MS spectra resulting from the mixture and incubation of different ratios of $SmtAZn_4$ and $SmtB_2Zn_4$
A) Full scan showing the different ionisation efficiencies of SmtA and SmtB. B) zoomed in section showing the peaks relating to $SmtB_2Zn_4$

It was not until a 13:1 ratio was incubated that both charge state distributions were observed and not until a 60:1 ratio that clear distinct distributions could easily be seen, Figure 5.13. The high ratio of SmtB to SmtA, required for the visualisation of the two species, is opposite to that which is thought to exist in the cyanobacterium when subjected to zinc ion stress. This observed ratio could also indicate that the loss of zinc ions from SmtB could be difficult to follow due to the low signal to noise. The gain of zinc ions by SmtA should therefore be taken forward.

From the results obtained, Figure 5.13, there were no observed interactions between $SmtB_2Zn_4$ and $SmtAZn_4$. This suggests that no unspecific interactions occur between these two species and they could be used as a control in future experiements. The most likely SmtA species to interact with SmtB would be the $SmtAZn_1$ intermediate species. This species has been shown to be more stable in solution than apo-SmtA and contains the zinc finger motif potentially required for interactions. The removal of zinc ions, as outlined in Chapter Three, using dialysis against EDTA is a laborious process and must be carried out in the absence of oxygen. Once obtained the zinc depleted SmtAZn sample has to be extensively washed to remove any residual EDTA before mixing with a zinc loaded $SmtB_2Zn_4$, otherwise EDTA could compete with SmtA for zinc ion removal. To remove the requirement of the use of EDTA a low pH technique could be employed. This would involve exposing SmtA to a low pH, removing dissociated zinc ions via PD-10 columns and then slowly raising the pH by the addition of Tris base, before buffer exchange into nitrogen saturated ammonium bicarbonate. The addition of this species in a ratio of 13:1 $SmtB_2Zn_4$:SmtAZn under anearobic conditions followed by frequent sampling and anaysis may reveal if SmtA's binding partner is SmtB and whether zinc tranfer between these two species is possible.

5.4 Conclusions

This work has demonstrated the applicability of mass spectrometry and travelling wave ion mobility for the study of the conformation of metalloproteins and their specific interactions with metal ions.

A combination of denatured and native mass spectrometry experiments revealed that the outlined expression and purification procedures produced the metallothionein SmtA. This protein was shown to have undergone the usual N-terminal Methionine cleavage but was otherwise full length. The binding stoichiometry was confirmed to be $SmtAZn_4$.

Ion mobility data conducted under native conditions produced estimated CCS data in agreement with the theoretical CCS data produced from the published NMR solution structure of $SmtAZn_4$. Exposure of SmtA to denaturing conditions resulted in an observed increase of the estimated CCS, in line with that observed in solution phase unfolding.

A combination of mass spectrometry and ion mobility mass spectrometry combined with zinc depletion experiments, using the chelator EDTA, revealed that the fully loaded species of SmtA was more compact than the zinc depleted species. Data obtained on the conformation of the zinc depleted species allowed a previously published SmtA zinc ion depletion model to be updated. The absence of a large change in CCS upon the removal of the final zinc ion, even at elevated energies, has led to two possible conclusions which require further investigation.

Significant steps have been taken towards the investigation of the possible interaction of SmtA with its metalloregulator SmtB. Differences in ionisation efficiency hampered efforts and a number of questions remain unanswered.

Chapter Six: Conclusions and Future work

6.1 Conclusions

A combination of MS and TWIMS experiments have been used to probe the structure, zinc stoichiometry and interaction partners of the cyanobacterial metalloregulator SmtB.

6.1.1 SmtB

An expression and purification protocol was developed to enable the production of full length zinc bound SmtB from *E. coli*. These samples allowed the investigation of SmtB produced in bacterial cytosol to be undertaken.

The zinc binding stoichiometry of the recombinant samples was determined using a combination of solution phase ICP-OES and gas phase ESI-MS experiments. The data obtained suggested a metal ion stoichiometry of more than three zinc ions per SmtB dimer. These results were not in agreement with the most recent literature which suggested that the predominant species was $SmtB_2Zn_2$, due to the action of uncooperative allostery (VanZile et al. 2002b; Eicken et al. 2003).

A fully loaded SmtB sample was compared with zinc depleted and zinc free samples using ESI-MS and TWIMS. Removal of zinc ions from SmtB either by EDTA or a reduction in pH produced an increasing intensity of monomeric species in addition to a small proportion of zinc depleted and zinc free dimeric species. A proportion of the monomeric species observed were highly charged suggesting an unfolding of the protein structure.

A comparison of calculated CCS obtained from conducting TWIMS experiments on all observable dimeric species demonstrated an increase in CCS in zinc depleted species. The broad ATD's observed for all stoichiometries suggest a high level of flexibility which is proposed to be due to the N-terminal domains and the beta-sheet regions.

Implementing a higher experimental cone voltage revealed a loss in dimeric stability upon zinc ion removal. A reduction in the solution phase dimeric association

constant had been observed previously in zinc free samples when compared to zinc loaded samples (Kar et al. 1997). The dissociation of dimeric apo-SmtB was proposed to be favoured in the gas phase, due to the strongly hydrophobic nature of the dimeric interface.

Tandem MS was used to investigate dimeric dissociation and the location of zinc binding residues. CID experiments produced fragments which were identified across the sequence of SmtB but in particular from the flexible N-terminal region. ETD experiments carried out on highly charged monomeric SmtB identified zinc bound fragments. The majority of these zinc bound fragments contained partial α5 sites whilst one possibly contained a partial α3N site.

6.1.2 SmtB:DNA complex

The interaction between SmtB and a proposed high affinity binding site was probed utilising ESI-MS. The S2/S1 binding site was successfully created from complementary single stranded oligonucleotides. The incubation of the dsDNA S2/S1 site with apo-SmtB dimer resulted in the observation of a $SmtB_2$:DNA complex with a stoichiometry of up to two dimers per binding site. A mixture of $SmtB_2Zn_4$ and the S2/S1 site failed to result in any observable protein:DNA complexes. These results were consistent with solution phase data.

6.1.3 SmtA

The metallothionein SmtA was studied due to its potential interaction with SmtB. The expression and purification of SmtA produced a full length protein with a zinc ion stoichiometry of $SmtAZn_4$. The use of TWIMS allowed the characterisation of zinc depleted SmtA conformations which, in agreement with a previously published mechanism, showed greater flexibility and an increase in collisional cross section upon zinc loss.

6.2 Future directions

This research has provided novel results in addition to complementary evidence surrounding research into metalloregulators and, in particular the protein SmtB. This area of research is still far from complete and would benefit from further mass spectrometry based experiments.

6.2.1 SmtB

A number of metal ions have been proposed to induce a conformational change in SmtB upon binding, including cadmium and copper (Morby et al. 1993), resulting in the derepression of the *smt* operon. The synthesis of these species, either by adapting the existing expression protocol or by the re-addition of metal ions to apo-SmtB, in combination with ESI-MS and IMS experiments may provide structural insights into the stoichiometry and conformational differences of these species.

6.2.2 SmtB:DNA complex

Aquiring IMS data from the two observable SmtB:DNA complexes could provide additional structural information. The CCS could be measured with an increased accuracy using the recently introduced protein calibration procedure (Bush et al. 2010). The addition of computational modelling and the use of the newly developed projected superposition approximation (PSA) (Bleiholder et al. 2011) to determine theoretical CCS of the complex would assist in determining the complexes architecture. The addition of zinc ions to the formed complex could provide a method of observing the complex dissociate and perhaps observing the biological mechanism in vitro.

A further potential avenue of interest would envolve the study of other metalloregulatory protein-DNA complexes. The MerR:DNA complex would be of particular interest as this protein is thought to undergo a conformational change, upon metal ion binding, but does not dissociate from its DNA (Guo et al. 2010). This would make it an ideal candidate for further ion mobility mass spectrometry-based studies.

6.2.3 SmtA

The ion mobility data obtained was found to agree with the existing model of zinc ion depletion. A few questions remain to be answered. These involve the conformational change associated with the removal of the final zinc ion, from the zinc finger motif of the protein. The stability of apo-SmtA and $SmtAZn_1$ could be probed further by implementing collisionally induced dissociation experiments. This may reveal if the protein is disordered or whether transient secondary structure is stabilising the protein in the gas phase. The combination of the information obtained from this type of experiment and molecular dynamic simulations could reveal whether the apo form has secondary structure or not.

Only limited progress was made towards characterising the interaction of SmtB with SmtA. Future steps would need to be taken to form apo-SmtA, either by dialysis or acidification and then gentle raising of the pH. The incubation of the resulting species with fully loaded SmtB or the production of the $SmtAZn_1$ intermediate by the addition of one molar equivalent of zinc ions could reveal if zinc transfer could occur between them and indeed whether an observable complex would be formed.

6.3 Concluding remarks

The use of native mass spectrometry coupled with the shape selective technique of ion mobility is now well establised as a complentary method to more traditional structural biology techniques such as NMR and X-ray crystallography. Data can be obtained quickly, at biologically relevant concentrations and with the added advantage of being able to probe heterogeneous mixtures whilst obtaining conformational information. In addition the conformational dynamics of specific regions within a protein can be probed utilising a hydrogen /deuterium exchange (HDX) MS work flow (Kaltashov et al. 2012).

Advances in instrumentation improving the dynamic range, sensitivity and resolution (both in MS and IMS) coupled with well developed experimental protocols will allow for detection of more subtle changes in individual proteins and protein mixtures. Improvements in determining kinetic and thermodynamic properties of

proteins and their complexes would also be supported by further instrument development (Hilton and Benesch 2012).

These advancements would significantly benefit the study of metalloproteins where binding and/or dissociation of small metal ions can lead to small but important changes in conformation or interaction partner.

References

Ahadi, E. and Konermann, L. (2012). Modeling the Behavior of Coarse-Grained Polymer Chains in Charged Water Droplets: Implications for the Mechanism of Electrospray Ionization. *The Journal of Physical Chemistry B.* **116**, 104-112.

Andreini, C., Bertini, I., Cavallaro, G., Holliday, G. and Thornton, J. (2008). Metal ions in biological catalysis: from enzyme databases to general principles. *Journal of Biological Inorganic Chemistry.* **13**, 1205-1218.

Arcella, A., Portella, G., Ruiz, M. L., Eritja, R., Vilaseca, M., Gabelica, V. and Orozco, M. (2012). Structure of Triplex DNA in the Gas Phase. *Journal of the American Chemical Society.* **134**, 6596-6606.

Armirotti, A., Benatti, U. and Damonte, G. (2009). Top-down proteomics with a quadrupole time-of-flight mass spectrometer and collision-induced dissociation. *Rapid Communications in Mass Spectrometry.* **23**, 661-666.

Arunkumar, A. I., Campanello, G. C. and Giedroc, D. P. (2009). Solution structure of a paradigm ArsR family zinc sensor in the DNA-bound state. *Proceeding of the National Academy of Sciences.* **106**, 18177-18182.

Atmanene, C., Petiot-Becard, S., Zeyer, D., Van Dorsselaer, A., Vivat Hannah, V. and Sanglier-Cianferani, S. (2012). Exploring Key Parameters to Detect Subtle Ligand-Induced Protein Conformational Changes Using Traveling Wave Ion Mobility Mass Spectrometry. *Analytical Chemistry.* **84**, 4703-4710.

Atmanene, C. d., Chaix, D., Bessin, Y., Declerck, N., Van Dorsselaer, A. and Sanglier-Cianferani, S. (2010). Combination of Noncovalent Mass Spectrometry and Traveling Wave Ion Mobility Spectrometry Reveals Sugar-Induced Conformational Changes of Central Glycolytic Genes Repressor/DNA Complex. *Analytical Chemistry.* **82**, 3597-3605.

Auld, D. S. (2001). Zinc coordination sphere in biochemical zinc sites. *BioMetals.* **14**, 271-313.

Badman, E. R., Myung, S. and Clemmer, D. E. (2005). Evidence for Unfolding and Refolding of Gas-Phase Cytochrome c Ions in a Paul Trap. *Journal of the American Society for Mass Spectrometry.* **16**, 1493-1497.

Baker, E. and Bowers, M. (2007). B-DNA Helix Stability in a Solvent-Free Environment. *Journal of the American Society for Mass Spectrometry.* **18**, 1188-1195.

Baker, E. S., Dupuis, N. F. and Bowers, M. T. (2009). DNA Hairpin, Pseudoknot, and Cruciform Stability in a Solvent-Free Environment. *The Journal of Physical Chemistry B.* **113**, 1722-1727.

Beck, J. L. (2011). Developments in Electrospray Ionization Mass Spectrometry of Non-Covalent DNA-Ligand Complexes. *Australian Journal of Chemistry.* **64**, 705-717.

Beck, J. L., Colgrave, M. L., Ralph, S. F. and Sheil, M. M. (2001). Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass Spectrometry Reviews.* **20**, 61-87.

Berezovskaya, Y., Armstrong, C. T., Boyle, A. L., Porrini, M., Woolfson, D. N. and Barran, P. E. (2011). Metal binding to a zinc-finger peptide: a comparison between solution and the gas phase. *Chemical Communications.* **47**, 412-414.

Bich, C., Baer, S., Jecklin, M. C. and Zenobi, R. (2010). Probing the Hydrophobic Effect of Noncovalent Complexes by Mass Spectrometry. *Journal of the American Society for Mass Spectrometry.* **21**, 286-289.

Biosciences, A. (2003). PD-10 Desalting column intsructions.

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.

Blencowe, D. K. and Morby, A. P. (2003). Zn(II) metabolism in prokaryotes. *FEMS Microbiology Reviews.* **27**, 291-311.

Blindauer, C. (2011). Bacterial metallothioneins: past, present, and questions for the future. *Journal of Biological Inorganic Chemistry.* **16**, 1011-1024.

Blindauer, C., Razi, M., Campopiano, D. and Sadler, P. (2007). Histidine ligands in bacterial metallothionein enhance cluster stability. *Journal of Biological Inorganic Chemistry.* **12**, 393-405.

Blindauer, C. A., Harrison, M. D., Parkinson, J. A., Robinson, A. K., Cavet, J. S., Robinson, N. J. and Sadler, P. J. (2001). A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity. *Proceedings of the National Academy of Sciences of the United States of America.* **98**, 9593-9598.

Blindauer, C. A., Harrison, M. D., Robinson, A. K., Parkinson, J. A., Bowness, P. W., Sadler, P. J. and Robinson, N. J. (2002). Multiple bacteria encode metallothioneins and SmtA-like zinc fingers. *Molecular Microbiology.* **45**, 1421-1432.

Blindauer, C. A. and Leszczyszyn, O. I. (2010). Metallothioneins: unparalleled diversity in structures and functions for metal ion homeostasis and more. *Natural Product Reports.* **27**, 720-741.

Boeri Erba, E., Barylyuk, K., Yang, Y. and Zenobi, R. (2011). Quantifying Protein–Protein Interactions Within Noncovalent Complexes Using Electrospray Ionization Mass Spectrometry. *Analytical Chemistry.* **83**, 9251-9259.

Breuker, K., Brüschweiler, S. and Tollinger, M. (2010). Electrostatic Stabilization of a Native Protein Structure in the Gas Phase. *Angewandte Chemie International Edition.* **50**, 873-877.

Breuker, K. and McLafferty, F. W. (2008). Stepwise evolution of protein native structure with electrospray into the gas phase, 10(-12) to 10(2) S. *Proceedings of the National Academy of Sciences of the United States of America.* **105**, 18145-18152.

Brewer, T. M. and Marcus, R. K. (2007). Determination of "free" iron and iron bound in metalloproteins via liquid chromatography separation and inductively coupled plasma-optical emission spectroscopy (LC-ICP-OES) and particle beam/hollow cathode-optical emission spectroscopy (LC-PB/HC-OES) techniques. *Journal of Analytical Atomic Spectrometry.* **22**.

Brocklehurst, K. R. and Morby, A. P. (2000). Metal-ion tolerance in Escherichia coli: analysis of transcriptional profiles by gene-array technology. *Microbiology-Uk.* **146**, 2277-2282.

Busenlehner, L. S. and Giedroc, D. P. (2006). Kinetics of metal binding by the toxic metal-sensing transcriptional repressor Staphylococcus aureus pI258 CadC. *Journal of Inorganic Biochemistry.* **100**, 1024-1034.

Busenlehner, L. S., Pennella, M. A. and Giedroc, D. P. (2003). The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. *FEMS Microbiology Reviews.* **27**, 131-143.

Busenlehner, L. S., Weng, T.-C., Penner-Hahn, J. E. and Giedroc, D. P. (2002). Elucidation of Primary ([alpha]3N) and Vestigial ([alpha]5) Heavy Metal-binding Sites in Staphylococcus aureus pI258 CadC: Evolutionary Implications for Metal Ion Selectivity of ArsR/SmtB Metal Sensor Proteins. *Journal of Molecular Biology.* **319**, 685-701.

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.

Campbell, D. R., Chapman, K. E., Waldron, K. J., Tottey, S., Kendall, S., Cavallaro, G., Andreini, C., Hinds, J., Stoker, N. G., Robinson, N. J. and Cavet, J. S. (2007). Mycobacterial Cells Have Dual Nickel-Cobalt Sensors. *Journal of Biological Chemistry.* **282**, 32298-32310.

Canutescu, A. A., Shelenkov, A. A. and Dunbrack, J. R. L. (2003). A graph theory algorithm for protein side-chain prediction. . *Protein Science.* **12**, 2001-2014.

Cavet, J. S., Borrelly, G. P. M. and Robinson, N. J. (2003). Zn, Cu and Co in cyanobacteria: selective control of metal availability. *FEMS Microbiology Reviews.* **27**, 165-181.

Cech, N. B. and Enke, C. G. (2000). Relating Electrospray Ionization Response to Nonpolar Character of Small Peptides. *Analytical Chemistry.* **72**, 2717-2723.

Chakravorty, D. K., Wang, B., Lee, C. W., Giedroc, D. P. and Merz, K. M. (2012). Simulations of Allosteric Motions in the Zinc Sensor CzrA. *Journal of the American Chemical Society*. **134**, 3367-3376.

Chary, K. V. R. and Govil, G. (2008). NMR in Biological Systems - From Molecules to Humans, Springer.

Chernushevich, I. V. and Thomson, B. A. (2004). Collisional Cooling of Large Ions in Electrospray Mass Spectrometry. *Analytical Chemistry*. **76**, 1754-1760.

Chowdhury, S. K., Katta, V. and Chait, B. T. (1990). Probing Conformational Changes in Proteins by Mass Spectrometry. *Journal of the American Chemical Society*. **112**, 9012-9013.

Clemmer, D. E. www.indiana.edu/~clemmer/Research/crosssectiondatabase/Oligonucleotides/oligonucleotides_cs.htm
Retrieved 27th April 2012.

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.

Clore, G. M. and Gronenborn, A. M. (1998). NMR structure determination of proteins and protein complexes larger than 20 kDa. *Current Opinion in Chemical Biology*. **2**, 564-570.

Clowers, B. H., Dwivedi, P., Steiner, W. E., Hill, J. H. H. and Bendiak, B. (2005). Separation of Sodiated Isobaric Disaccharides and Trisaccharides Using Electrospray Ionization-Atmospheric Pressure Ion Mobility-Time of Flight Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*. **16**, 660-669.

Cobine, P., Wickramasinghe, W. A., Harrison, M. D., Weber, T., Solioz, M. and Dameron, C. T. (1999). The Enterococcus hirae copper chaperone CopZ delivers copper(I) to the CopY repressor. *Febs Letters*. **445**, 27-30.

Cobine, P. A., George, G. N., Jones, C. E., Wickramasinghe, W. A., Solioz, M. and Dameron, C. T. (2002). Copper transfer from the Cu(I) chaperone, CopZ, to the repressor, Zn(II)CopY: Metal coordination environments and protein interactions. *Biochemistry*. **41**, 5822-5829.

Cook, W. J., Kar, S. R., Taylor, K. B. and Hall, L. M. (1998). Crystal structure of the cyanobacterial metallothionein repressor SmtB: a model for metalloregulatory proteins. *Journal of Molecular Biology*. **275**, 337-346.

Cooke, R. M. (1997). Protein NMR extends into new fields of structural biology. *Current Opinion in Chemical Biology*. **1**, 359-364.

Coon, J. J., Syka, J. E. P., Schwartz, J. C., Shabanowitz, J. and Hunt, D. F. (2004). Anion dependence in the partitioning between proton and electron transfer in ion/ion reactions. *International Journal of Mass Spectrometry*. **236**, 33-42.

Costello, L. C., Guan, Z., Franklin, R. B. and Feng, P. (2004). Metallothionein can function as a chaperone for zinc uptake transport into prostate and liver mitochondria. *Journal of Inorganic Biochemistry.* **98**, 664-666.

Cotter, R. J. (1999). Peer Reviewed: The New Time-of-Flight Mass Spectrometry. *Analytical Chemistry.* **71**, 445A-451A.

Cox, E. H. and McLendon, G. L. (2000). Zinc-dependent protein folding. *Current Opinion in Chemical Biology.* **4**, 162-165.

Creaser, C. S., Griffiths, J. R., Bramwell, C. J., Noreen, S., Hill, C. A. and Thomas, C. L. P. (2004). Ion mobility spectrometry: a review. Part 1. Structural analysis by mobility measurement. *Analyst.* **129**, 984-994.

Cristoni, S. and Bernardi, L. R. (2003). Development of new methodologies for the mass spectrometry study of bioorganic macromolecules. *Mass Spectrometry Reviews.* **22**, 369-406.

Daniels, M. J., Turner-Cavet, J. S., Selkirk, R., Sun, H., Parkinson, J. A., Sadler, P. J. and Robinson, N. J. (1998). Coordination of Zn2+ (and Cd2+) by Prokaryotic Metallothionein. Involement of His-Imidazole. *J. Biol. Chem.* **273**, 22957-22961.

De Hoffmann, E. and Stroobant, V. (2007). Mass Spectrometry Principles and Applications, Wiley.

Deterding, L. J., Kast, J., Przybylski, M. and Tomer, K. B. (2000). Molecular Characterization of a Tetramolecular Complex between dsDNA and a DNA-Binding Leucine Zipper Peptide Dimer by Mass Spectrometry. *Bioconjugate Chemistry.* **11**, 335-344.

Dodds, E. D. (2012). Gas-phase dissociation of glycosylated peptide ions. *Mass Spectrometry Reviews.* In Press.

Dodds, E. D., Blackwell, A. E., Jones, C. M., Holso, K. L., O'Brien, D. J., Cordes, M. H. J. and Wysocki, V. H. (2011). Determinants of Gas-Phase Disassembly Behavior in Homodimeric Protein Complexes with Related Yet Divergent Structures. *Analytical Chemistry.* **83**, 3881-3889.

Douglas, D. J., Frank, A. J. and Mao, D. M. (2005). Linear ion traps in mass spectrometry. *Mass Spectrometry Reviews.* **24**, 1-29.

Eckers, C., Laures, A. M. F., Giles, K., Major, H. and Pringle, S. (2007). Evaluating the utility of ion mobility separation in combination with high-pressure liquid chromatography/mass spectrometry to facilitate detection of trace impurities in formulated drug products. *Rapid Communications in Mass Spectrometry.* **21**, 1255-1263.

Eicken, C., Pennella, M. A., Chen, X., Koshlap, K. M., VanZile, M. L., Sacchettini, J. C. and Giedroc, D. P. (2003). A Metal-Ligand-mediated Intersubunit Allosteric Switch in Related SmtB/ArsR Zinc Sensor Proteins. *Journal of Molecular Biology.* **333**, 683-695.

Eide, D. J. (2006). Zinc transporters and the cellular trafficking of zinc. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research.* **1763**, 711-722.

Ellman, G. L. (1959). Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics.* **82**, 70-77.

Endo, G. and Silver, S. (1995). CadC, the Transcriptional Regulatory Protein of the Cadmium Resistance System of Staphylococcus aureus Plasmid pI258. *Journal of Bacteriology.* **177**, 4437-4441.

Erales, J., Gontero, B., Whitelegge, J. and Halgand, F. (2009). Mapping of a copper-binding site an the small CP12 chloroplastic protein of Chlamydomonas reinhardtii using top-down mass spectrometry and site-directed mutagenesis. *Biochemical Journal.* **419**, 75-82.

Erbe, J. L., Taylor, K. B. and Hall, L. M. (1995). Metalloregulation of the cyanobacterial smt locus: indentification of SmtB binding sites and direct interaction with metals. *Nucl. Acids Res.* **23**, 2472-2478.

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.

Fedor, M. J. (2002). The role of metal ions in RNA catalysis. *Current Opinion in Structural Biology.* **12**, 289-295.

Feng, X., Liu, X., Luo, Q. and Liu, B.-F. (2008). Mass spectrometry in systems biology: An overview. *Mass Spectrometry Reviews.* **27**, 635-660.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science.* **246**, 64-71.

Finney, L. A. and O'Halloran, T. V. (2003). Transition metal speciation in the cell: Insights from the chemistry of metal ion receptors. *Science.* **300**, 931-936.

Fowler, B. A., Hildebrand, C. E., Kojima, Y. and Webb, M. (1987). Nomenclature of metallothionein. *Experientia. Supplementum.* **52**, 19-22.

Gabelica, V., Galic, N., Rosu, F., Houssier, C. and De Pauw, E. (2003). Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. *Journal of Mass Spectrometry.* **38**, 491-501.

Gabelica, V. r., Shammel Baker, E., Teulade-Fichou, M.-P., De Pauw, E. and Bowers, M. T. (2007). Stabilization and Structure of Telomeric and c-myc Region Intramolecular G-Quadruplexes:The Role of Central Cations and Small Planar Ligands. *Journal of the American Chemical Society.* **129**, 895-904.

Gasteiger E., H. C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005). The Proteomics Protocols Handbook, Protein Identification and Analysis Tools on the ExPASy Server, Humana Press

Gattiker, A., Bienvenut, W. V., Bairoch, A. and Gasteiger, E. (2002). FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification. *Proteomics.* **2**, 1435-1444.

GEHealthcare (2007). HiTrap affinity columns Instructions.

Gehrig, P. M., You, C., Dallinger, R., Gruber, C., Brouwer, M., Kagi, J. H. and Hunziker, P. E. (2000). Electrospray ionization mass spectrometry of zinc, cadmium, and copper metallothioneins: evidence for metal-binding cooperativity. *Protein Sci.* **9**, 395-402.

Gidden, J., Baker, E. S., Ferzoco, A. and Bowers, M. T. (2005). Structural motifs of DNA complexes in the gas phase. *International Journal of Mass Spectrometry.* **240**, 183-193.

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.

Giles, K., Williams, J. P. and Campuzano, I. (2011). Enhancements in travelling wave ion mobility resolution. *Rapid Communications in Mass Spectrometry.* **25**, 1559-1566.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research.* **38**, W695-W699.

Gourdon, P., Liu, X.-Y., Skjorringe, T., Morth, J. P., Moller, L. B., Pedersen, B. P. and Nissen, P. (2011). Crystal structure of a copper-transporting PIB-type ATPase. *Nature.* **475**, 59-64.

Grass, G., Wong, M. D., Rosen, B. P., Smith, R. L. and Rensing, C. (2002). ZupT is a Zn(II) uptake system in Escherichia coli. *Journal of Bacteriology.* **184**, 864-866.

Guedon, E. and Helmann, J. D. (2003). Origins of metal ion selectivity in the DtxR/MntR family of metalloregulators. *Molecular Microbiology.* **48**, 495-506.

Guerra, A. J. and Giedroc, D. P. (2012). Metal site occupancy and allosteric switching in bacterial metal sensor proteins. *Archives of Biochemistry and Biophysics.* **519**, 210-222.

Guharay, S. K., Dwivedi, P. and Hill, H. H. (2008). Ion mobility spectrometry: Ion source development and applications in physical and biological sciences. *Ieee Transactions on Plasma Science.* **36**, 1458-1470.

Gumerov, D. R. and Kaltashov, I. A. (2001). Dynamics of Iron Release from Transferrin N-Lobe Studied by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **73**, 2565-2570.

Gumerov, D. R., Mason, A. B. and Kaltashov, I. A. (2003). Interlobe Communication in Human Serum Transferrin: Metal Binding and Conformational Dynamics Investigated by Electrospray Ionization Mass Spectrometry. *Biochemistry.* **42**, 5421-5428.

Guo, H.-B., Johs, A., Parks, J. M., Olliff, L., Miller, S. M., Summers, A. O., Liang, L. and Smith, J. C. (2010). Structure and Conformational Dynamics of the Metalloregulator MerR upon Binding of Hg(II). *Journal of Molecular Biology*. **398**, 555-568.

Guo, X., Bruist, M. F., Davis, D. L. and Bentzley, C. M. (2005). Secondary structural characterization of oligonucleotide strands using electrospray ionization mass spectrometry. *Nucleic Acids Research.* **33**, 3659-3666.

Gupta, A., Morby, A. P., Turner, J. S., Whitton, B. A. and Robinson, N. J. (1993). Deletion within the metallothionein locus of cadmium-tolerant Synechococcus PCC 6301 involving a highly iterated palindrome (HIP1). *Molecular Microbiology.* **7**, 189-195.

Gupta, A., Whitton, B. A., Morby, A. P., Huckle, J. W. and Robinson, N. J. (1992). Amplification and Rearrangement of a Prokaryotic Metallothionein Locus smt in Synechococcus PCC 6301 Selected for Tolerance to Cadmium. *Proceedings: Biological Sciences.* **248**, 273-281.

Gupta, R., Kapur, A., Beck, J. L. and Sheil, M. M. (2001). Positive ion electrospray ionization mass spectrometry of double-stranded DNA/drug complexes. *Rapid Communications in Mass Spectrometry.* **15**, 2472-2480.

Hall, Z., Politis, A., Bush, M. F., Smith, L. J. and Robinson, C. V. (2012). Charge-State Dependent Compaction and Dissociation of Protein Complexes: Insights from Ion Mobility and Molecular Dynamics. *Journal of the American Chemical Society.* **134**, 3429-3438.

Han, X., Jin, M., Breuker, K. and McLafferty, F. W. (2006). Extending Top-Down Mass Spectrometry to Proteins with Masses Greater Than 200 Kilodaltons. *Science.* **314**, 109-112.

Hantke, K. (2001). Bacterial zinc transporters and regulators. *BioMetals.* **14**, 239-249.

Hantke, K. (2005). Bacterial zinc uptake and regulators. *Current Opinion in Microbiology.* **8**, 196-202.

Harrison, M. D., Jones, C. E., Solioz, M. and Dameron, C. T. (2000). Intracellular copper routing: the role of copper chaperones. *Trends in Biochemical Sciences.* **25**, 29-32.

Hartmer, R., Kaplan, D. A., Gebhardt, C. R., Ledertheil, T. and Brekenfeld, A. (2008). Multiple ion/ion reactions in the 3D ion trap: Selective reagent anion production for ETD and PTR from a single compound. *International Journal of Mass Spectrometry.* **276**, 82-90.

Heck, A. J. R. (2008). Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods.* **5**, 927-933.

Heck, A. J. R. and Van den Heuvel, R. H. H. (2004). Investigation of intact protein complexes by mass spectrometry. *Mass Spectrometry Reviews.* **23**, 368-389.

Helbig, K., Grosse, C. and Nies, D. (2008). Cadmium Toxicity in Glutathione Mutations of *Escherichia coli*. *Journal of Bacteriology.* **190**, 5439-5454.

Hilton, G. R. and Benesch, J. L. P. (2012). Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of the Royal Society Interface.* **9**, 801-816.

Hoaglund, C. S., Liu, Y. S., Ellington, A. D., Pagel, M. and Clemmer, D. E. (1997). Gas-phase DNA: Oligothymidine ion conformers. *Journal of the American Chemical Society.* **119**, 9051-9052.

Hopper, J. T. S. 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.

Hu, P., Ye, Q.-Z. and Loo, J. A. (1994). Calcium Stoichiometry Determination for Calcium Binding Proteins by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **66**, 4190-4194.

Huckle, J. W., Morby, A. P., Turner, J. S. and Robinson, N. J. (1993). Isolation of a prokaryotic metallothionein locus and analysis of transcriptional control by trace metal ions. *Molecular Microbiology.* **7**, 177-187.

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.

Irving, H. and Williams, R. J. P. (1953). 637. The stability of transition-metal complexes. *Journal of the Chemical Society (Resumed).* 3192-3210.

Jackson, S. N., Ugarov, M., Egan, T., Post, J. D., Langlais, D., Schultz, J. A. and Woods, A. S. (2007). MALDI-ion mobility-TOFMS imaging of lipids in rat brain tissue. *Journal of Mass Spectrometry.* **42**, 1093-1098.

Jarrold, M. F. (1999). Unfolding, refolding, and hydration of proteins in the gas phase. *Accounts of Chemical Research.* **32**, 360-367.

Jarrold, M. F. (2000). Peptides and proteins in the vapor phase. *Annual Review of Physical Chemistry.* **51**, 179-207.

Jennings, K. R. (1968). Collision-induced decompositions of aromatic molecular ions. *International Journal of Mass Spectrometry and Ion Physics.* **1**, 227-235.

Jensen, L. T., Peltier, J. M. and Winge, D. R. (1998). Identification of a four copper folding intermediate in mammalian copper metallothionein by electrospray ionization mass spectrometry. *Journal of Biological Inorganic Chemistry.* **3**, 627-631.

Jonscher, K. R. and Yates, J. R. (1997). The quadrupole ion trap mass spectrometer - A small solution to a big challenge. *Analytical Biochemistry.* **244**, 1-15.

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.

Kaltashov, I., Zhang, M., Eyles, S. and Abzalimov, R. (2006). Investigation of structure, dynamics and function of metalloproteins with electrospray ionization mass spectrometry. *Analytical and Bioanalytical Chemistry.* **386**, 472-481.

Kaltashov, I. A., Bobst, C. E., Zhang, M., Leverence, R. and Gumerov, D. R. (2012). Transferrin as a model system for method development to study structure, dynamics and interactions of metalloproteins using mass spectrometry. *Biochimica et Biophysica Acta (BBA) - General Subjects.* **1820**, 417-426.

Kanu, A. B., Dwivedi, P., Tam, M., Matz, L. and Hill, Jr, H. H. (2008). Ion mobility-mass spectrometry. *Journal of Mass Spectrometry.* **43**, 1-22.

Kapur, A., Beck, J. L., Brown, S. E., Dixon, N. E. and Sheil, M. M. (2002). Use of electrospray ionization mass spectrometry to study binding interactions between a replication terminator protein and DNA. *Protein Science.* **11**, 147-157.

Kar, S. R., Adams, A. C., Lebowitz, J., Taylor, K. B. and Hall, L. M. (1997). The cyanobacterial repressor SmtB is predominantly a dimer and binds two Zn2+ ions per subunit. *Biochemistry.* **36**, 15343-15348.

Kar, S. R., Lebowitz, J., Blume, S., Taylor, K. B. and Hall, L. M. (2001). SmtB-DNA and Protein-Protein Interactions in the Formation of the Cyanobacterial Metallothionein Repression Complex: Zn2+ Does Not Dissociate the Protein-DNA Complex in Vitro. *Biochemistry.* **40**, 13378-13389.

Karas, M., Bachmann, D., Bahr, U. and Hillenkamp, F. (1987). Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes.* **78**, 53-68.

Karas, M. and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry.* **60**, 2299-2301.

Kebarle, P. and Peschke, M. (2000). On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions. *Analytica Chimica Acta.* **406**, 11-35.

Keltner, Z., Meyer, J. A., Johnson, E. M., Palumbo, A. M., Spence, D. M. and Reid, G. E. (2010). Mass spectrometric characterization and activity of zinc-activated proinsulin C-peptide and C-peptide mutants. *Analyst.* **135**, 278-288.

Knochenmuss, R. (2006). Ion formation mechanisms in UV-MALDI. *Analyst.* **131**, 966-986.

Konermann, L., Rosell, F. I., Mauk, A. G. and Douglas, D. J. (1997). Acid-induced denaturation of myoglobin studied by time-resolved electrospray ionization mass spectrometry. *Biochemistry.* **36**, 6448-6454.

Koomen, J., Ruotolo, B., Gillig, K., McLean, J., Russell, D., Kang, M., Dunbar, K., Fuhrer, K., Gonin, M. and Schultz, A. (2002). Oligonucleotide analysis with MALDI-ion-mobility-TOFMS. *Analytical and Bioanalytical Chemistry.* **373**, 612-617.

Koppenaal, D. W., Barinaga, C. J., Denton, M. B., Sperline, R. P., Hieftje, G. M., Schilling, G. D., Andrade, F. J., Barnes, J. H. and Iv, I. V. (2005). MS Detectors. *Analytical Chemistry.* **77**, 418 A-427 A.

Kosada, T., Morita, E. H., Miura, A., Yamazaki, T., Hayashi, H. and Kyogoku, Y. (1999). Backbone NMR assignments of a cyanobacterial transcriptional factor, SmtB, that binds zinc ions. *Journal of Biomolecular NMR.* **14**, 191-192.

Krezel, A., Lesniak, W., Jezowska-Bojczuk, M., Mlynarz, P., Brasun, J., Kozlowski, H. and Bal, W. (2001). Coordination of heavy metals by dithiothreitol, a commonly used thiol group protectant. *Journal of Inorganic Biochemistry.* **84**, 77-88.

Lane, E., J. Holden, A. and A. Coward, R. (1999). Determination of copper and zinc in blood plasma by ion chromatography using a cobalt internal standard. *Analyst.* **124**, 245-249.

Lane, T. W., Saito, M. A., George, G. N., Pickering, I. J., Prince, R. C. and Morel, F. M. M. (2005). A cadmium enzyme from a marine diatom. *Nature.* **435**, 42.

Leal, S. S., Botelho, H. M. and Gomes, C. M. (2012). Metal ions as modulators of protein conformation and misfolding in neurodegeneration. *Coordination Chemistry Reviews.* **256**, 2253-2270.

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. B. (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.

Lee, S., Arunkumar, A. I., Chen, X. and Giedroc, D. P. (2006). Structural Insights into Homo- and Heterotropic Allosteric Coupling in the Zinc Sensor S. aureus CzrA from Covalently Fused Dimers. *Journal of the American Chemical Society.* **128**, 1937-1947.

Lei, Q. P., Cui, X., Kurtz, D. M., Amster, I. J., Chernushevich, I. V. and Standing, K. G. (1998). Electrospray Mass Spectrometry Studies of Non-Heme Iron-Containing Proteins. *Anal. Chem.* **70**, 1838-1846.

Leszczyszyn, O. I., Evans, C. D., Keiper, S. E., Warren, G. Z. L. and Blindauer, C. A. (2007). Differential reactivity of individual zinc ions in clusters from bacterial metallothioneins. *Inorganica Chimica Acta.* **360**, 3-13.

Leszczyszyn, O. I., White, C. R. J. and Blindauer, C. A. (2010). The isolated Cys2His2 site in EC metallothionein mediates metal-specific protein folding. *Molecular BioSystems.* **6**.

Light-Wahl, K. J., Schwartz, B. L. and Smith, R. D. (1994). Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. *Journal of the American Chemical Society.* **116**, 5271-5278.

Liu, J., Dutta, S. J., Stemmler, A. J. and Mitra, B. (2005). Metal-Binding Affinity of the Transmembrane Site in ZntA: Implications for Metal Selectivity. *Biochemistry.* **45**, 763-772.

Liu, X. and Theil, E. C. (2005). Ferritins:Dynamic Management of Biological Iron and Oxygen Chemistry. *Accounts of Chemical Research.* **38**, 167-175.

Loo, J. A. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews.* **16**, 1-23.

Loo, J. A., Berhane, B., Kaddis, C. S., Wooding, K. M., Xie, Y., Kaufman, S. L. and Chernushevich, I. V. (2005). Electrospray Ionization Mass Spectrometry and Ion Mobility Analysis of the 20S Proteasome Complex. *Journal of the American Society for Mass Spectrometry.* **16**, 998-1008.

Loo, J. A., Hu, P. F. and Smith, R. D. (1994). Interaction of Angiotensin Peptides and Zinc Metal Ions Probed by Electrospray Ionisation Mass Spectrometry. *Journal of the American Society for Mass Spectrometry.* **5**, 959-965.

Loo, J. A., Loo, R. R. O., Udseth, H. R., Edmonds, C. G. and Smith, R. D. (1991). Solvent-induced conformational changes of polypeptides probed by electrospray-ionization mass spectrometry. *Rapid Communications in Mass Spectrometry.* **5**, 101-105.

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

Loo, R. R. O., Goodlett, D. R., Smith, R. D. and Loo, J. A. (1993). Observation of a noncovalent ribonuclease S-protein/S-peptide complex by electrospray ionization mass spectrometry. *Journal of the American Chemical Society.* **115**, 4391-4392.

Lorenzen, K., Olia, A. S., Uetrecht, C., Cingolani, G. and Heck, A. J. R. (2008). Determination of Stoichiometry and Conformational Changes in the First Step of the P22 Tail Assembly. *Journal of Molecular Biology.* **379**, 385-396.

Lorenzen, K., Vannini, A., Cramer, P. and Heck, A. J. R. (2007). Structural Biology of RNA Polymerase III: Mass Spectrometry Elucidates Subcomplex Architecture. *Structure.* **15**, 1237-1245.

Luh, F. Y., Archer, S. J., Domaille, P. J., Smith, B. O., Owen, D., Brotherton, D. H., Raine, A. R. C., Xu, X., Brizuela, L., Brenner, S. L. and Laue, E. D. (1997). Structure of the cyclin-dependent kinase inhibitor p19Ink4d. *Nature.* **389**, 999-1003.

Ma, Z., Jacobsen, F. E. and Giedroc, D. P. (2009). Coordination Chemistry of Bacterial Metal Transport and Sensing. *Chemical Reviews.* **109**, 4644-4681.

Mamyrin, B. A. (2001). Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry.* **206**, 251-266.

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure.* **29**, 291-325.

Mattapalli, H., Monteith, W., Burns, C. and Danell, A. (2009). Zinc deposition during ESI-MS analysis of peptide-zinc complexes. *Journal of The American Society for Mass Spectrometry.* **20**, 2199-2205.

Matz, L. M., Asbury, G. R. and Hill, H. H. (2002). Two-dimensional separations with electrospray ionization ambient pressure high-resolution ion mobility spectrometry/quadrupole mass spectrometry. *Rapid Communications in Mass Spectrometry.* **16**, 670-675.

McIntyre, D. (2004). Advantages of Wide Dynamic Range on an Orthogonal Acceleration Time-of-Flight Mass Spectrometer. *Technical Overview, Agilent Technologies.*

McLafferty, F. W. and Bryce, T. A. (1967). Metastable-ion characteristics: characterization of isomeric molecules. *Chemical Communications (London).* 1215-1217.

Merenbloom, S., Flick, T. and Williams, E. (2012). How Hot are Your Ions in TWAVE Ion Mobility Spectrometry? *Journal of The American Society for Mass Spectrometry.* **23**, 553-562.

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. *Journal of Physical Chemistry.* **100**, 16082-16086.

Mirgorodskaya, O. A., Shevchenko, A. A., Chernushevich, I. V., Dodonov, A. F. and Miroshnikov, A. I. (1994). Electrospray-ionization time-of-flight mass spectrometry in protein chemistry. *Analytical Chemistry.* **66**, 99-107.

Mirza, U. A., Cohen, S. L. and Chait, B. T. (1993). Heat-induced conformational changes in proteins studied by electrospray ionization mass spectrometry. *Analytical Chemistry.* **65**, 1-6.

Moradian, A., Scalf, M., Westphall, M. S., Smith, L. M. and Douglas, D. J. (2002). Collision cross sections of gas phase DNA ions. *International Journal of Mass Spectrometry.* **219**, 161-170.

Morby, A. P., Turner, J. S., Huckle, J. W. and Robinson, N. J. (1993). SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene smtA: identification of a Zn inhibited DNA-protein complex. *Nucl. Acids Res.* **21**, 921-925.

Morsa, D., Gabelica, V. r. and De Pauw, E. (2011). Effective Temperature of Ions in Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry.* **83**, 5775-5782.

Mukhopadhyay, R. (2008). IMS/MS: its time has come. *Analytical Chemistry.* **80**, 7918-7920.

Myung, S., Badman, E. R., Lee, Y. J. and Clemmer, D. E. (2002). Structural Transitions of Electrosprayed Ubiquitin Ions Stored in an Ion Trap over Similar to 10 ms to 30 s. *The Journal of Physical Chemistry A.* **106**, 9976-9982.

Namuswe, F. and Berg, J. M. (2012). Secondary interactions involving zinc-bound ligands: Roles in structural stabilization and macromolecular interactions. *Journal of Inorganic Biochemistry.* **111**, 146-149.

Neilands, J. B. (1995). Siderophores: Structure and Function of Microbial Iron Transport Compounds. *Journal of Biological Chemistry.* **270**, 26723-26726.

Nies, D. and Silver, S. (2007). Bacterial Transition Metal Homeostasis
Molecular Microbiology of Heavy Metals, Springer Berlin / Heidelberg. **6:** 117-142.

Nordhoff, E., Kirpekar, F. and Roepstorff, P. (1996). Mass spectrometry of nucleic acids. *Mass Spectrometry Reviews.* **15**, 67-138.

Novagen (2004). User protocol.

Novichkov, P. S., Laikova, O. N., Novichkova, E. S., Gelfand, M. S., Arkin, A. P., Dubchak, I. and Rodionov, D. A. (2010). RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. *Nucleic Acids Research.* **38**, D111-D118.

Nyborg, J. K. and Peersen, O. B. (2004). That zincing feeling: the effects of EDTA on the behaviour of zinc-binding transcriptional regulators. *The Biochemical journal.* **381**, e3-4.

Olafson, R. W. (1986). Physiological and Chemical Characterization of Cyanobacterial Metallothioneins. *Environmental Health Perspectives.* **65**, 71-75.

Olafson, R. W., McCubbin, W. D. and Kay, C. M. (1988). Primary- and secondary-structural analysis of a unique prokaryotic metallothionein from a Synechococcus sp. cyanobacterium. . *Biochemical Journal.* **251**, 691-699.

Outten, C. E., O'Halloran and Thomas, V. (2001). Femtomolar Sensitivity of Metalloregulatory Proteins Controlling Zinc Homeostasis. *Science.* **292**, 2488-2492.

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Science.* **4**, 2411-2423.

Pan, P. and McLuckey, S. A. (2003a). The effect of small cations on the positive electrospray responses of proteins at low pH. *Analytical Chemistry.* **75**, 5468-5474.

Pan, P. and McLuckey, S. A. (2003b). Electrospray ionization of protein mixtures at low pH. *Analytical Chemistry.* **75**, 1491-1499.

Patzer, S. I. and Hantke, K. (1998). The ZnuABC high-affinity zinc uptake system and its regulator Zur in Escherichia coli. *Molecular Microbiology.* **28**, 1199-1210.

Paul, W. (1990). Electromagnetic Traps for Charged and Neutral Particals. *Reviews of Modern Physics.* **62**, 531-540.

Pennella, M. A. and Giedroc, D. P. (2005). Structural Determinants of Metal Selectivity in Prokaryotic Metal-responsive Transcriptional Regulators. *Biometals.* **18**, 413-428.

Politis, A., Park, A. Y., Hyung, S.-J., Barsky, D., Ruotolo, B. T. and Robinson, C. V. (2010). Integrating Ion Mobility Mass Spectrometry with Molecular Modelling to Determine the Architecture of Multiprotein Complexes. *PLoS ONE.* **5**, e12080.

Pribnow, D. (1975). Nucleotide Sequence of an RNA Polymerase Binding Site at an Early T7 Promoter. *Proceedings of the National Academy of Sciences of the United States of America.* **72**, 784-788.

Pringle, S. D., Giles, K., Wildgoose, J. L., Williams, J. P., Slade, S. E., Thalassinos, 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.

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.

Pusey, M. L., Liu, Z.-J., Tempel, W., Praissman, J., Lin, D., Wang, B.-C., Gavira, J. A. and Ng, J. D. (2005). Life in the fast lane for protein crystallization and X-ray crystallography. *Progress in Biophysics and Molecular Biology.* **88**, 359-386.

Rensing, C., Mitra, B., Nies, D. and Silver, S. (2007). Zinc, Cadmium, and Lead Resistance and Homeostasis
Molecular Microbiology of Heavy Metals, Springer Berlin / Heidelberg. **6:** 321-341.

Reyes-Caballero, H., Campanello, G. C. and Giedroc, D. P. Metalloregulatory proteins: Metal selectivity and allosteric switching. *Biophysical Chemistry.* **156**, 103-114.

Rhoads, T. W., Lopez, N. I., Zollinger, D. R., Morre, J. T., Arbogast, B. L., Maier, C. S., DeNoyer, L. and Beckman, J. S. (2011). Measuring copper and zinc superoxide dismutase from spinal cord tissue using electrospray mass spectrometry. *Anal. Chem.* **415**, 52-58.

Robinson, C. V., Chung, E. W., Kragelund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M. and Dobson, C. M. (1996). Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein-CoA Ligand Binding and Assembly. *Journal of the American Chemical Society.* **118**, 8646-8653.

Robinson, N. J., Gupta, A., Fordhamskelton, A. P., Croy, R. R. D., Whitton, B. A. and Huckle, J. W. (1990). Prokaryotic Metallothionein Gene Characterization and Expression - Chromosome Crawling by Ligation-Mediated PCR. *Proceedings of the Royal Society of London Series B-Biological Sciences.* **242**, 241-247.

Roepstorff, P. and Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biological Mass Spectrometry.* **11**, 601-601.

Rosu, F., Gabelica, V., Houssier, C., Colson, P. and De Pauw, E. (2002). Triplex and quadruplex DNA structures studied by electrospray mass spectrometry. *Rapid Communications in Mass Spectrometry.* **16**, 1729-1736.

Rozman, M. and Gaskell, S. J. (2011). Charge state dependent top-down characterisation using electron transfer dissociation. *Rapid Communications in Mass Spectrometry.* **26**, 282-286.

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.

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.

Salbo, R., Bush, M. F., Naver, H., Campuzano, I., Robinson, C. V., Pettersson, I., Jorgensen, T. J. D. and Haselmann, K. F. (2012). Traveling-wave ion mobility mass spectrometry of protein complexes: accurate calibrated collision cross-sections of human insulin oligomers. *Rapid Communications in Mass Spectrometry.* **26**, 1181-1193.

Scarff, C. A., Thalassinos, K., Hilton, G. R. and Scrivens, J. S. (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.

Scarff, C. A., Thalassinos, K., Hilton, G. R and Scrivens, J. S (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.

Schmelzeisen-Redeker, G., Buttfering, L. and Rollgen, F. W. (1989). Desolvation of ions and molecules in thermospray mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes.* **90**, 139-150.

Shelimov, K. B., Clemmer, D. E., Hudgins, R. R. and Jarrold, M. F. (1997). Protein structure in vacuo: Gas-phase confirmations of BPTI and cytochrome c. *Journal of the American Chemical Society.* **119**, 2240-2248.

Sherman, F., Stewart, J. W. and Tsunasawa, S. (1985). Methionine or Not Methionine at the Beginning of a Protein. *Bioessays.* **3**, 27-31.

Shi, J. G., Lindsay, W. P., Huckle, J. W., Morby, A. P. and Robinson, N. J. (1992). Cyanobacterial Metallothionein Gene Expressed in Escherichia-Coli - Metal-Binding Properties of the Expressed Protein. *Febs Letters.* **303**, 159-163.

Shi, W. and Chance, M. (2008). Metallomics and metalloproteomics. *Cellular and Molecular Life Sciences (CMLS).* **65**, 3040-3048.

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.

Shvartsburg, A. A. and Smith, R. D. (2008). Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Analytical Chemistry.* **80**, 9689-9699.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D. and Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* **7**.

Smith, C. J., Clarke, A. R., Chia, W. N., Irons, L. I., Atkinson, T. and Holbrook, J. J. (1991). Detection and characterization of intermediates in the folding of large proteins by the use of genetically inserted tryptophan probes. *Biochemistry.* **30**, 1028-1036.

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.

Sobott, F., Hernandez, H., McCammon, M. G., Tito, M. A. and Robinson, C. V. (2002). A Tandem Mass Spectrometer for Improved Transmission and Analysis of Large Macromolecular Assemblies. *Analytical Chemistry.* **74**, 1402-1407.

Squires, G. (1998). Francis Aston and the mass spectrograph. *Journal of the Chemical Society, Dalton Transactions.* 3893-3900.

Stafford Jr, G. C., Kelley, P. E., Syka, J. E. P., Reynolds, W. E. and Todd, J. F. J. (1984). Recent improvements in and analytical applications of advanced ion trap

technology. *International Journal of Mass Spectrometry and Ion Processes.* **60**, 85-98.

Stola, M., Musiani, F., Mangani, S., Turano, P., Safarov, N., Zambelli, B. and Ciurli, S. (2006). The Nickel Site of Bacillus pasteurii UreE, a Urease Metallo-Chaperone, As Revealed by Metal-Binding Studies and X-ray Absorption Spectroscopyâ€ *Biochemistry.* **45**, 6495-6509.

Stryer, L. (1995). Biochemistry, Freedman.

Summers, A. O. (2009). Damage control: regulating defences against toxic metals and metalloids. *Current Opinion in Microbiology.* **12**, 138-144.

Tagle, M. V., Pozebon, D., Garcia, R. H., Pinar, F. C., Durruthy Rodriguez, M. D. and Alfonso, M. P. (2011). Methodology for the Determination of Stoichiometry and Metal Impurities in New PZT Ceramics by Inductively Coupled Plasma Optical Spectrometry (ICP OES). *Spectroscopy Letters.* **44**, 138-145.

Takach, E. J., Hines, W. M., Patterson, D. H., Juhasz, P., Falick, A. M., Vestal, M. L. and Martin, S. A. (1997). Accurate Mass Measurements Using MALDI-TOF with Delayed Extraction. *Journal of Protein Chemistry.* **16**, 363-369.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. and Matsuo, T. (1988). Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry.* **2**, 151-153.

Taudte, N. and Grass, G. (2010). Point mutations change specificity and kinetics of metal uptake by ZupT from Escherichia coli. *BioMetals.* **23**, 643-656.

Thalassinos, 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.

Thalassinos, K., Slade, S. E., Jennings, K. R., Scrivens, J. H., Giles, K., Wildgoose, J., Hoyes, J., Bateman, R. H. and Bowers, M. T. (2004). Ion mobility mass spectrometry of proteins in a modified commercial mass spectrometer. *International Journal of Mass Spectrometry.* **236**, 55-63.

Thelwell, C., Robinson, N. J. and Turner-Cavet, J. S. (1998). An SmtB-like repressor from Synechocystis PCC 6803 regulates a zinc exporter. *Proceedings of the National Academy of Sciences.* **95**, 10728-10733.

Touboul, D. and Zenobi, R. (2009). A simple model for exploring conformation of highly-charged electrosprayed single-stranded oligonucleotides. *Chemical Communications.* 298-300.

Turner, J. S., Glands, P. D., Samson, A. C. and Robinson, N. J. (1996). Zn2+-sensing by the cyanobacterial metallothionein repressor SmtB: different motifs mediate metal-induced protein-DNA dissociation. *Nucl. Acids Res.* **24**, 3714-3721.

Turner, J. S., Morby, A. P., Whitton, B. A., Gupta, A. and Robinson, N. J. (1993). Construction of Zn2+ / Cd2+ Hypersensitive Cyanobacterial Mutants Lacking a Functional Metallothionein Locus. *The Journal of Biological Chemistry.* **268**, 4494-4498.

Turner, J. S., Morby, A. P., Whitton, B. A., Gupta, A. and Robinson, N.J (1993). Construction of Zn2+ / Cd2+ Hypersensitive Cyanobacterial Mutants Lacking a Functional Metallothionein Locus. *The Journal of Biological Chemistry.* **268**, 4494-4498.

Turner, J. S., Robinson, N. J. and Gupta, A. (1995). Construction of Zn2+/Cd2+-tolerant cyanobacteria with a modified metallothionein divergon: Further analysis of the function and regulation of *smt*. *Journal of Industrial Microbiology & Biotechnology.* **14**, 259-264.

Uetrecht, C., Versluis, C., Watts, N. R., Wingfield, P. T., Steven, A. C., and Heck, A. J. R. (2008). Stability and Shape of Hepatitis B Virus Capsids In Vacuo13. *Angewandte Chemie International Edition.* **47**, 6247-6251.

Vallee, B. L. and Ulmer, D. D. (1972). Biochemical Effects of Mercury, Cadmium, and Lead. *Annual Review of Biochemistry.* **41**, 91-128.

van der Burgt, Y. E. M., Palmblad, M., Dalebout, H., Heeren, R. M. A. and Deelder, A. M. (2009). Electron capture dissociation of peptide hormone changes upon opening of the tocin ring and complexation with transition metal cations. *Rapid Communications in Mass Spectrometry.* **23**, 31-38.

van Duijn, E. and Heck, A. J. R. (2006). Mass spectrometric analysis of intact macromolecular chaperone complexes. *Drug Discovery Today: Technologies.* **3**, 21-27.

VanZile, M. L., Chen, X. and Giedroc, D. P. (2002a). Allosteric Negative Regulation of smt O/P Binding of the Zinc Sensor, SmtB, by Metal Ions: A Coupled Equilibrium Analysis. *Biochemistry.* **41**, 9776-9786.

VanZile, M. L., Chen, X. H. and Giedroc, D. P. (2002b). Structural characterization of distinct alpha 3N and alpha 5 metal sites in the cyanobacterial zinc sensor SmtB. *Biochemistry.* **41**, 9765-9775.

VanZile, M. L., Cosper, N. J., Scott, R. A. and Giedroc, D. P. (2000). The Zinc Metalloregulatory Protein Synechococcus PCC7942 SmtB Binds a Single Zinc Ion per Monomer with High Affinity in a Tetrahedral Coordination Geometry. *Biochemistry.* **39**, 11818-11829.

Vestal, M. L., Juhasz, P. and Martin, S. A. (1995). Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry.* **9**, 1044-1050.

Vincenzetti, S., Cambi, A., Neuhard, J., Garattini, E. and Vita, A. (1996). Recombinant Human Cytidine Deaminase: Expression, Purification, and Characterization. *Protein Expression and Purification.* **8**, 247-253.

Waldron, K. J. and Robinson, N. J. (2009). How do bacterial cells ensure that metalloproteins get the correct metal? *Nature Reviews Microbiology.* **7**, 25-35.

Wang, C., Vernon, R., Lange, O., Tyka, M. and Baker, D. (2010). Prediction of structures of zinc-binding proteins through explicit modeling of metal coordination geometry. *Protein Science.* **19**, 494-506.

WHATIF. "http://swift.cmbi.ru.nl/servers/html/index.html. ."

Wiley, W. C. and McLaren, I. H. (1955). Time-of-Flight Mass Spectrometer with Improved Resolution. *Review of Scientific Instruments.* **26**, 1150-1157.

Wilkins, M. C., Sanchez, J.-C., Gooley, A. A., D., A. R., Humphrey-Smith, I., Hochstrasser, D. F. and Williams, K. L. (1996). Progress with Proteome Projects: Why all Proteins Expressed by a Genome Should be Identifed and How to Do It. *Biotechnology & Genetic Engineering Reviews*
13, 19-50.

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.

Wilm, M. and Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. *Analytical Chemistry.* **68**, 1-8.

Wortmann, A., Rossi, F., Lelais, G. and Zenobi, R. (2005). Determination of zinc to beta-peptide binding constants with electrospray ionization mass spectrometry. *Journal of Mass Spectrometry.* **40**, 777-784.

Wyttenbach, T. and Bowers, M. T. (2012). Structural Stability from Solution to the Gas Phase: Native Solution Structure of Ubiquitin Survives Analysis in a Solvent-Free Ion Mobility-Mass Spectrometry Environment. *The Journal of Physical Chemistry B.* **115**, 12266-12275.

Xu, C. and Rosen, B. P. (1997). Dimerization Is Essential for DNA Binding and Repression by the ArsR Metalloregulatory Protein of Escherichia coli. *Journal of Biological Chemistry.* **272**, 15734-15738.

Ybarra, G. R. and Webb, R. (1999). Effects of divalent metal cations and resistance mechanisms of the cyanobacterium Synechococcus Sp. Strain PCC7942. *Journal of Hazardous Substance Research.* **2**, 1-9.

Yin, S. and Loo, J. A. (2011). Top-down mass spectrometry of supercharged native protein-ligand complexes. *International Journal of Mass Spectrometry.* **300**, 118-122.

Yu, X., Wojciechowski, M. and Fenselau, C. (1993). Assessment of metals in reconstituted metallothioneins by electrospray mass spectrometry. *Anal. Chem.* **65**, 1355-1359.

Zhang, W. Z., Niu, S. F. and Chait, B. T. (1998). Exploring infrared wavelength matrix-assisted laser desorption/ionization of proteins with delayed-extraction time-of-flight mass spectrometry. *Journal of the American Society for Mass Spectrometry.* **9**, 879-884.

Zubarev, R. A., Kelleher, N. L. and McLafferty, F. W. (1998). Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society.* **120**, 3265-3266.