Use of electrospray ionization mass spectrometry to study protein conformation and protein-protein interactions

Stephen J. Watt
University of Wollongong

Use of Electrospray Ionization Mass Spectrometry to Study Protein Conformation and Protein-Protein Interactions

A thesis submitted in (partial) fulfilment of the

requirements for the award of the degree

Doctor of Philosophy

from

University of Wollongong

by

Stephen J. Watt
Bachelor of Medicinal Chemistry (Honours)

Department of Chemistry
2005

CERTIFICATION

I, Stephen J. Watt, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy, in the Department of Chemistry, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualification at any other academic institution.

Stephen J. Watt
23th August 2005

For my family

PUBLICATIONS

Watt, S.J., Oakley, A., Sheil, M.M., and Beck, J.L. (2005) Comparison of Negative and Positive Ion ESI Mass Spectra of Calmodulin and its Complex with Trifluoperazine. *Rapid Communications in Mass Spectrometry,* **19**, 2123-2130.

Williams, N.K., Liepinsh, E., Watt, S.J., Prosselkov, P., Matthews, J.M., Attard, P., Beck, J.L., Dixon N.E. and Otting G. (2005) Stabilization of Native Protein Fold by Intein-Mediated Covalent Cyclization, *Journal of Molecular Biology,* **346**, 1095-1108**.**

ABSTRACT

The polymerisation of a polypeptide chain from an encoded genetic sequence allows the formation of structured molecules, known as proteins. These are essential components for a range of processes including molecular recognition, DNA replication and enzymatic functions. In this thesis, the ability of electrospray ionization (ESI) mass spectrometry (MS) to be used as a tool to determine functional properties of proteins has been explored. The coupling of ESI-MS to hydrogen deuterium exchange has been used to show how restriction of the C- and N-termini by cyclization of the polypeptide backbone can affect the ability of a protein to sample unfolded or partially unfolded states. The development of appropriate methodologies for analysis of linear (uncyclized) and cyclized systems identified a slowing of the rate of unfolding due to cyclization. The implication for the unfolding processes of proteins are discussed. An increased thermal stability of the cyclized protein was also demonstrated. This property was used to analyse the ability of ESI-MS to identify changes in protein structure from shifts in ion distributions. Important observations regarding the polarity of ionization used in these experiments are highlighted.

The effect opposite polarity ionization has on the ability to detect conformational changes in proteins and interactions with small ligands was explored using the well-characterized calmodulin-calcium-antipsychotic drug system. Important considerations regarding the binding of metal ions to protein structures are discussed in relation to the ability to unequivocally identify a conformational transition in protein structure from ESI mass spectra. An inability to detect complexes of calcium loaded-calmodulin with the antispychotic drug trifluroperazine in the negative ion mode was observed, a result believed to be due to the Coulombic repulsions between acidic residues of calmodulin.

Finally, the non-covalent complex and interactions of the E. coli helicase (DnaB) were probed by nanoESI-MS and MS/MS studies. Development of suitable conditions allowed for identification of a previously unresolved heptamer in addition to the expected hexamer. The interaction of DnaB with its loading partner DnaC and the possible roles of ATP and ADP in this interaction were also probed with findings being related to the biological functions of these proteins.

ACKNOWLEDGMENTS

There have been a number of people who have helped contribute to this thesis. I would especially like to thank:

- *My supervisors Jenny Beck and Margaret Sheil. I could not have asked for two better supervisors. Jenny, you were an inspiration to work with. Your commitment to your work and tireless efforts are second to none. Thankyou for all the time you have put into my project. I will always regard you as a fantastic scientist, attentive supervisor and great friend for many years to come.*

  Margaret, I have always enjoyed being your student. You have taught me more scientifically and personally than I could have ever imagined. Thankyou for always making time to see me and for all your support.

- *Stephen Blanksby for his thoughtful advice and pleasant nature. You have been a huge bonus for the mass spectrometry team.*

- *Larry Hick, the person who keeps the wheels turning. You never cease to amaze me how great your knowledge of these instruments is. Thankyou for all you have taught me and never failing to help me out whenever you were required. See you down the beach, Lazza.*

- *Raj Gupta, Thitima Urathamakul, Todd Mitchell, Andrew Aquilina, David Harman, Michael Thomas, Jihan Talib, Peter Hains, Amit Kapur, Roger Kanitz, John Korth, Karin Maxwell and all other past and present members of the Mass Spectrometry group. You have all grown to be a part of my family. I wish you all the best for the future and thankyou for providing such a great atmosphere to work in.*

- *Nick Dixon and his group from the Research School of Chemistry, Australian National University for so kindly providing a number of the proteins used in this study.*

- *The members of the Department of Chemistry for being such a great department to work in. I hope someday to find a work place with the same level of support, guidance and friendship available here; if another such place exists. Thankyou for all you have done for me over the past 8 years.*

- *The Australian Research Council for providing funding for my scholarship, instrumentation and this project.*

- *My wife Dianne (PhD in support). You have been with me every step of this roller coaster ride. Your patience, understanding and support over the last few years has meant more to me than words can express. You are a priceless treasure my heart will never tire of. Thankyou and I love you dearly.*

- *Finally, Mum, Dad, Jen, Nathan, Alice and all other members of my family. You have always been my inspiration. Thankyou for your endless support. The strength I have used to get me to this point has stemmed from you all.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADP	Adensine-5′-diphosphate
AMP-PNP	Adenosine 5′-(β,γ-imido)triphosphate
ATP	Adenosine-5′-triphosphate
ATPγS	Adenosine-5′-(γ-thio)-triphosphate
AS	Aerospray
βγATP	β-γ-methyleneadenosine-5′-triphosphate
BSA	Bovine serum albumin
CAD	Collisionally Activated Dissociation
CaM	Calmodulin
CD	Circular Dichroism
CI	Chemical Ionization
CRM	Charged Residue Model
CryoEM	Cryoelectron Micrograph
CSD	Charge State Distribution
DnaB-N	N-Terminal of DnaB
ds	Double Stranded
DSC	Differential Scanning Calorimetry
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EI	Electron Ionization
EH	Electrohydrodynamic
ESI	Electrospray Ionization
FIB	Fast Ion Bombardment
FAB	Fast Atom Bombardment

HDX	Hydrogen Deuterium Exchange
IEM	Ion Evaporation Model
IPA	Imipramine
ITC	Isothermal Titration Calorimetry
LD	Laser Desorption
MALDI	Matrix Assisted Laser Desorption Ionization
MHC	Major Histocompatibility Complexes
MS	Mass Spectrometry
m/z	Mass to Charge
NMR	Nuclear Magnetic Resonance
oa	Orthogonal Acceleration
OAc	Acetate
PD	Plasma Desorption
PDB	Protein Data Bank
Q-ToF	Quadrupole Time of Flight
ss	Single Stranded
SPR	Surface Plasmon Resonance
TFP	Trifluoperazine
ToF	Time-of-Flight
TS	Thermospray

Chapter 1
General Introduction

Protein shape or conformation is important on a variety of different levels. These include the dynamics of molecular recognition and enzymatic activity, through to the processes involved in the formation and functions of macromolecular machinery. This thesis will examine the ability of electrospray ionization (ESI) mass spectrometry (MS) to help understand protein folding and protein-protein and protein-ligand interactions, identifying strengths and weaknesses in its application.

1.1 Protein Conformational Dynamics

From the enzymes that synthesise a new DNA strand, to the antibodies that recognise foreign antigens within a host organism, the role and scope of proteins in the molecular machinery of life is extensive. Protein molecules are comprised of 20 naturally occurring amino acids (sometimes post-translationally modified), whose sequence and composition dictates their role(s) within cells. Considerable effort has been devoted to the question(s) of how a linear sequence of amino acids can fold into secondary and tertiary structures in the functional molecule, then, relay "signals" to other proteins through local structural fluctuations and/or large-scale conformational changes.[1,2] Understanding these processes will ultimately aid in the identification and treatment of conditions where the molecular machinery has malfunctioned.

1.1.1 The Protein Folding Problem

Transcription of messenger RNA (mRNA) from the genetic material DNA is followed by protein synthesis. This process is known as translation and is highly complex, requiring transfer RNA (tRNA), mRNA and more than a hundred proteins including ribosomes.[3] Translation is thus carried out in microenvironments where the concentration of newly synthesised proteins is high. In order to prevent aggregation of separate polypeptide chains, folding processes are required to be extremely fast.

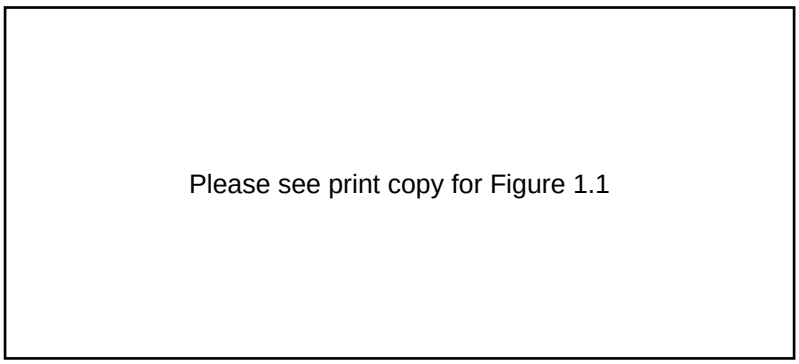


Figure 1.1 The secondary protein structures. (a) An α helix and (b) β sheets. α helixes are stabilized by intrastrand hydrogen bonding. β sheets are stabilized by intra- or intermolecular H-bonds between parallel or antiparallel strands. Adapted from Branden and Tooze.[4]

Owing to the steric constraints of the amide bonds linking adjacent amino acids, secondary structural elements of proteins are often in the forms of an α-helix or β-sheets (figure 1.1). Both these secondary structures are stabilized by hydrogen bonding between amide and carbonyl groups of the polypeptide backbone. In α-helices these bonds are largely intramolecular whereas β-sheet structures found in proteins are stabilized by intermolecular bonding between parallel or antiparallel β-sheets.

Computational studies combined with experiments using techniques that can follow rapid folding/unfolding, have shown that α-helices form *in vitro* in less than 100 nanoseconds (ns), and β-sheets in as little as 1 μs.[5,6] Small, single domain proteins (<100 residues) which subsequently form α-helix bundles have been determined to fold in 50 μs, with other small proteins based on β-sheet structures taking orders of magnitude longer.[7-9] These times for folding are astonishingly fast. Considering that a polypeptide chain of 101 residues allowing three rotation angles between connecting amino acids,[10,11] could exist in $3^{100} = 5 \times 10^{47}$ configurations. Even if a protein could sample a new configuration at a rate of 10^{13} per second, it would take 10^{27} years to try them all. Levinthal[12,13] therefore concluded that proteins cannot not fold via random searches. Protein folding is more likely to occur through molecular recognition events requiring the cooperative action of a number of non-covalent interactions rather than random searches.[14-16]

The amino acids of proteins can be defined into three groups based on their side chains (commonly called R groups). Those that are polar, apolar and charged. These side chains can interact with each other through non-covalent interactions such as hydrogen bonding, electrostatic interactions, van der Waal forces and dipole-diople interactions and through covalent disulfide bonds formed by oxidation of cysteine residues. Protein folding simulations[14,17,18] demonstrate that these interactions greatly increase the probability of finding a native conformation within biologically significant times. How exactly a protein adopts native (or lowest energy) conformations is still the subject of debate.[15,19-21]

1.1.2 Energy Landscape Model for Protein Folding

An energy landscape model has been put forward to explain protein folding, this is shown in figure 1.2.[15,16,21]

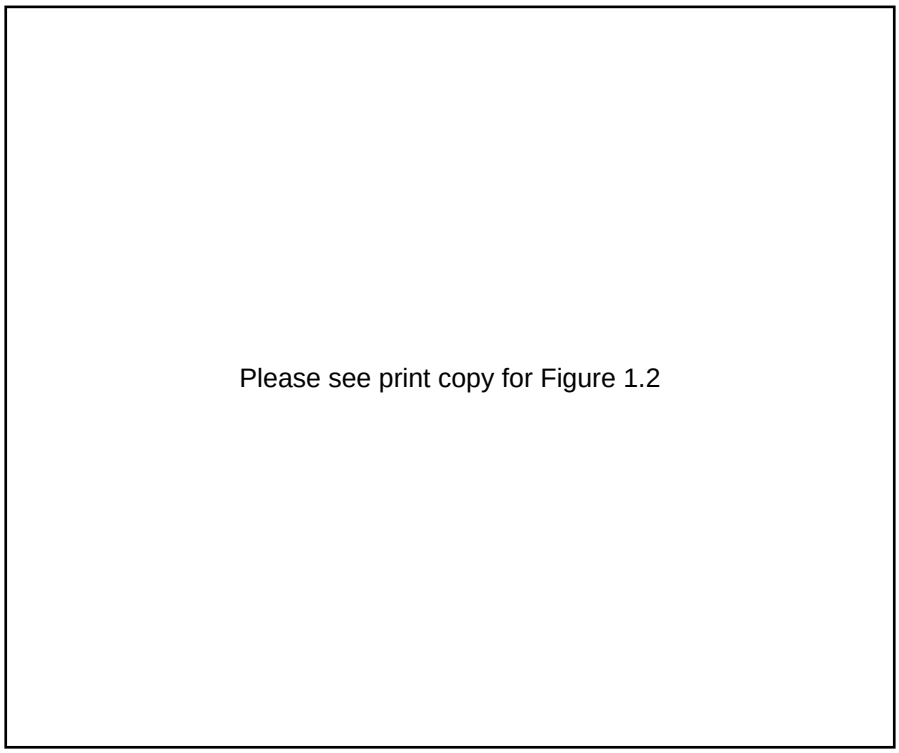
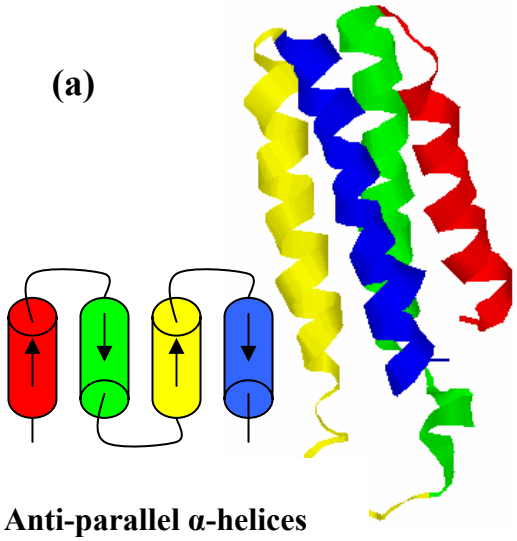


Figure 1.2 A schematic diagram of the landscape model for protein folding. Amino acid sequences with a number of starting conformations, funnel into structures with fewer more specific native interactions. A critical region in this relatively simple protein folding pathway is at the saddle point or transition state. This is an energy barrier which all molecules must cross if they are to fold to the native structure. An example of acylphosphatase[22] is shown to the right where at the transition state a selective number of native interactions are formed (as seen by the yellow spheres) which allow for folding to a native structure. Taken from Dobson.[16]

Since there are a large number of interactions accessible to the unfolded polypeptide chain, folding can begin from a number of different conformers. The funnel-like depiction of the folding process (as shown in figure 1.2) is the result of the progressive

organisation of partially folded structures through which the protein passes on its way to the native structure.[21] Comparative folding studies of native proteins and proteins carrying point mutations,[23,24] as well as proteins with similar topologies but with as little as 16% sequence homology[25,26] have shown that provided these conformers pass though a "saddle point" (or transition state) where native contacts between specific residues of the protein are formed, a fully folded, native structure, can be obtained. Interestingly, the average separation in sequence between residues that are in contact with each other in the native structure (commonly defined as the "contact order") has been shown to have a linear relationship with folding rates of small proteins, suggesting why some proteins can fold faster than others.[9]

The energy landscape model also suggests that the folding of a protein is driven by a significant drop in entropy. This is the result of favourable hydrophobic and hydrophilic interactions of the polymer. Although these interactions drive the formation of native structures, local minima can occur during folding where populations of molecules become transiently stuck during folding. These states have become a large focus of research as they may provide details of the protein-folding process.[27]



(b)

Please see print copy for Figure 1.3

(c)

Please see print copy for Figure 1.3

Figure 1.3. The structure of some commonly observed structural motifs. (a) An anti-parallel α-helix. (b) A β-sheet motif that may form β-beta barrels. (c) A β-α-β motif which forms part of α/β barrels. Figures (b) and (c) were taken from Nelson and Cox.[28]

1.1.3 The Formation of Tertiary Structure

The formation of a number of specific interactions between separate secondary structural units gives rise to tertiary structure. This provides a platform for the interaction of spatially separated amino acids enabling a variety of cellular processes including ligand binding, molecular recognition and enzymatic activities. Tertiary structures can be classified into a number of structural motifs including those shown in figure 1.3.

1.1.4 Dynamics of a Folded Protein

The formation of a tertiary structure commonly does not produce a static structure, instead species with degrees of flexibility important in a variety of cellular processes requiring molecular recognition. For example, in order for foreign bodies to be recognized by the immune system, major histocompatibility complexes (MHC) present peptides to T-cell receptor complexes. The dynamics of the T-cell receptor are such that it can sample hundreds of presented peptides and either form a specific interaction triggering an immune response when foreign peptides are detected or dissociate away from the MHC complex when an immune response is not required.[3,29,30]

The oxygen binding properties of hemoglobin (Hb) provide another example of the importance of protein conformational change. The oxygenated and deoxygenated forms of Hb represent two different conformations.[31-34] In the deoxygenated form the quaternary structure is quite tense producing a low affinity state. Upon binding oxygen a conformational shift in the structure produces a high affinity state. This allows oxygen to be transported to peripheral cells, but once released, Hb reverts back to its tense state. The conformational dynamics of this complex are thus essential for its activity. Interestingly, it has also been shown that organic phosphates (including bisphosphoglycerate) are capable of interacting with hemoglobin locking it into a low affinity state.[3] This is an important physiological control as it allows for a more efficient transfer of oxygen to peripheral cells. These organic phosphates are elevated in people who have lived at high altitudes where levels of oxygen are low and transfer to cells is required to be more efficient.

Conformational dynamics are also important in the catalytic activities of enzymes such as dihydrofolate reductase. This enzyme catalyses the reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate by hydride transfer from NADPH, an important step in thymidylate synthesis and hence cell growth and proliferation. Five kinetically observable conformational intermediates are involved in this process which are regulated by the movements of the Met20 loop[35] of the enzyme. These changes of the enzyme structure help to lock substrates and cofactors in the active site of the enzyme and produce flexibility when turnover of the reacting species is required. The knowledge of the dynamics of this enzyme has contributed to the development of new anticancer agents[36] and antibiotics.[37]

1.1.5 Techniques Used to Determine Conformational Dynamics of Folded Proteins

X-ray crystallography is an important and powerful tool for determining protein structure as it can provide information at atomic resolution. Unfortunately, difficulties are presented for growing crystals of proteins that are largely unstructured in certain areas or in proteins where conformational forms are in equilibrium. Furthermore, X-ray structures trap the protein in a single conformation and therefore do not reveal delicate fluctuations in protein structure essential for the biological activity in solution.[38-40]

Nuclear magnetic resonance (NMR) spectroscopy has over the last 20 years been at the forefront of techniques for determining protein structure and observing conformational dynamics. Heteronuclear NMR spectroscopy (proteins isotopically enriched with ^{2}H, ^{13}C and or ^{15}N) has enabled structural determination at near atomic resolution in many cases. These techniques together with NMR relaxation methods, enable conformational

dynamics to be probed over time frames from microseconds to milliseconds.[41,42] For example, in the analysis of ribonuclease A, dynamics in the range of micro- to millisecond time frames were observed for residues involved in the catalytic region of the enzyme.[43] This not only allowed for the identification of protein conformational intermediates sampled in native and near native conditions, but contributed to an understanding of the structural changes regulating overall rates of catalysis.[43] Similar experiments have been performed to study the mechanisms of ribonuclease binase[44] and triosephosphate isomerase.[45]

Protein conformational changes can also be studied by fluorescence spectroscopy.[46,47] Changes in the intrinsic fluorescence of tryptophan residues can accompany conformational transitions of proteins. An example of this is in the analysis of the kinetic mechanism of the 70 kDa molecular chaperones, molecules which assist in the refolding of proteins in an ATP-dependent manner.[48] Fluorescence studies have helped identify ATP-induced conformational changes which promote transition from a high affinity state to a low affinity state. This transition is important for the binding and release of target proteins. For useful information to be obtained from fluorescence of proteins, tryptophan residues in environments sensitive to conformational changes are required. Advances in detecting single molecule fluorescence have significantly enhanced these studies.[49] Further, by utilising either fluorescent products to monitor enzymatic activity[50,51] or covalently attaching fluorescent dyes[52,53] to specific sites of proteins, information regarding the molecular degrees of freedom, electrostatic environment and macromolecular interactions can be obtained. A limitation of this technique is that "handles" usually need to be introduced to label the protein. This can be achieved by site-directed mutagenesis to introduce cystine[52] or biosynthetically to

induce ketone handles for specific labelling of a protein.[53] Any modifications may result in structural changes to the proteins and/or result in inactivation of the protein.

A number of other complementary techniques are also available to probe conformational dynamics of proteins. These include circular dichroism (CD)[54] and infra-red [55] and Raman[55,56] spectroscopy. CD has widely been used to observe changes in both secondary and tertiary structure, to characterise folding intermediates, to evaluate the effects of site-directed mutagenesis on protein structure and to determine denaturation curves.[54] All these techniques, however, have a disadvantage when compared to NMR as they can only monitor cumulative changes in structure over the bulk sample.

1.1.6 Mass Spectrometry Techniques used to Probe Conformational Dynamics of Proteins

Over the past 17 years mass spectrometry has emerged as a tool used to probe conformations of biological molecules. This has largely been due to the development of electrospray ionization (ESI)[57] (and to a lesser extent matrix assisted laser desorption ionization (MALDI))[58] mass spectrometry which provide soft ionization of very large intact biomolecules in the gas phase. In this section the development of these ionization techniques will be reviewed as well as their applications to determining conformational properties of biomolecules by coupling to techniques such as hydrogen deuterium exchange and chemical modification, and from analysing the charge state distribution produced during ESI itself.

1.1.6.1 Development of Mass Spectrometry for the Analysis Biomolecules

Once molecules are ionized and present in the gas phase, a mass spectrometer can detect and analyse the molecules based on a ratio of mass-to-charge (*m/z*). Early ionization methods included electron ionization (EI) or chemical ionization (CI), which allowed the analysis of volatile molecular species. These techniques generally require analytes to be heated into the gas phase. These methods are therefore not suitable for the analysis of labile biomolecules (e.g. proteins and oligonucleotides) which fragment and degrade upon slow heating.

Broadly, two alternative approaches were pursued to maintain intact molecular species during ionization of labile molecules. The first is an energy transfer technique in which an analyte is dispersed over a surface while energy which may be in the form a radioactive isotope or beam of high energy ions is applied rapidly to the surface to desorb the analyte.[59] These techniques include plasma desorption (PD),[60,61] fast ion and atom bombardment (FIB and FAB)[62,63] and laser desorption (LD).[64] The addition of matrices like nitrocellulose for PD, thioglycerol for FIB and FAB and nicotinic acid for LD was found to have profound effects on desorption fidelity and a reduction in fragmentation. It is from these techniques (particularly LD) that MALDI,[58] a now routinely used technique in the analysis of biomolecules, emerged.

A second class of ionization requires the application of a strong electric field to aid in the extraction of ions from a substrate (i.e. solvent). These techniques include field desorption (FD),[65] electrohydrodynamic ionization (EH), thermospray (TS),[66] aerospray (AS),[68] and electrospray ionization (ESI).[57] FD and EH have not been widely used due

to tedious sample preparation and the difficulties in finding suitable non-volatile solvents, respectively. TS, AS and ESI all have similar methods of ionization, however, it is ESI that has now come to the forefront of biomolecule analysis essentially superceding the other techniques. ESI has been shown to produce ions from an aqueous or organic solvent which are not prone to fragmentation, even for labile biomolecules allowing it to be coupled to liquid chromatography techniques and other solution based methods including hydrogen-deuterium exchange.

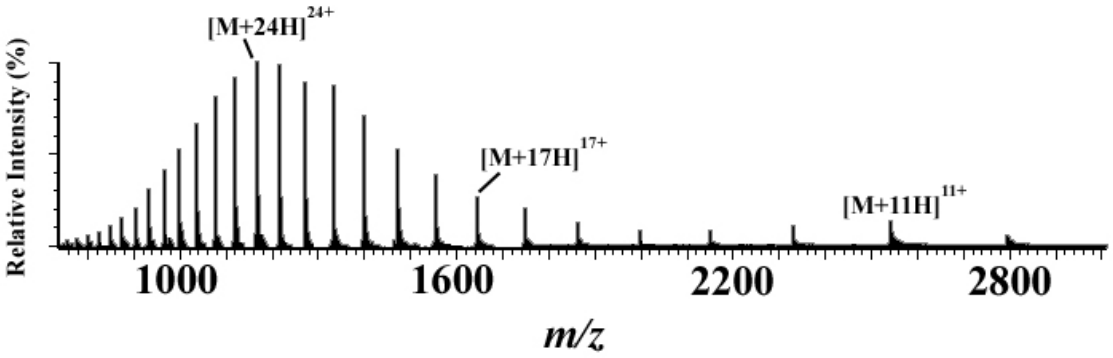


Figure 1.4 An example of a charge state distribution produced by electrospray ionization mass spectrometry. This spectrum is of the *E. coli* protein DnaC in 1% formic acid. Ions with 24, 17 and 11 positive charges are labelled as $[M+24H]^{24+}$, $[M+17H]^{17+}$ and $[M+11H]^{11+}$ respectively. Spectrum was acquired using a Waters Q-ToF-2^{TM} mass spectrometer.

1.1.6.2 The Charge State Distribution

The development of ESI-MS by Fenn and his co-workers,[57] followed the pioneering work of Dole.[67] Fenn identified that a characteristic high proportion of charges from species ionized by this method often appeared in a Gaussian distribution as shown for example in figure 1.4. This profile is frequently referred to as the charge state distribution (CSD). As mass spectrometers detect ions based on their *m/z*, different charge states can easily be resolved and provide the molecular mass of a multiply charged species using the following equations (for positive ions).

$$(m/z)_1 = \frac{M + nH}{n} \quad \textbf{(1.1)}$$

$$(m/z)_2 = \frac{M + (n+1)H}{n+1} \quad \textbf{(1.2)}$$

Where $(m/z)_1$ is the mass/charge of the first charge state, $(m/z)_2$ is the mass/charge of the second charge state, and M is the mass of the analyte. Following determination of n (the charge state) M can be calculated using equation 1.3.

$$M = n(m/z)\text{-}nH \quad \textbf{(1.3)}$$

The high proportion of charges produced by this technique in comparison to older techniques and MALDI aided the rapid uptake of ESI as sources were readily coupled to mass analysers with low cost, and low *m/z* ranges (e.g. quadrupoles, ion traps).

1.1.6.3 Mechanism of ESI

At present there are two ionization methods which dominate discussions about the ESI mechanism: The Charged Residue Model (CRM) proposed by Dole[67] and the Ion Evaporation Model (IEM) of Iribarne and Thomson.[68,69] The models describe similar initial stages of the ESI process where analytes are infused to the ionization source via a narrow capillary. This capillary has a high electric field applied at the tip which, when positively charged, causes positive ions (or negative ions when negatively charged) from the solutions to be enriched at the capillary tip and separate away from counterions. The force of the electric field and the accumulation of the positive ions expands the liquid into a Taylor cone, which subsequently breaks up into droplets containing positive charges at the surface. It is also largely agreed that these droplets continue to decrease in size via solvent evaporation and Coulombic fissions.

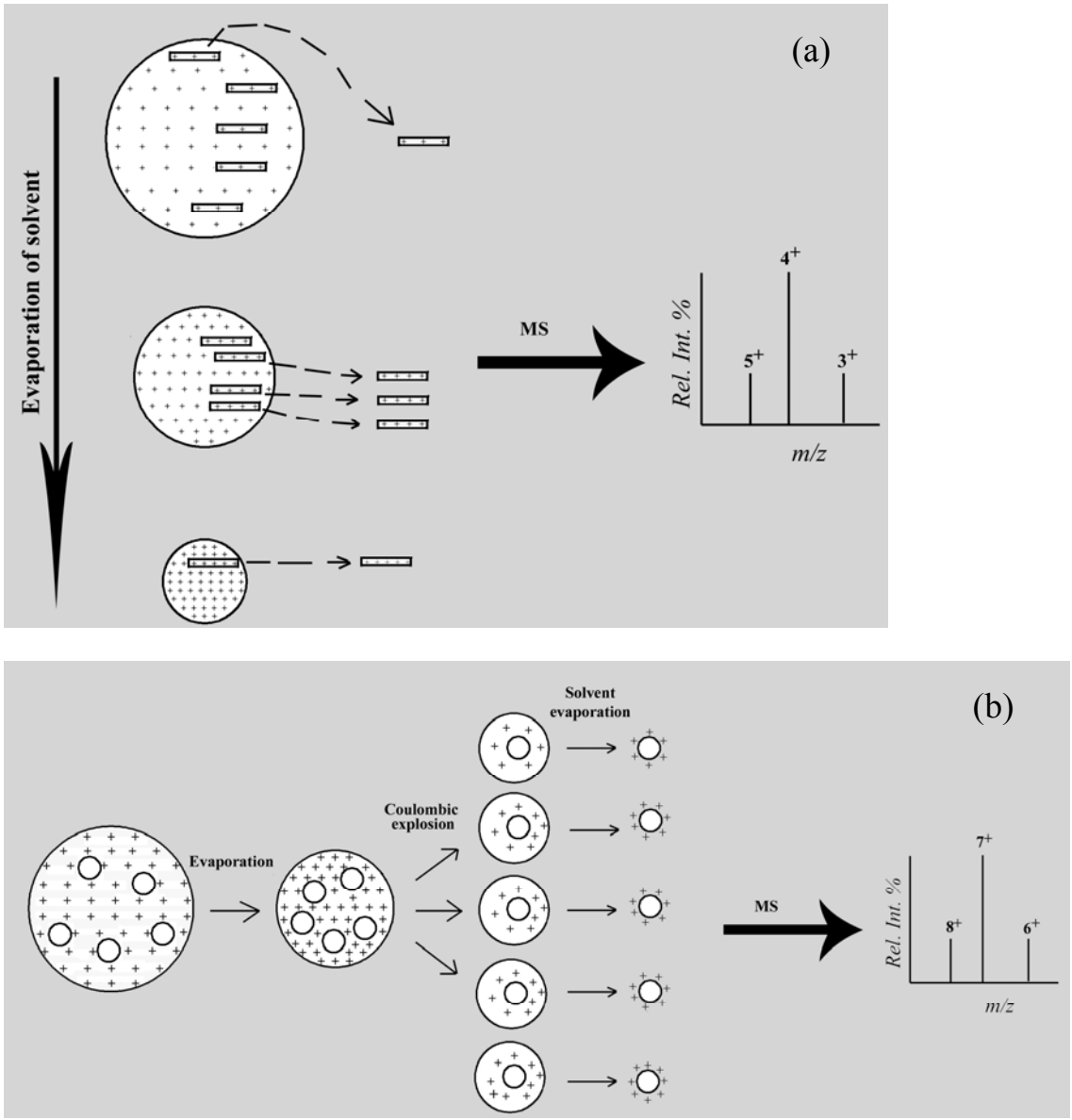


Figure 1.5 A schematic picture of the process of: (a) the Ion Evaporation Model and, (b) the Charged Residue Model for ESI. See text for description.

The IEM proposes that the radii of droplets produced from the ESI capillary decrease in size until a certain point prior to the Rayleigh limit is reached where Coulombic repulsions are sufficiently high to eject solvated ions into the gas phase.[70] Fenn[71] later extended this model to account for the Gaussian distribution of charges produced by proteins when sprayed from an ESI source. In this theory, Fenn proposed that ions with fewer charges leave the droplet earlier in the evaporation sequence than those with many charges.[71] Assuming in the initial stages of ionization the abundances of both

analyte and charges remain constant, the subsequent evaporation of solvent from the droplet will yield a higher surface charge density. This will greatly increase the number of charges within reach of the charge sites on the molecule.

In accordance with Iribarne and Thomson, Fenn assumed that evaporation or desorption of analyte ions can be treated as an active process in which an activation energy (Δ_{Giz} where i is a molecule of the analyte and z is its charge) is required to escape from the surface of a droplet. This activation energy dictates the rate of escape, as those molecules with more charges will have greater repulsions with the cations in solution and escape more readily. As shown in figure 1.5 (a) this suggests that initially the rate of desorption will be quite slow as the concentration of cations is not sufficient to access a large number of sites on the analyte. As the surface area of the droplet decreases the subsequent increase in the charge density will increase the abundance of charges present on the analyte surface, increasing the rate of desorption. As the droplet continues to decrease in size, the factors described above will generate molecules with higher charges, however, the concentration of solute or even the charge carried by the droplet (taken by desorbing molecular species) will decrease producing a relatively low abundance of molecules with high abundances of charges. This process will ultimately produce a mass spectrum with a Gaussan distribution of charges for a single molecular species with a number of accessible charge sites.

In the CRM, droplets containing analytes remain spherical throughout ionization except perhaps during the brief moments when they undergo a Coulombic explosion. The hydrophilic properties of a folded protein greatly reduce the likelihood of desorption by IEM mechanisms, causing the analyte to remain inside the droplet until complete

evaporation of the solvent occurs.[72] When this situation occurs, following evaporation of solvent, the analyte will have available to it the same charge the droplet had moments prior to evaporation. This is known as the Rayleigh limit of charge (q_R) and can be defined as seen in equation 1.4

$$q_R = 8\pi(\gamma\varepsilon_0 R_i^3)^{1/2} \quad \textbf{(1.4)}$$

where γ is the surface tension of the droplet, ε_0 is the permittivity of vacuum and R_i is the radius of the ion. Using data acquired by Schwartz *et al.*[73] and Tolic *et al.*,[74] Fernandez de la Mora has shown that globular proteins maintain a maximum charge of between 65% and 110% of q_R when ionized from native conditions.[75] This has given additional support for the CRM being the likely ionization mechanism for native proteins. The rather large error for levels of q_R have been explained by Fernandez de la Mora as a result of precursor droplets originating from fissions of parent droplets which contain a lower ratio of charge to analyte (see figure 1.5(b)).[76] If these precursor droplets then dry shortly after explosion, the analyte may contain as little as 65-70% of q_R. The above process would also account for the different charge states observed in ESI-MS spectra of native proteins.

Since Iribarne and Thompson based their calculations on the desorption of small ions and since there are substantial experimental observations supporting the CRM, it is generally thought that native proteins are ionized by the latter mechanism.[75,77,78] The mechanism is important when considering ESI mass spectra of folded (or native conformations) compared with unfolded or denatured proteins (see also 1.1.6.6).

1.1.6.4 Hydrogen Deuterium Exchange and Mass Spectrometry

Hydrogen deuterium exchange (HDX) is an acid-base catalysed reaction where labile protons are exchanged for deuterons (or deuterons for protons).[79] Exchange involves protons capable of hydrogen bonding (i.e. NH, OH and SH groups), however amide protons are of particular relevance as they are located along the polypeptide backbone and have exchange rates that are readily examinable. Structural information can be obtained by monitoring HDX rates of amide protons as an increase in the degree of hydrogen bonding and limited solvent accessibility will slow HDX rates. Mass spectrometry can be used to monitor this exchange as the incorporation of a deuteron for a hydrogen gives rise to mass increase of 1.006 Da.

The utility of HDX to probe protein structure was first demonstrated by Lenormant and Blout[80] and Hvidt and Linderstrom-Lang[81] who used HDX coupled with infrared spectroscopy and ultrafiltration (density gradient measurements) to obtain qualitative and quantitative data, respectively, about protein structure. Later Leach and Hill[82] replaced deuterium (^{2}H) with tritium (^{3}H) to directly monitor exchange in ribonuclease by scintillation techniques. This coincided with the first applications of NMR to monitor HDX.[83] At this time the technique was limited by low frequency NMR spectrometers. It was not until some twenty years later that NMR could be widely applied to monitor HDX owing to the development of high field and multi-dimensional NMR techniques. These developments rectified the earlier problems with peak overlap in 1D NMR experiments.[84,85] Using two-dimensional J-correlated (COSY) spectroscopy, the incorporation of a deuterium for a hydrogen results in the loss of signal (or appearance of a signal when a hydrogen replaces a deuterium). This allows for exchange rates to be monitored by obtaining spectra at various time points. The

development of rapid mixing techniques has allowed the application of NMR to examine amide exchange characteristics of proteins in transient folding states.[86]

Although NMR has provided an incredible amount of site-specific information on protein structure, there are a number of drawbacks limiting its use to some extent. When protein sizes are greater then 30 kDa, proton signal overlaps make interpretation of spectra difficult. The presence of paramagnetic cofactors such as ferric ions in proteins also limits studies by NMR. Further, large quantities and concentrations of proteins are required for the analyses.

It has only been in the last 15 years with the development of soft ionization techniques such as ESI, MALDI and FAB that HDX in proteins has been analyzed by mass spectrometry. Katta and Chait were the first to couple HDX with mass spectrometry in their analysis of acid- and alcohol-denatured ubiquitin.[87] This was quickly followed by Smith and co-workers[88] and Robinson and co-workers.[89] These reports showed that mass spectrometry offered complementary information to NMR.

Mass spectrometry can readily provide information on the presence of various protein conformations using small amounts of protein.[89] NMR requires large quantities of protein but can provide information on the identity of the amide protons that are exchanged. Attempts have been made to determine which amide protons are exchanged in HDX monitored by mass spectrometry by coupling these experiments with protease digestion. The peptides with increased mass as a result of HDX separated by either "off-line" or "on-line" HLPC can be identified by sequencing using ESI-MS/MS.[90]

Zhang and Smith[91] were the first to couple enzymatic digestion (using pepsin at pH 2-3 and 0^{o}C) with HDX and MS to obtain sequence specific information. Other more recent examples include the analysis of conformations of bovine insulin[92] and the conformations of reduced and oxidized *E. coli* thioredoxins.[93] A problem of carrying out this technique is that even with immobilized pepsin columns, up to 10% back exchange (deuterons that replaced protons are replaced by protons) can occur during protease digestion.[94]

Peptic digests of protein subjected to HDX have also been analyzed by MALDI-MS.[95] This requires only a slight variation in sample preparation from conventional MALDI target preparation methods and eliminates the need for a HPLC separation prior to analysis. Mandell *et al.* revealed that despite complete dryness and the presence of high vacuum that some back exchange did occur, possibly from the carboxylic acids of the matrix and/or water crystallization in the solid state.[95] The authors commented that these values were similar to back exchange rates determined by Bai *et al.*[79] for rates of exchange under quenching conditions (pH 2.5, 0 ^{o}C). HDX-MALDI techniques are now being used in SUPREX (stability of unpurified proteins from rates of H/D exchange) thermodynamic studies of proteins[96-98] and the analysis of conformational transitions.[99]

Recently, high resolution Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) has been used to fragment long polypeptides that were subjected to HDX, eliminating the need for protease digestion.[100-103] In some of these experiments, the deuterium incorporation level of individual amino acids was difficult to obtain, particularly when low energy fragmentation mechanisms (i.e. repeated low energy

collisions) were employed. This was the result of mobile protons (or hydrogen scrambling) moving along the polypeptide backbone during fragmentation.[104] Another method, electron capture dissociation,[105-108] been shown to give rise to specific fragmentation pathways (only c- and z-ions are produced)[109] and appears to lead to minimal hydrogen scrambling. This technique, along with expression of proteins without C-13 and N-15 isotopes has aided in the analysis of the protein ubiquitin by increasing signal responses 2-4 fold and allowed site specific MS/MS results following solution phase H/D exchange.[110]

1.1.6.5 Chemical Modifications

Chemical modification of proteins is a well-established technique for determining structure-function/activity relationships.[111,112] Recently, these methods have been combined with mass spectrometry to enable precise determination of which amino acids are modified.[113] Reagents such as the lysine-specific bis(sulfosuccinimidyl) substrate (BS^3) have been used to form cross-links with single proteins to determine which folds of the protein are in close proximity to one another[114] and to examine the binding of protein-protein complexes.[113,115,116] The addition of a cross-linking reagent causes an increase in mass which is readily detected by mass spectrometry, however, in order to determine the sites of modification so structural information can be obtained, an enzymatic digestion followed by a tandem MS/MS experiment is usually required. Identification of cross-linked peptides can be a difficult and time consuming process, however, by using cross-linking reagents with cleavable thiol groups,[117] cross-linking with isotopic labelled reagents[118,119] or the addition of O-18 water in the protein digestion,[120] interpretation of MS spectra can be simplified.

1.1.6.6 Analysis of Protein Conformations from the Charge State Distribution Observed in ESI Mass Spectra

Chowdhury *et al.*[121] and Loo *et al.*[122] were the first to report that when proteins were sprayed from solutions of varied pH or in the presence of organic solvent, a shift in the observed charge state distribution (CSD) would result. The more denaturing the conditions were, the higher the charge states compared to more neutral/native conditions. This raised the possibility that ESI-MS could be used to identify folding intermediates of proteins.

In the study by Loo *et al.*, a correlation was made between structural changes observed in protein solutions by CD and NMR spectroscopy and changes in the CSD observed in ESI mass spectra when minimal amounts of organic solvents were added to mixtures.[122] This study, however, did not specify what change(s) in the protein structure were being monitored. Later, Konermann and Douglas[123] carried out a detailed analysis of the tertiary and secondary structural changes under different solvent conditions using CD and fluorescence spectroscopy. These were compared to changes observed in the CSD of positive ion ESI mass spectra of cytochrome c in similar solvent conditions. The results of these experiments were interpreted to support that changes in CSD of positive ion ESI mass spectra reflected changes in tertiary structure of cytochrome c. Further studies by Konermann[124] and Grandori[125,126] have supported these findings.

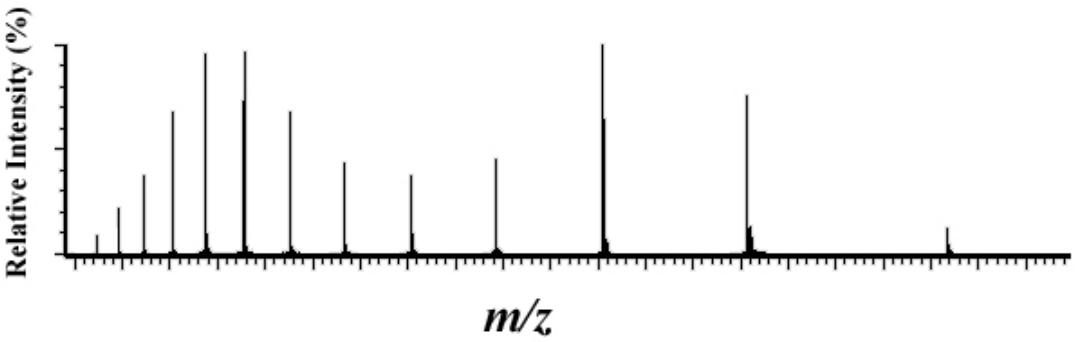


Figure 1.6 An example of bimodal charge state distribution produced by electrospray ionization mass spectrometry. Spectrum is of the N-terminal domain of the DnaB protein, acquired in the negative ion mode with a desolvation temperature of 60 ^{o}C.

Bimodal distributions of charge states as seen in figure 1.6 have been observed in ESI mass spectra indicating the presence of two conformational states of a protein under certain conditions. Konermann exploited this observation to analyse transient unfolding intermediates of cytochrome c by utilizing rapid mixing techniques. In that study Konermann *et al.*[127] induced a rapid re-folding of acid-denatured cytochrome c. ESI mass spectra were able to present lifetime measurements for subpopulations of cytochrome c which reflected intermediates in the refolding process. These results were in accordance with previous solution phase measurements. These studies can become quite complex to interpret and quantify since peaks may overlap. Dodo and Kaltashov[128] have developed a method to deconvolute CSDs in a fairly fast and straightforward manner, without requiring access to dedicated instrumentation or computer facilities. This has aided in the analysis of holo- and apomyoglobin in distinguishing between four different states of acid unfolding as seen in other solution phase experiments.[129,130]

The ability of ESI-MS to detect conformational changes has been explained in the following ways: (1) Without the structural order of proteins, as in the case of a denatured protein, excess charges are obtained by higher solvent accessibility to acidic and basic residues. The subsequent decrease in steric hindrance and Coulombic repulsions enables proteins to maintain a high quantity of charges.[122,131] Alternatively, this has been explained as resulting from the denatured proteins becoming more linear so that the droplets are no longer spherical and the Rayleigh limit q_R (see section 1.1.6.3) does not apply. (2) In the case of folded proteins, where steric hindrance and Coulombic repulsion appear to be a driving factor for the reduced charge state seen in the spectra of native proteins,[132,133] Grandori has pointed out that conformational dependent neutralisation (CDN) may be a key factor affecting the suppression of high *m/z* ions of the CSD.[134] Grandori's CDN theory suggests that the highest abundance charge state observed for a protein sprayed from a native/neutral solution should be the difference between positive and negative charge sites in the primary sequence. Nesatyy and Suter[78] have shown, however, that this only occurs in a limited number of cases reported in literature and that the Rayleigh limiting charge theory (RLCT) proposed for spherical protein structures[135] accounts for a larger number of cases observed in the literature to date, and supports that native proteins are ionized largely as explained by the CRM.

1.2 Non-Covalent Complexes

Non-covalent complexes between biomolecules regulate many cellular processes. These complexes range from metals in proteins that are essential for activity, to enzyme–substrate (small molecule) complexes through to complexes involving two or

more macromolecules (e.g. nucleic acids and proteins). These interactions allow enzymatic process such as the hydration of carbon dioxide by carbonic anhydrase (an enzyme-metal-substrate non-covalent complex)[136] and the regulation of gene expression as seen in the non-covalent interaction between the λ repressor protein and λ bacteriophage DNA, a process that controls transition between the lytic (cell lysing) or lysogenic (dormant) phase of this virus.[3]

Non-covalent complexes derived from multiple protein, nucleic acids and/or other molecules are commonly being identified and classed as molecular machines. For example, ATP synthase is a molecular machine which synthesises ATP, the nucleoside triphosphate used for many purposes including muscle contraction and the active transport of molecules and ions.[3] Two separate protein motors, the F_1-ATPase and the F_0-ATPase complexes drive the synthesis of ATP through the cooperative actions of the catalytic F_1-ATPase motor and the proton channelling properties of the F_0-ATPase motor.[137-139] This cooperative action is highly complex, requiring the interactions of a number of protein domains, interactions with the mitochondrial membrane and interactions with ADP, phosphate molecules and ATP. These forms of cooperative non-covalent interactions are also present on the cell surface where signal cascades can begin. Multiple kinases and phosphatases are anchored via non-covalent interactions to ion channels. This anchoring process allows a sensitive and specific regulation of phosphorylation and dephosphorylation events when an influx of ions occurs.[140]

Having components of cellular machinery in close proximity is also very important to allow processes like protein synthesis and DNA replication to be carried out in an efficient fashion. Ribosomes, which function to synthesise polypeptides based on

information from messenger RNA sequences require the cooperative actions of transfer RNA to match mRNA codons. Assemblies of proteins, RNA molecules and amino acids are linked by non-covalent interactions to direct this process.[141-143]

The replicative machinery comprises a functionally conserved set of components including polymerases, editing exonucleases, helicases and accessory proteins that synthesise new strands of DNA to reproduce the chromosome. These components interact non-covalently and guide the melting of duplex DNA to single strands, replicate it and proofread it allowing for replication to occur efficiently. To facilitate this process a number of these complexes and individual components have multiple conformations and are able to interact with a number of small molecules (e.g. metals and nucleotides) to allow this catalytic process to fluently take place.[144-146]

In order to understand processes such as ATP synthesis, hormone-receptor interactions, protein synthesis and DNA replication, understanding of the interactions between the various components of these complexes is required.

1.2.1 Techniques to Characterise Non-Covalent Macromolecular Complexes

Currently there is a range of techniques available for characterisation of non-covalent complexes of macromolecules and for providing mechanistic information about their assembly. NMR spectroscopy and X-ray crystallography are two methods that can identify a number of structural and functional properties of macromolecular complexes. These may include residues involved in catalytic sites, the identity of residues which interact with different subunits, and identification of fluctuations in structures upon ligand binding. However, as discussed in section 1.1.5 if complexes are large,

complexed with paramagnetic ions, or difficult to crystallise, other techniques are required to probe their structure.

An alternative imaging technique currently being employed is cryo-electron microscopy (cryoEM). It produces images with resolution to approximately 4 Å.[147,148] The advantage of this technique is that it does not require the growth of crystals of X-ray diffraction quality. Instead samples are super-cooled by the application of liquid nitrogen or liquid helium. As discussed by Henderson[149] the cooling of these particles is essential as radiation by electrons can lead to destruction of the covalent bonds. This has allowed for the overall structural characterisation of molecules such as the tubulin dimer[147] and the membrane water channel aquaporin.[148] When high resolution images are not able to be produced (a problem for very large >250 kDa or asymmetric structures) a combination of cryoEM with bioinformatics sequencing strategies, can provide a detailed characterisation of structures within complexes.[150] This has been used to determine structural aspects of the macromolecular rice dwarf virus to 6.8 Å.[151]

Like X-ray crystallography, cryoEM provides a static image of biomolecules. A number of other techniques are available to probe the relationship between molecular structure and the biological function of macromolecules.[152-155] Ultracentrifugation is making a return as a technique capable of probing a number of structural and dynamic properties of macromolecules. Analysis by this technique measures rates of diffusion of molecules through solvents under the influence of a centrifugal force. Absorbance or interference optics inside the instrument, enable detection of intermediate strength interactions (K_D between 10^{-3} and 10^{-8} M) between biomolecules, ligand binding-promoted conformational changes, macromolecular shape changes and measurement of

equilibrium constants.[152] An example can be seen in the application of this technique to determining of the size, shape and stoichiometry of the calponin-actin complex.[156]

Traditional biochemical techniques such as size exclusion chromatography, non-denaturing gel electrophoresis, and spectroscopic techniques (e.g. fluorescence and UV-Vis) have been used to monitor and measure interactions between biomolecules as well as interactions of biomolecules with small ligands. Size exclusion and non-denaturing gels, are techniques that allow the relative sizes of macromolecules to be determined. When complexes form, commonly structures undergo a conformational shift, changing the apparent surface area of the complex in a non-linear manner. This poses a problem for techniques which require comparisons of elution times or migration distances (depending on the technique being used), as a change in the molecular shape of a protein will change the scale of the interaction with the polymer. As a result, errors in mass determination may occur which will vary depending on the standards used to calculate a calibration curve. Fluorescence techniques are also able to probe conformational transitions of complexes as discussed in section 1.1.5.

An alternative method for the analysis of the interactions of complexes is through the use of surface plasmon resonance (SPR; e.g. BIAcore). The attachment of one binding partner of a complex to a gold surface allows a ligand to be passed over the immobilized binding partner in a variety of different solvent systems. Interactions between species can then be monitored by observing changes in the refractive index at the sensor surface as a result of binding.[157,158] A range of biological molecules have been applied to the metal surface in SPR studies including proteins, nucleic acids, lipids, carbohydrates and small molecules. SPR allows for dissociation equilibria with dissociation constants

between 10^{-4} to 10^{-11} M to be determined,[152] and monitoring of conformational changes.[159] This has been used to analyze a large number of molecular interactions including analysis of the processes associated with a number of aspects of the viral life cycle. These include the initial docking events between the host CD4 molecule and the viral gp120 protein, analysis of the nucleic acid binding properties of the helix clamp region of reverse transcriptase and mapping epitopes in the interaction between antibodies and the HIV capsid protein.[160]

Isothermal titration calorimetry (ITC) and differential scanning calorimetry (DSC) are two of the most direct techniques measure thermodynamics of complex formation. ITC monitors the heat change of a complex at constant temperature allowing the magnitude of binding and the enthalpy and entropy changes to be determined. This is achieved by the titration of one of the binding partners of the reaction into a reaction cell containing the other reactant. By comparing the energy released or absorbed by the reaction cell to a reference cell, the amount of power (microcalories) required to maintain a constant temperature can be determined. This allows for dissociation constants between 10^{-8}-10^{-20} M to be measured.[161] ITC also can be useful in understanding the roles of residues in the stability of complexes when coupled with site-directed mutagenesis[162,163] or variations in the ligand sequences.[164]

DSC resolves the energetics of conformational transitions of biological macromolecules.[153] By monitoring the heat capacity change of a molecule as a function of temperature, DSC allows the identification of the thermodynamic properties that govern a conformational transition between a folded and unfolded state and the transitions accompanying ligand binding. This can be seen in the analysis of the of the

lambda Cro repressor dimer where DSC studies have revealed that equilibrium unfolding occurs via a highly-populated intermediate state corresponding to polypeptide tetramers.[165] Together, DSC and ITC are useful tools to analyze the thermodynamics of the association of macromolecules.

These techniques do have some limitations. Often they require the purification of large amounts of proteins, analysis of complexes can be time consuming, and binding characteristics may not be able to be obtained without complex calculations. Mass spectrometry, however, offers an alternative and complementary method to explore non-covalent interactions in a highly sensitive and straightforward manner.

1.2.2 ESI-MS of Non-Covalent Complexes

As discussed in section 1.1.6, the development of ESI offered an ionization method that allows analysis of intact, large biomolecules by mass spectrometry. Not only are fragile molecules such as proteins and DNA maintained during ESI, early reports were published that suggested specific non-covalent interactions correlating with solution data could be observed. These observations were first made by Ganem *et al.* in their analysis of the complex between the human FKBP receptor (protein) and the FK506 and/or rapamycin immunosuppressants.[166] Shortly after Katta and Chait reported specific interactions between heme and myoglobin in its native state.[167] These reports were followed by numerous reports of applications of ESI-MS of non-covalent complexes including protein-protein,[168] protein-DNA,[169,170] DNA-drug[171] and protein-cofactor[172] complexes. An excellent review by Loo has been published looking at aspects of these early studies[173] and reviews by Beck *et al*.[174] and Veenstra[175] examine interactions involving DNA in particular.

ESI-MS has been described by Loo as having the "S" advantages over other commonly used techniques (see section 1.2.1) when analysing non-covalent complexes.[173,176] The first of these "S" advantages is speed. In comparison with techniques such as NMR and X-ray crystallography, mass spectrometry can acquire data of a non-covalent complexes in extremely short time frames. The advantage of sensitivity is also extensively exploited with femto-attomoles of sample capable of being detected.[177,178] In many NMR experiments nanomoles (~ 1 mg/mL of protein in ~ 1 mL) are required. The final "S" advantage is specificity. Two components can come together and form specific interactions, which can be readily detected by ESI-MS enabling determination of the stoichiometry and relative binding affinity information in an easily interpreted manner. This is provided that appropriate solvent and instrumental conditions are used when attempting studies by ESI-MS.

1.2.2.1 Conditions for Observing Non-Covalent Complexes by ESI- MS

For non-covalent complexes to be detected by mass spectrometry, gentle solvent and instrumental conditions need to be used. Measurement of the mass of biomolecules by ESI-MS with high intensity and sharp resolution, requires intermediate concentrations of organic solvent (e.g. acetonitrile or methanol between 30-70%) in addition to acidification (e.g. formic or acetic acid) to take the pH between 2-4 for detection of positive ions. For detection of negative ions, base (e.g. ammonia) is added to produce a pH of 8-10.[175] As non-covalent interactions rely on maintaining native structures, conditions such as these are not applicable for their analyses. Instead, solutions containing the volatile salts ammonium acetate or ammonium bicarbonate are commonly used in water to maintain the pH of the solution close to physiological conditions and allow the detection of these complexes without salt adducts.

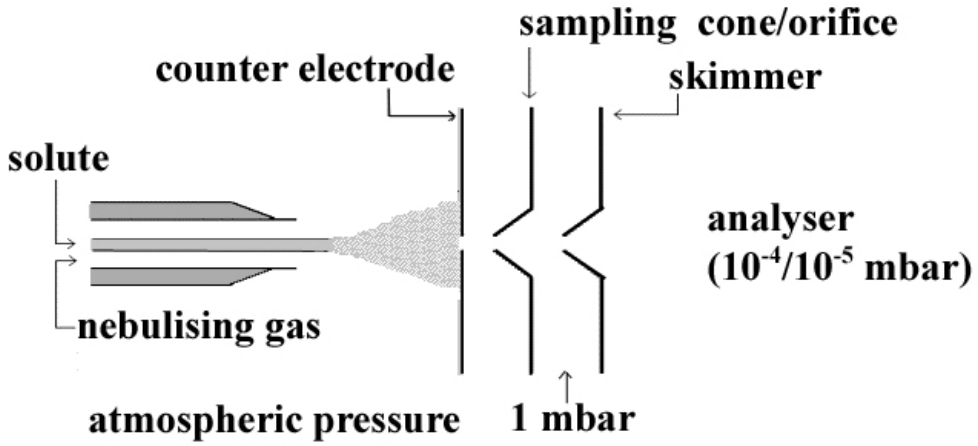


Figure 1.7. A schematic of an electrospray ionization source. Droplets produced from the electrospray capillary are desolvated with help of a nebulising gas. Lowering of potentials between the capillary, sample cones and skimmers and the decrease in pressure through the various stages of the source assists in guiding ions towards mass analysers.

Instrument parameters also have to be closely monitored for the analysis of non-covalent complexes. Owing to the use of aqueous solvent systems it can be tempting to use strong desolvation conditions (e.g. higher flow rates of dry nitrogen, increased orifice/cone voltages and source temperatures) to help resolve ions more readily. As shown by Robinson *et al.*[179] in their analyses of acyl CoA binding protein complexed with acyl CoA derivatives increasing temperatures in the source of a mass spectrometer from 20 to 80 ^{o}C can cause destabilisation of complexes. Veenstra *et al.*[180] in the analysis of binding of calcium to calmodulin have also reported that the kinetic energy applied to ions at extractor orifices in mass spectrometers (figure 1.7) can also affect the stability of complexes. With an increase in the potential applied to ions at the orifices, fragmentation can occur as complexes obtain too much kinetic energy. This has also been seen in the analysis of the γ- interferon homodimer (K_D = 50 nM), where increasing the cone voltage from 30-80 V significantly reduced the abundance of ions

from the mass of the homodimer.[181] Interestingly, reduction in the flow rate and the nebulizing gas pressure favoured the appearance of homodimer ions. These observations have been noted and explored in the design of nanospray sources[182] to analyse non-covalent interactions, where flow rates and droplet size are significantly reduced, diminishing the need for extensive desolvation of ions.

1.2.2.2 Mass Analysers for Detecting Non-Covalent Complexes

A range of mass analysers have been used to probe aspects of non-covalent interactions. These have included quadrupole, ion trap, FTICR, sector and time of flight analysers. Commonly quadrupole mass analysers are limited to *m/z* ranges of 3000-4000 and have only unit *m/z* mass resolution and have therefore not been extensively used in analysis of non-covalent complexes. This is because when complexes form, the number of accessible charge sites is reduced, such that complexes are usually observed at higher *m/z* than the free protein. By decreasing the frequency applied to quadrupole mass analysers[183-185] complexes such as the ~200 kDa hexagonal bilayer hemoglobins have, however, been analyzed.[185] Ion trap instruments have similar mass limitations as quadrupoles, however, they are also able to have their *m/z* range extended.[186] Additionally they have the advantage of being able to trap ions in the gas phase to allow aspects of their structures to be probed by collision-induced dissociation[187,188] FTICR allows for similar experiments to be carried out. These analysers offer high mass resolution except when high *m/z* ions are being detected (>3000 *m/z*) as seen in the analysis of the trisaccharide (globotriaoside) complexed with the Shiga-like toxin (SLT) of pathogenic *E. coli*.[189]

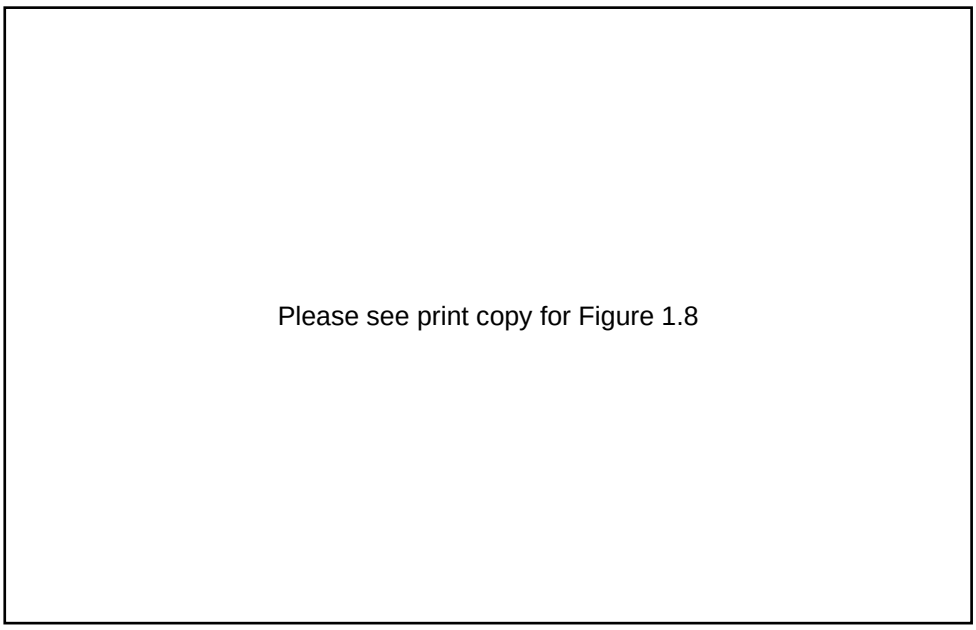


Figure 1.8. A schematic picture of a hybrid quadrupole time-of-flight mass spectrometer. Take from the Micromass Q-ToF 2™ users guide.

Sector instruments have been beneficial to analyse non-covalent interactions as they enable tandem mass spectrometry experiments at high collision energies to be carried out.[190,191] These instruments have largely been superceded for this application by hybrid-time-of-fight instruments. Time-of-flight (ToF) mass spectrometers are perhaps currently the most commonly used instruments used for analysis of non-covalent interactions. This is largely because of the higher *m/z* range and resolution compared with quadrupoles and traps. ToF instruments that utilize orthogonal acceleration (oa)[192] can be effectively coupled with continuous ion sources such as ESI to yield ToF analysers with good resolution and moderately high *m/z* range (typically < *m/z* 20,000). Additionally, oa-ToF analysers have commonly been coupled to quadrupole analysers to from a hybrid instrument[193] which allows for tandem mass spectrometry. Tandem or

MS/MS involves two stages of mass analysis. The first stage is used to select a precursor ion of a given *m/z*; thereafter the precursor ions are passed through a collision cell where they undergo collisions with an inert gas causing them to fragment. The resulting fragment or product ions are analysed in the second statge. A schematic of a hybrid Q-ToF is shown in figure 1.8.

For the analysis of large multiprotein assemblies it has been shown that collisional dampening/cooling of ions through the initial stages of the mass spectrometer significantly aids in the survival of high *m/z* (>1500) species.[194-198] By increasing the pressure in the initial stages of the mass spectrometer (seen in figure 1.8 as the hexapole ion bridge), collisions occur which reduce the kinetic energy of these large macromolecules allowing them to remain intact when focused through the remaining parts of the mass spectrometer.[199] The use of hybrid instruments (i.e. quadrupole time-of-flight, Q-ToF) is particularly useful for analysis of large complexes as extended mass range quadrupoles can be used to select ions of up to *m/z* 32,000 and perform tandem mass spectrometry experiments on these complexes.[200]

1.2.3 Solution and Gas Phase Characteristics of Non-Covalent Complexes Between Biomolecules and Small Molecules

Much research has been aimed at comparing the properties of non-covalent complexes in the gas phase, with their solution phase characteristics. That is, to determine whether the stoichiometry and relative binding affinities seen by ESI-MS studies relate to observations made on solutions. A large number of systems have been used to study these properties including protein-metal,[180,201] DNA-ligand,[202,203] protein-peptide,[167,204,205] and protein-sugar[206,207] complexes. Many of these have shown that observed stoichiometries in solution correlate with those determined by ESI-MS. In

contrast, some studies including the analysis of double stranded DNA complexes between non-complementary base pairs,[208] cyclodextrin inclusion complexes[209] and the observation of complexes with inactive enantiomers of substrate–enzyme[210] have showed that non-specific interactions can also be observed in ESI mass spectra. Non-specific interactions were thought to arise from the more concentrated environment which occurs following ionization or when particularly high concentrations of analyte were employed.

ESI-MS has been used to probe the stability of small non-covalent complexes by examining relative binding affinities. By analysing the relative intensities of ions produced when components of a complex have been varied, and performing competition experiments, relative binding affinities of components in the gas phase can be compared to solution phase data. An example of this can be seen in the analysis of DNA-drug complexes which offer an excellent candidate to monitor the relative binding affinities as there are a variety of different ligands capable of interacting with DNA strands (reviewed by Beck *et al.*[174]). Gabelica *et al*. have compared the binding affinities of a number of DNA intercalating and minor groove binders and have shown that the relative affinities of the drugs for sequences of double-stranded DNA compare well with solution data.[211] Protein-peptide interactions in Src SH2 domain protein-phosphopeptide complexes were investigated to compare the relative binding affinities in the gas phase and solution.[205] These complexes also showed a good correlation between solution and gas phase binding affinities.

The effect of the removal of water from non-covalent complexes involving hydrophobic interactions has been tested in several studies. In the analysis of acyl CoA binding

protein interacting with a series of acyl CoA derivatives by Robinson and co-workers it was noted that increasing the size of the hydrophobic acyl chain did not cause any significant change in the binding affinity of these ligands in the gas phase, despite variations in dissociation constants (10^{-7} to 10^{-10} M) measured in solution.[179] Wu *et al.* also observed similar behaviour in their analysis of interactions between carbonic anhydrase II and para-substituted benzenesulfonamide inhibitors.[212] These studies appear to indicate a decrease in the ability to detect hydrophobic non-covalent interactions in the gas phase. Gupta *et al.* were able to detect a protein-protein complex (θ-ε_{186}) by ESI-MS.[213] These subunits of *E. coli* DNA polymerase III interact largely through hydrophobic amino acids. This result suggests that hydrophobic interactions can be maintained on transfer into the gas phase, however, these interactions are not strengthened or may be weakened by removal of solvent. An increased dependence on electrostatic interactions is, however, thought to be the one cause of the formation of non-specific complexes in the gas phase.[208-210]

ESI-MS studies of the calmodulin-Ca^{2+} complexes[214] and the spermine and spermine-binding peptide complex[215] have indicated an increased dependence on electrostatic interactions. In solution, calmodulin has been shown to contain between four and six auxiliary metal binding sites in addition to the four Ca^{2+} sites essential for many calmodulin functions.[216,217] The auxiliary binding sites of calmodulin appear to bind Ca^{2+} quite weakly when mesured in solution (K_D ~10^{-3} M), however, these interactions are readily observed in the presence of high concentrations of Ca^{2+} by ESI-MS studies suggesting these interactions are stronger in the gas phase than solution.[214] In solution, the electrostatic interaction between spermine and spermine-binding peptide was 10^{-4} M.[173] ESI mass spectra of this complex were readily obtained and attempts to dissociate

the complex in collisionally activated dissociation (CAD) experiments resulted in the breakage of covalent bonds before the dissociation of the non-covalent complex.[215] This observation is consistent with the proposal that electrostatic interactions are strengthened in the gas phase.

There have been several studies where ESI mass spectra have been used to estimate dissociation (or association) constants.[206,207,218-222] In some of this work ESI-MS data have been used to generate Scatchard plots to determine dissociation constants.[218-220] This method has been applied to analyse the binding of tobramycin and paromomycin (antibiotics) to a 27-nucleotide RNA construct of the ribosome.[220] Intensities of peaks in the ESI mass spectra were used to calculate dissociation constants which were in agreement with those measured using solution based methods. In fact, the majority of the reported ESI-MS procedures to determine equilibrium constants have been consistent with solution phase data. Recently, however, De Vriendt *et al.* showed that equilibrium constant varied when determined in solution compared with the gas phase.[206] In their analysis the binding of three derivatives of the disaccharide inhibitor of the catalytic domain of cellobiohydrolase was examined. Non-specific interactions were observed for one on these inhibitors which resulted in a decrease in the dissociation constant from 35 μM in solution to 6.42 μM measured by mass spectrometry.[206] The authors accredited this to the non-specific interaction forming in the gas phase. As a result of these observations they suggested that care should be taken to carry out control experiments to determine the specificity of non-covalent interactions analyzed in the gas phase before attempting to determine equilibrium constants.

Smith and Light-Wahl[223] have attempted to address the ability to distinguish the occurrence of gas phase non-specific interaction from specific interactions by adding a number of additional steps in analysis. To be more confident that a specific interaction is being observed: (1) Changes in interface conditions which allow the favouring of complex ions over the dissociated species and in the preferred stoichiometry should be observed in ESI-MS spectra. (2) Complexes arising from weak interactions should readily dissociate under more severe interface conditions or tandem mass spectrometry experiments. (3) Modification of solution conditions (i.e. pH, temperature, addition of organic solvent) should result in dissociation of any specific interactions; and (4) changes of one of the complex components should cause a decrease in the abundance of complex ions. Daniel *et al*. have also suggested experiments where the substrate binding site is blocked by chemical modification, however, this method may be quite difficult to carry out without disturbing the overall native structure of the biomolecule.[221]

The ESI-MS analysis of non-covalent complexes where small ligands bind biomolecules appears to be more problematic. The increase in dependence of electrostatic interactions and the loss to some degree of hydrophobic interactions will have an effect on some complexes but not others. The extent of these variations and additional factors which should be taken into account for gas phase analyses have yet to be fully understood and will be a focal point of this study.

1.2.4 Mass Spectrometry Studies of Large Macromolecular Complexes

Macromolecular complexes may be comprised of proteins, DNA or RNA molecules and smaller ligands (including metal ions, heme ligands and nucleotide bases). If these

complexes remain intact in the gas phase, mass spectrometry provides one of the best methods for determining the precise stoichiometries of these complexes. For example, the hemoglobin $(\alpha\beta)_2$ tetramer[168,183] and insulin hexamer[224,225] produce the same stoichiometries in solution and gas phase experiments. For large complexes like the 500 kDa vanillyl-alcohol oxidase octamer[195] and the 100 kDa transthyretin-retinol binding protein complex,[226] correlations between X-ray crystallography and mass spectrometry data have also been observed.

Large macromolecular complexes involve a greater number of individual non-covalent interactions than smaller non-covalent complexes. Taking this into account, it has been hypothesized that a closer agreement between gas phase and solution phase data will be observed for macromolecular complexes than for smaller complexes, as the net effects of these interactions between complexes out-weighs the differences in gas phase stabilities of electrostatic and hydrophobic interactions.[227] This may allow for ESI-MS experiments to be carried out with higher confidence, enabling elucidation of macromolecular structures and functions.

The use of ESI-MS in the analysis of these complexes has been extremely beneficial in examining the supramolecular assemblies of macromolecules. Mutation of His61 of vanillyl-alcohol oxidase (a flavoenzyme) results in expression of flavin adenine dinucleotide (FAD)-free complexes. ESI-MS showed that the FAD-free protein was a dimer, in agreement with size exclusion chromatography data.[228] Addition of FAD to form a protein-ligand complex promoted formation of an octamer that is the biologically active form of the enzyme.[228] This study may suggest how other members of the flavoenzyme family become activated.

Another example of how ESI-MS can be used to study aspects of complex dissociation and provide structural information can be seen in the analysis of urease. Urease catalyses the hydrolysis of urea to ammonium and carbamate.[229] The crystal structure of urease from *Helicobacter pylori* shows that the protein is comprised of 12 αβ units which associate to form four $(\alpha\beta)_3$ complexes.[230] ESI-MS of this complex revealed a mass correlating to a $\alpha_{12}\beta_{12}$ complex when gentle source conditions were used.[231] By increasing the voltage applied to the cone and extractors at the source of the QToF instrument, dissociation of the $\alpha_{12}\beta_{12}$ complex resulted in the formation of $(\alpha\beta)_3$ species. Identification of these intermediates in the dissociation of the urease complex will aid in understanding some of the structural aspects of this complex. ESI-MS of macromolecules can therefore provide information about the association mechanisms and the higher order structures of proteins in native conditions.

The ability of ESI mass spectrometry to analyse large non-covalent complexes has also been demonstrated in studies on amyloid formation.[225,232-234] Amyloid fibrils are aggregates of partially unfolded proteins that have been linked to Alzheimer's disease and other conditions. Transthyretin is a plasma protein that has been shown to form amyloid fibrils in humans.[235,236] Under acidic conditions it appears that transthyretin tetramers dissociate into monomers which may be the precursor to amyloid fibril formation.[237] Robinson and co-workers have used ESI-MS to probe the stability of the wild type transthyretin tetramer compared with point mutated variants.[232] Under acidic conditions, wild-type monomers and tetramers appeared to be in equilibrium as judged by peaks in ESI mass spectra for both species. Wild-type tetramers were also more stable than the variants under less gentle MS conditions. This was confirmed by HDX

experiments where higher deuterium incorporations rates were observed for variants than wild-type proteins indicating variants were more unstructured. Additional aspects of this complex were later analyzed by the same group to probe the ability of 18 different ligands to bind and stabilize tetrameric structures of transthyretin.[233] The results pointed to the cooperative nature of ligand binding and supported the report by Peterson *et al.* that ligand binding can stabilize native conformations of transthyretin reducing the yields of amyloid fibrils.[238]

Insulin oligomers that may be precursors to amyloid have also been examined by ESI-MS.[225] At a pH of 2.0 and at a concentration of 2 mM, oligomers of up to 12 molecules of insulin were observed. Oligomeric species in biology are dynamic structures and are often in equilibrium with monomers or other oligomeric forms.[225] Exchange of insulin monomers was measured indirectly by hydrogen exchange monitored by ESI-MS. Exchange rates for oligomers were similar to those observed for the monomeric states suggesting that subunits were in a rapid equilibrium with monomeric insulin. A Q-ToF ESI mass spectrometer has been used to measure protein subunit exchange directly. This study exploited the high *m/z* range available for precursor *m/z* selection and used nanospray to minimise sample consumption. The exchange of subunits between the small heat shock protein (sHSP) complexes of PsHSP18.1 from pea and TaHSP16.9 from wheat was observed.[234] ESI-MS showed that both PsHSP18.1 and TaHSP16.9 were dodecameric assemblies. When these complexes were mixed, subunit exchange was observed to occur primarily by sequential incorporation of dimeric species. The authors proposed that *in vivo*, this dynamic process may provide a mechanism for these molecular chaperones to interact with partially unfolded proteins.

The stability of TaHSP16.9 dodecamers was also tested by attachment of a temperature-controlled nanoESI probe to a factory nanospray source.[239] As the temperature of the spraying solution was increased, predominantly monomers and small amounts of dimer were observed. At higher temperatures, larger species, being almost exclusively tetradecamers are observed. Therefore as the dimer dissociated at elevated temperatures, it may re-interact with a dodecamer species, dissociate to monomers, or remain as a dimers. This correlated well with earlier studies which showed subunit exchange of the complex occurred primarily by incorporation of dimers.[234]

Interestingly, performing collision-induced dissociation of the dodecamer complex within the collision cell of the mass spectrometer revealed a different dissociation mechanism. Monomers were released in which the 14^+ charge state was most abundant. No dimers or tetradecamers were observed in the gas phase dissociation. This suggests that the solution phase and the gas phase dissociation give very different results. Further studies of these systems may therefore aid in the understanding of the gas phase properties of macromolecular complexes.

1.3 Scope of this Project

In the present chapter, the application of ESI-MS as an alternative method for the analysis of protein dynamics has been discussed. Owing to the gentle nature of the electrospray process, biomolecules can be introduced into the gas phase without fragmentation, and possibly maintain the non-covalent forces that stabilise protein structure and interactions between ligands. The added ease of coupling ESI to techniques such as HDX and tandem mass spectrometry has allowed a variety of aspects of protein structure and function to be probed.

This thesis will examine the capabilities of electrospray mass spectrometry to probe protein structural fluctuations, protein ligands interactions and the dynamics of interactions within a macromolecular complexes. The strengths and weaknesses of using ESI-MS for these applications will be discussed. In chapter 3, HDX of cyclized and linear forms of N-terminal of DnaB have been examined to explore how native structural variations can be identified and quantified by ESI-MS. Complementary NMR data and additional information gained from HDX-ESI-MS studies have been used to show the effects of restriction of C- and N- termini by cyclization on the stability and folding of this protein.

Using this same system, the experiments in chapter 4 explore the ability of ESI-MS to detect changes in protein structure from changes in charge state distribution (CSD). The effect of polarity of ionization on the ability of ESI-MS to detect conformational changes of low pI proteins when sprayed from a neutral solution is examined. The possibility of significant structural differences following opposite polarity ionization will also be discussed.

To complement the studies described in chapter 4, chapter 5 examines the effect positive and negative ionization on detecting protein-ligand non-covalent complexes between calmodulin, calcium ions and antipsychotic drugs. The relative effects of charge neutralization versus conformational changes on CSDs is addressed. Interactions between calmodulin and antipsychotic drugs show how Coulombic repulsions between negative charge sites may affect the ability to detect protein-ligand complexes and conformational changes as a result of ligand binding.

Finally, the macromolecular assembly between subunits of the primary *E. coli* helicase DnaB has been analyzed using an extended mass range QToF instrument. The results of these experiments are shown in chapter 6. The ESI-MS data reveal previously undetected oligomers of DnaB subunits, which may be important for understanding the physiological functions of DnaB *in vivo*. Additionally, the association and dissociation mechanisms between DnaB and its binding partner DnaC have been probed by ESI-MS and ESI-MS/MS and these results are discussed in this final chapter.

Chapter 2
Materials and Methods

2.1. Materials

All reagents and solvents used were of the highest grade commercially available. MilliQ^{TM} Water (Millipore, Molsheim, France) was used for all experiments.

The proteins DnaB, DnaC and DnaB-N (N-terminal domain, residues 24-136, of DnaB) and mutants were kindy provided by Dr Nicholas Dixon (Research School of Chemistry, Australian National University). Ribonuclease A, α-lactalbumin, bovine serum albumin (BSA), magnesium acetate ($Mg(OAc)_2$), ethylenediaminetetraacetic acid (free acid, EDTA), calcium acetate ($Ca(OAc)_2$), trifluoperazine (dihydrochloride, TFP), imipramine (monohydrochloride, IPA) and β-γ-methyleneadenosine-5′-triphosphate (disodium salt, βγ-CH_2-ATP) were obtained from Sigma-Aldrich (St. Louis, USA). Dithiothreitol (DTT), adenosine-5′-triphosphate free acid (ATP), adenosine-5′-diphosphate free acid (ADP), and iodoacetamide were purchased from ICN Biomedicals (Aurora, USA). Deuterium oxide (D_2O) and acetonitrile-D3 were obtained from Cambridge Isotope Laboratories (Andover, USA). Ammonium acetate (NH_4OAc), ammonium hydroxide, acetic acid, formic acid, isopropanol and acetonitrile were obtained from Ajax Finechem (Seven Hills, Australia). Ultrafree centrifugal filters and Microcon centrifugal filters were obtained from Millipore (Bedford, USA), Dialysis tubing (10,000 molecular weight cut-off) was purchased from Crown Scientific (Moorebank, Australia). Bio-Rad DC Protein assay was obtained from Bio-Rad

(Hercules, USA). Nanoelectrospray capillaries (Au/Pd coated, medium size) were obtained from Proxeon (Odense, Denmark).

2.2 Protein Preparation

2.2.1 DnaB-N

DnaB-N (cyclized and linear forms; see chapter 3) stored in 50 mM Tris-HCl buffer (pH 7.6), 15% (*v/v*) glycerol, 100 mM NaCl and 5 mM $MgCl_2$, were taken from a –80 ^oC freezer, thawed on ice, and dialysed against 2 L of a 10 mM NH_4OAc solution at pH 7.2 (for deuterium exchange experiments, see chapter 3) or pH 7.6 (for charge state distribution studies, see chapter 4) at 4 ^oC. The dialysate was changed twice over 24 hours followed by a final change into a 10 mM NH_4OAc solution at the desired pH (total dialysis time = 36 hours). The proteins were then concentrated using Millipore Ultrafree centrifugal filters for deuterium exchange experiments and Microcon centrifugal filters for charge state distribution studies. For HDX experiments, the final concentration of DnaB-N mutants was $\geq$ 1 mM. For charge state distribution studies, concentrations were between 15 and 50 μM, and were diluted to 10 μM with ice-cold 10 mM NH_4OAc (pH 7.6) prior to ESI-MS analysis. Protein samples were kept on ice until used.

2.2.2 Ribonuclease A and α-Lactalbumin

Ribonuclease A and α-lactalbumin stock solutions were prepared at concentrations of 73 and 71 μM, respectively, in 10 mM NH_4OAc, pH 7.1, and kept on ice. "Native" samples were prepared by diluting the ribonuclease or α-lactalbumin stock solutions with ice cold 10 mM NH_4OAc, pH 7.1, to produce solutions at 10 μM and 5 μM, respectively. Samples were kept on ice until subject to MS analysis.

Reduced α-lactalbumin (5 μM) was prepared by addition of appropriate volumes of 6.5 mM DTT (in 10 mM NH_4OAc, pH 7.1) and 10 mM NH_4OAc pH 7.1 solutions. The final concentration of DTT was 2.8 mM (560-fold excess over α-lactalbumin). This solution was kept at room temperature while reduction occurred (commonly one hour). Samples were injected into the mass spectrometer without further manipulation to stop further reduction.

Conditions for preparing fully reduced ribonuclease A were as follows. A 10 μM solution of ribonuclease A was prepared in 30% (*v/v*) CH_3CN, 0.1 M DTT (from a freshly prepared 6.48 mM stock in 10 mM NH_4OAc, pH 7.1) and 70% (v/v) 10 mM NH_4OAc, pH 7.1. This solution was then heated at 45 ^oC for 45 minutes in a Polystat (Crown Scientific, Moorebank, Australia) controlled water bath and analyzed by mass spectrometry without further dilution or purification.

The number of disulfide bonds reduced by treatment with DTT were determined by mixing 100 μL of the α-lactalbumin/DTT solution or ribonuclease A/DTT/CH_3CN solution with 200 μL of 50 mM iodoacetamide at various time intervals and reacting for 30 minutes in the dark at room temperature. These mixtures were then dialysed against 3 x 2 L of 10 mM NH_4OAc adjusted to pH 3 with formic acid. Samples were then analyzed by ESI-MS. This method is slightly modified from the procedure outlined by Carver *et al*.[240]

2.2.3 Calmodulin

Apocalmodulin (apoCaM) was prepared for ESI mass spectrometry by dissolving an appropriate amount in water followed by dialysis at 4°C against: 4 x 2 L 10 mM NH_4OAc, 2 mM in EDTA, pH 5.6; 4 x 2 L 10 mM NH_4OAc, pH 5.1; and 1 x 2 L 10 mM NH_4OAc, pH 7.6. Typically, the concentration of the resulting apoCaM was ~100 μM. Calmodulin with its full complement of calcium (Ca_4CaM) was prepared by adding calcium acetate (5 mM in 10 mM NH_4OAc, pH 7.6) to apoCaM to give an 8-fold excess of calcium over protein. The final concentration of Ca_4CaM was ≥ 10 μM, and the mixture was treated on ice for 15 min. Drug/Ca_4CaM (or Drug/CaM) complexes were prepared by addition of an aliquot of the drug (5 mM TFP or IPA in 10 mM NH_4OAc, pH 7.6) to the protein giving a final protein concentration of 10 μM and a final concentration of the drug of 50 μM.

2.2.4 DnaB for Preparation of $(DnaB)_6$ and $(DnaB)_6(DnaC)_x$

DnaB (~7 mg/mL) was stored at –80 ^oC in 50 mM Tris-HCl buffer, pH 7.6, 25 mM NaCl, 10 mM $MgCl_2$, 2 mM DTT, 0.1 mM ATP, 1 mM EDTA and 20% (*v/v*) glycerol. Exchange into buffers suitable for mass spectrometry was carried out by taking 20-60 μL of stock DnaB, thawing on ice, and diluting to 400 μL with ice-cold 50 mM NH_4OAc, pH 7.6, 1 mM $Mg(OAc)_2$ and 0.1 mM ATP or ADP. Buffer exchange was then carried out using Millipore Biomax centrifugal filters (5,000 molecular weight cut off). Filtration units were equilibrated in this same buffer prior to application of the protein solution. Exchange was then carried out by centrifuging the sample in an Eppendorf bench top centrifuge (4 ^oC, 11,000 x*g*). Three additional 400 μL aliquots of the 50 mM NH_4OAc, pH 7.6, 1 mM $Mg(OAc)_2$ and 0.1 mM ATP or ADP buffer were applied before samples were concentrated to ~20 μL.

For nanoESI-MS the concentration of $(DnaB)_6$ was 40 μM. $(DnaB)_6$ was prepared by addition of appropriate volumes of $(DnaB)_6$, 3 M NH_4OAc (pH 7.6), 1 mM ATP or ADP (in 50 mM NH_4OAc (pH 7.6)), 10 mM $Mg(OAc)_2$ (in 50 mM NH_4OAc pH 7.6)) and 50 mM NH_4OAc (pH 7.6) to give final concentrations of 200-1500 mM NH_4OAc, 1 mM $Mg(OAc)_2$, 0.1 mM ATP or ADP and 10 μM protein in 10 μL (pH 7.6). The concentrations of these reagents used in individual experiments are given in chapter 6. These solutions were left on ice for at least 5 minutes prior to ESI-MS analysis.

2.2.5 DnaC

DnaC in 50 mM Tris pH 7.6, 2.5 mM NaCl, 10 $MgCl_2$, 2 mM DTT 1 mM EDTA and 20% (v/v) glycerol was removed from a –80 ^oC freezer and thawed on ice. To obtain nucleotide-free DnaC, samples were first diluted to 400 μL with 10 mM NH_4OAc, pH 5.5 (adjusted with acetic acid), before dialysis overnight against 2 L of 10 mM NH_4OAc at pH 5.5 (adjusted with acetic acid) at 4 ^oC, followed by 3 x 2 L changes of 10 mM NH_4OAc, pH 7.6.

Analyses of the interactions of DnaC with ATP, ADP or βγ-CH_2-ATP were carried out by preparing the individual nucleotides at approximately 2 mM in 10 mM NH_4OAc, pH 7.6. Appropriate volumes of DnaC and nucleotides were then diluted with 10 mM NH_4OAc, pH 7.6, to give a concentration of DnaC of 5 μM and a typical concentration of 15 μM for nucleotides. All solutions were kept at room temperature before analysis.

For the analysis of DnaB/DnaC complexes, DnaC was prepared as above, however, following dialysis, DnaC was concentrated to ~ 20 μL using Millipore Biomax centrifugal filters which had been equilibrated in 10 mM NH_4OAc at pH 7.6. For these

samples the concentrations of proteins were determined using the Bio Rad DC Protein assay as described in section 2.3.

2.2.6 $(DnaB)_6(DnaC)_x$ Complexes

DnaC was prepared as described in section 2.2.1.5, however, DnaB was exchanged into a 50 mM NH_4OAc, pH 7.6, 0.1 mM $Mg(OAc)_2$ and 0.1 mM ATP solution by the same method described in section 2.2.4. Complexes were prepared by taking sufficient volumes of $(DnaB)_6$ to produce 10 μM in 10 μL. An appropriate volume of a solution of 300 mM NH_4OAc, pH 7.6, 100 μM ATP and 100 μM $Mg(OAc)_2$ was then added to the $(DnaB)_6$ solution before an appropriate volume of DnaC solution was added to give ratios between 1:1 to 1:8 $(DnaB)_6$ to DnaC monomers. The final volume in each mixture was 10 μL. The solutions were then analyzed by nanoESI-mass spectrometry.

2.3 Protein Concentration Determination

Concentrations of proteins were determined by UV spectroscopy using either a Cary 500 scanning (Varian, Mulgrave, Australia) or a Shimadzu UV-1700 (Suzhou, China) spectrophotometer. Extinction coefficients at given wavelengths are given in table 2.1.

Table 2.1 The wavelengths and extinction coefficients used to determine concentrations of proteins following exchange into solutions suitable for mass spectrometry.

Please see print copy for Table 2.1

Typically a quartz cuvette with a 0.1 cm path length was used. This allowed analysis of protein concentrations in volumes as little as 300 μL. Owing to the decreased path length, samples were not commonly diluted to allow maximum absorbances to be measured. Absorbances were determined at room temperature.

Concentrations of DnaB and DnaC for analysis of $(DnaB)_6$ and DnaB/DnaC complexes were determined using a Bio Rad DC Protein assay. Standard solutions of BSA (1-15 mg/mL) were prepared in the same solvent used for DnaB and DnaC. An aliquot of 3 μL was taken from each standard and the DnaB or DnaC solutions and 37.5 μL of Bio Rad Reagent A was added to each aliquot. Samples were then mixed before 300 μL of Bio Rad reagent B was added. Samples were again mixed, pulse centrifuged, and left at room temperature for at least 10 minutes, but no longer than 40 minutes, before being analyzed on the Shimadzu UV-Vis spectrophotometer at 750 nm. BSA standards were analyzed in duplicate. From the standard curve determined for BSA, the concentrations of DnaB and DnaC monomers were interpolated.

2.4 Hydrogen/Deuterium Exchange (HDX)

The following procedure was found to be most reproducible when performing HDX experiments. Additional changes for particular experiments are outlined in the respective chapters.

2.4.1 Preparation of Ammonium Acetate in Deuterium Oxide (D_2O) Solution

An NH_4OAc solution in D_2O was prepared by adding 51 µL of a 2.93 M solution of NH_4OAc in H_2O to 15 mL of deuterium oxide (99.9% D). The pH was then adjusted to the desired value using deuterium oxide solutions containing 3% (*v/v*) ammonium hydroxide and/or 3% (*v/v*) acetic acid. When adjusting the pH, N_2 gas was constantly flushed over the surface of this solution. A thermostat-controlled water bath (Pharmacia Biotech, Uppsala, Sweden) was used to keep this at the desired temperature, prior to incubation with protein solutions.

2.4.2 Quenching Method for Deuterium Exchange

Protein solutions were prepared at 1 mM (unless stated otherwise). Aliquots (0.5 µL) of the protein solutions and 10 mM NH_4OAc in the D_2O solutions at the required pH were then separately equilibrated at the specified temperature for at least 5 minutes in a thermostat-controlled water bath. Protein samples were diluted 1 in 100 with the NH_4OAc in D_2O solution to produce a 10 µM protein solution and a deuterium percentage of 98.5±0.5%. At various times, deuterium exchange was quenched by diluting 4 µL of the exchanging solution into 36 µL of ice-cold quenching mixture (water: methanol: formic acid (~90:9:1 by vol) at a pH between pH 2.1 and 3.5. The pH of this solution was varied by modifying the amount of formic acid added. A 25 µL ice cold syringe was then used to inject 10 µL of the quenching solution into the mass spectrometer through a Rheodyne injector with a 10 µL sample loop. Ice-cold quenching mixture at the same pH used for quenched was used as the mobile phase at a flow rate of 50 µL/min using a Harvard Model 22 syringe pump (Natick, MA, USA).

2.4.3 Direct Injection for Deuterium Exchange

Direct injection experiments were carried out by preparing 5 μL aliquots of protein at a concentration of 1 mM. Protein solutions were then diluted with 10 mM NH_4OAc in D_2O at the required pH to give a protein concentration of 10 μM and deuterium percentage of 98.5±0.5%. This solution was then injected at room temperature (~25 ^oC) directly into the mass spectrometer at 10 μL/min using a Harvard Model 22 syringe pump.

2.5 Structure Visualisation

Visualisation of structures was performed using the RasMol™ software (version 2.6-ucb). Protein structures were downloaded from the RCSB Protein Data Bank (PDB).[244] PDB identification codes for proteins modelled in this thesis are given in table 2.2.

Table 2.2. The protein structures and the Protein Data Bank identification codes for proteins studied in this thesis.

Protein Structure	PDB ID
α-lactalbumin	1HFX
Apo-calmodulin	1CFC
Ca_4-Calmodulin	1CLL
Ca_4-Calmodulin + TFP	1CTR
Ca_4-Calmodulin + myosin peptide	1CDL

2.6 Mass Spectrometry

All mass spectra were acquired on either a Micromass Q-ToF2™ (Wyntheshawe, UK) or on a Waters extended mass range Q-ToF Ultima™ (Wyntheshawe, UK) mass spectrometer, each with a Z-spray source and an effective *m/z* range of 10,000 and ~150,000, respectively. For the Q-ToF-2™, protein solutions were injected using a Harvard Model 22 syringe pump (Natick, USA) at flow rates between 10 and 20 μL/min for ESI-MS (unless stated otherwise). Table 2.3 lists experimental conditions for the analysis of proteins using the QToF-2™ mass spectrometer. Variations from these conditions are noted in the respective sections of this thesis.

Table 2.3 ESI-MS conditions used for the analysis of protein samples on the Waters Q-ToF-2™ mass spectrometer. Identical conditions were used in positive and negative ion modes.

	Sample					
	DnaB-N	**Calmodulin**	**α-Lactalbumin**	**Ribonuclease A**	**DnaC**	**HDX**
Capillary (V)	2600	2600	2600	2600	2600	2600
Cone (V)	30	30-50	50	50	30	50
Source Block temperature (^oC)	40	40	40	40	40	40
Desolvation temperature (^oC)	40	40	120	120	120	120
HM/LM resolution	N/A	10	N/A	N/A	10	N/A
Collision cell (eV)	4	4	4	4	4	4
Transport (V)	2	2	2	2	2	2
Aperture 2 (V)	13	13	13	13	13	13
Acquisition time (sec)	5	5	30	5	5	$5^a(30^b)$
Mass range (*m/z*)	500-5000	500-5000	500-3000	500-3000	500-4000	500-3000
Acquisition / spectrum	20-30	20-30	2	20-30	20-30	$4^c(2)^d$

[a] An acquisition time of 5 seconds was used for samples which has been quenched.
[b] An acquisition time of 30 seconds was used for direct injection methods.
[c] For quenched samples 4 acquisitions were combined for each spectrum.
[d] For directly injected samples 2 acquisitions were combined for each spectrum.

Table 2.4 MS profile used to analyse protein samples using the Waters Q-ToF-2™ mass spectrometer.

	Dwell Time (% of each scan)	**Scan Time (% of each scan)**
mass range minimum	20	60
0.3 x maximum of mass range	20	0
0.3 x maximum of mass range		

Spectra acquired using the Q-ToF-2™ were externally calibrated using a solution of CsI (750 μM, in water) over the same *m/z* range outlined for each sample in table 2.2. MS profiles were as shown in table 2.4. Typically spectra were summed, baseline subtracted using a polynomial order of 11, 40% below the curve and smoothed using a Savitzky Golay smoothing method, smoothing 20 channels twice for each spectrum.

For the analysis of HDX kinetics, the electrospray ion series from baseline-subtracted, data were transformed to a mass scale using the transform function in the MassLynx Software™. Spectra were subsequently smoothed using the Savitzky Golay algorithm, smoothing 30 channels twice for each spectrum. Peak areas were determined using the MassLynx peak integration tool.

The factory-modified Q-ToF Ultima™ with a 32,000 *m/z* resolving quadrupole was used for the analysis of the DnaB and DnaB/DnaC protein complexes (in chapter 6). Experiments were performed on this instrument using a nanospray source. Typical conditions for MS and MS/MS analyses using this instrument are listed in table 2.5.

Table 2.5 The ESI-MS conditions used for the analysis of DnaB and DnaB/C complexes on the Waters Q-ToF Ultima™.

Instrument parameter	**MS**	**MS/MS**
Capillary (V)	1200-1600	1500-1600
Cone (V)	550	650-700
RF lens1 (eV)	190	190
Source block temperature (^{o}C)	N/A	N/A
Collision cell (eV)	5	80-110
RM/LM resolution	2	5
Transport(V)	5	5
Aperture(V)	12	12
Ion optics pressure (mbar)	0.1	3×10^{-3}
Collision cell gas pressure (bar)	0.75	0.75-1
Acquisition time (Sec)	2	2
Mass range (m/z)	500-20000	500-30000
Acquisition /spectrum	15-30	10-30

Spectra acquired on the Q-ToF Ultima™ were externally calibrated using a 10 mg/mL solution of CsI in 70 % isopropanol. Manual MS profile was turned off to allow for acquisition over the range (500-20000 *m/z*). Processing of spectra was carried out as described above for the Q-ToF-2™.

Chapter 3

Deuterium Exchange Characteristics of Linear and Cyclized Mutants of the N-terminal of DnaB (DnaB-N)

3.1. Introduction

The secondary and tertiary structures of proteins are stabilized by interactions between amino acid side chains such that hydrophobic residues in the protein core are protected from an aqueous environment by an outer hydrophilic layer. C- and N- terminal residues of proteins are often highly flexible owing to a lack of steric restraint and can be considered a weak point in protein structures where protein unfolding may begin. This poses the question as to what are the effects of restricting the C- and N- termini of proteins.

Answers to some of this question became accessible in the mid 1990's when a number of protein structures were identified which had a head-to-tail cyclized polypeptide backbone.[245-251] These molecules showed a high tolerance towards thermal and enzymatic degradation, and it has been suggested that they may have roles as antibacterial and antifungal agents[245,247,249] and as protection against proteolysis.[251] These molecules are also significantly more stable than their linear counterparts and are therefore prime candidates as therapeutics with improved bioavailability and as thermodynamically more stable enzymes that could be used in industrial processes at higher temperatures to improve efficiency.[252-255]

A number of methods have been used to link C- and N-termini *in vitro*. Goldenburg and Creighton were the first to attempt this by chemically cross-linking the termini of bovine pancreatic trypsin inhibitor.[256] This method, however, did not provide a general route for production of homogeneous samples. An alternative method involves chemical or enzymatic ligation.[257] Chemical ligation methods carried out under native conditions require the incorporation of an N-terminal cysteine and a C-terminal α-thioester in replacement of a carboxylic acid group. The N-terminal cysteine residues can be readily incorporated into these structures by use of recombinant protein technologies.[258] The α-thioester residue previously had to be synthetically attached,[259] a problem for large proteins, however, it can now be generated using protein splicing processes which themselves can generate cyclized proteins.[257,260,261]

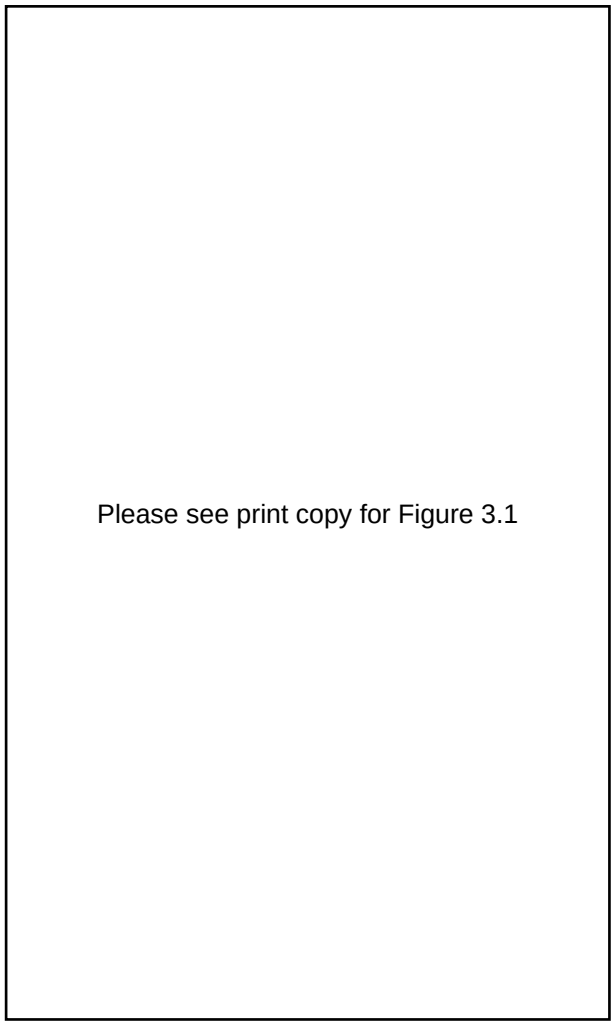


Figure 3.1 The process of cyclization by a split-intein system. Activation of the N-terminal splicing junction occurs in step one by an N-O or N-S acyl shift. In the second step, the ester or thiol ester bond is attacked by the hydroxyl or thiol group of a Ser, Thr or Cys residue at the C-terminal of the extein. Step three results in the cleavage of C-terminal intein domain. Finally, a second acyl shift occurs resulting in the production of a cyclized protein. Adapted from Trabi and Craik.[262]

Protein splicing reagents are sequences of polypeptide chains, known as inteins, which are able to excise themselves post-translationally from a nascent polypeptide chain. This results in the joining of the flanking regions of the polypeptide (exteins) through the formation of a peptide bond. The site of this splicing reaction can be identified from the conserved amino acid residues at the intein-extein junction where at either side of the intein there is a thiol or hydroxyl side chain; next to this there is an asparagine residue only at the C- terminal side of the intein (see figure 3.1). The reaction involves a series of four steps, (described by Paulus[260,261]) three of which require the catalytic properties of the intein.

Inteins can be expressed *in vivo* in two separate halves (i.e. N- and C- terminal domains, as seen in figure 3.1) as for the *Mycobacterium tuberculosis* RacA intein[263,264] and the cyanobacterium *Synechocystis sp*. DnaE intein[265] which when brought back together form an active molecule. These systems provided the inspiration for recombinant production of inteins, which produce cyclized proteins. Proteins that have been cyclized by this method include the green fluorescent protein,[266] β-lactamase,[267] and dihydrofolate reductase.[255] Reasonable quantities (>mg amounts) of cyclized proteins have been obtained by this method allowing a variety of protein stability measurements to be carried out. For example, cyclized and linear forms of proteins were subjected to heat and urea-denaturation.[266-269] It was determined that the increase in stability commonly observed for cyclized proteins is the result of a decrease in the entropy of the protein in folded and intermediate states.[267,268] Further, cyclized proteins hold great promise for determining the roles of termini in protein folding and unfolding motions. Molecular motions play an important role in protein functions including molecular recognition and enzymatic activity. Evidence of these motions can be seen through

hydrogen exchange experiments where exchange requires unlocking of hydrogen bonding networks through local structural changes. Using cyclized proteins the roles of the C- and N- termini in these motions may be elucidated.

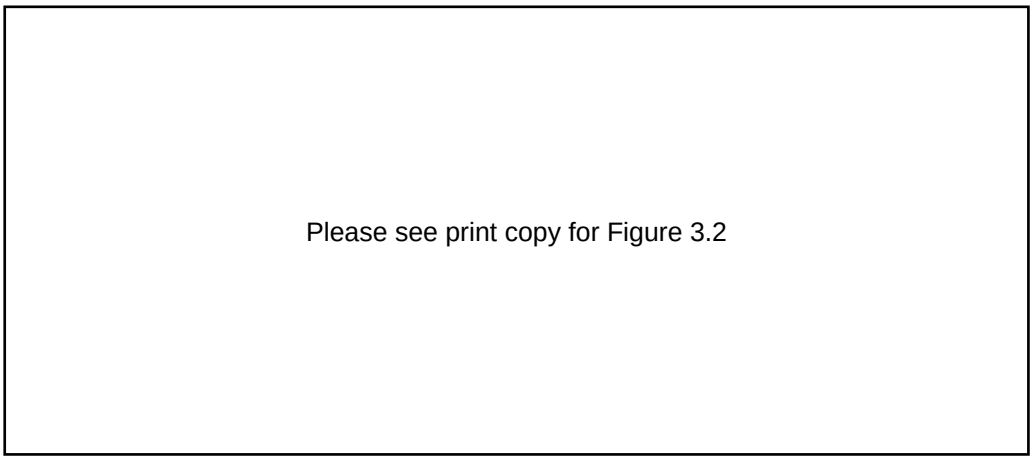


Figure 3.2 The NMR determined structures of: (a) Cyclized (Cz), (b) X-Linear (X-Lin), and (c) native (Lin) DnaB-N. Adapted from Williams *et al.*.[270]

Dixon and co-workers have prepared linear and cyclized forms of the N-terminal domain (residues 24-136) of DnaB. The cyclized protein was prepared using a split mini intein system from *Synechocystis sp.*[241,270] DnaB is the primary helicase of *E. coli*, which in its active state exists as a hexamer. The specific functions of this protein will be discussed in chapter 6. The DnaB monomer is comprised of three regions: a hydrophilic region (residues 1-20), an N-terminal compact domain with a possible hinge region (residues 21-172), and an ATPase active site in the C-terminal domain (residues 173-470).[271,272] The structure of the N-terminal domain has been determined by X-ray crystallography[273] and NMR[274] showing that this protein is largely structured between residues 30-134 with six helical segments. The hydrophobic core of this protein is centred around the C-terminal end of helix 1 (see figure 3.2 (c)) and charged residues have been shown to distributed over the surface of the protein.[274] The N- and C-

terminal residues are in close proximity to each other and appear to be quite flexible.[274,275]

Williams *et al.* compared a non-cyclized and cyclized form of this domain and found from differential scanning calorimetry, that the cyclized form was significantly stabilized ($\Delta\Delta G$ = 1.9 kcal/mol) compared to the linear protein form.[241] These structures were not ideal candidates for the comparison between linear and cyclized proteins as the cyclized form contained a additional nine-amino acid linker to allow cyclization of this protein occur. Further, a dimer of the protein was observed in solution at high concentrations.

3.2 Scope of Project

Linear (X-Lin) and cyclized (Cz) forms of the N-terminal of DnaB (residues 24-136, DnaB-N) were prepared in the Dixon group (Australian National University; figures 3.2 (b) and (a), respectively). The X-Lin form also contained the additional nine amino acid sequence, therefore, Cz-DnaB-N and X-Lin-DnaB-N have the same amino acid composition and differ only by the mass of one water molecule. NMR spectroscopy showed that addition of linker residues did not significantly change the protein structure with only a three residue extension of the C-terminal helix and a slight reorientation of the side chain of Tyr 60.[241] Additionally, Phe102 was changed to Glu to minimize dimerization. These alterations did not produce any significant changes in the structures of these proteins.[270]

In this work, the rates of unfolding of linear and cyclized DnaB-N were measured by diluting proteins into ammonium acetate solutions prepared in D_2O, and following the increase in mass by ESI-MS. The increase in mass corresponded to exchange of amide protons for deuterons. Two methods were employed to monitor hydrogen-deuteium exchange (HDX) by ESI-MS: (i) direct injection of reaction mixtures, and (ii) quenching of HDX by addition of acid. The relative merits and disadvantages of these methods are discussed.

The rate of HDX for cyclized DnaB-N was 10-fold slower than for the linear version of the protein. Further, there was evidence consistent with the existence of an intermediate in the unfolding process. The exchange profile was consistent with an EX1 unfolding mechanism.[276,277] The implications of these results with respect to the mechanism of unfolding and the structure of DnaB-N are also discussed.

3.3 Results and Discussion

As proteins unfold, their amide protons are exposed. If the protein is diluted into D_2O, the readily exchangeable protons exchange for deuterons, increasing the mass of the protein.[94] The readily exchanged protons include amide protons and protons of acidic and basic (including the N-and C-terminal groups) sites. Commonly, a quenching solution (at pH 2-3 and 0 ^oC) is added to stop exchange. The rate of exchange for amide protons is decreased by ~10^5 under these conditions.[94] Englander and co-workers have calculated the half life of an amide proton to be between 15 and 30 minutes under these conditions.[79,278] Owing to a significantly greater rate of exchange at acidic and basic sites (e.g. Asp, Arg) under these conditions, incorporated deuterons become re-protonated by hydrogens, so that the deuterons remain only on the exposed amide

backbone. Coupling this quenching method to ESI-MS has allowed for the analysis of calmodulin conformational dynamics as it interacts with calcium,[279,280] and the identification of folding intermediates of lysozyme.[89] Additionally, the quenching method allows sufficient time for desalting and even digestion of protein samples prior to analysis.[91-93]

ESI-MS has the advantage over other techniques that it can allow exchanging solutions to be injected straight into the mass spectrometer without the need for quenching, allowing all exchanged protons to be detected. This is referred to as the direct injection method. Simmons *et al*. have successfully applied the direction injection method in conjunction with time-resolved techniques to study myoglobin unfolding dynamics.[281] This allowed analysis of exchange characteristics to be determined over a millisecond time scale.

3.3.1 Development of HDX Techniques for Analysis of DnaB-N

Figure 3.3 shows the results of HDX experiments where Cz-DnaB-N was diluted into 10 mM NH_4OAc in D_2O at a pH of 7.2.

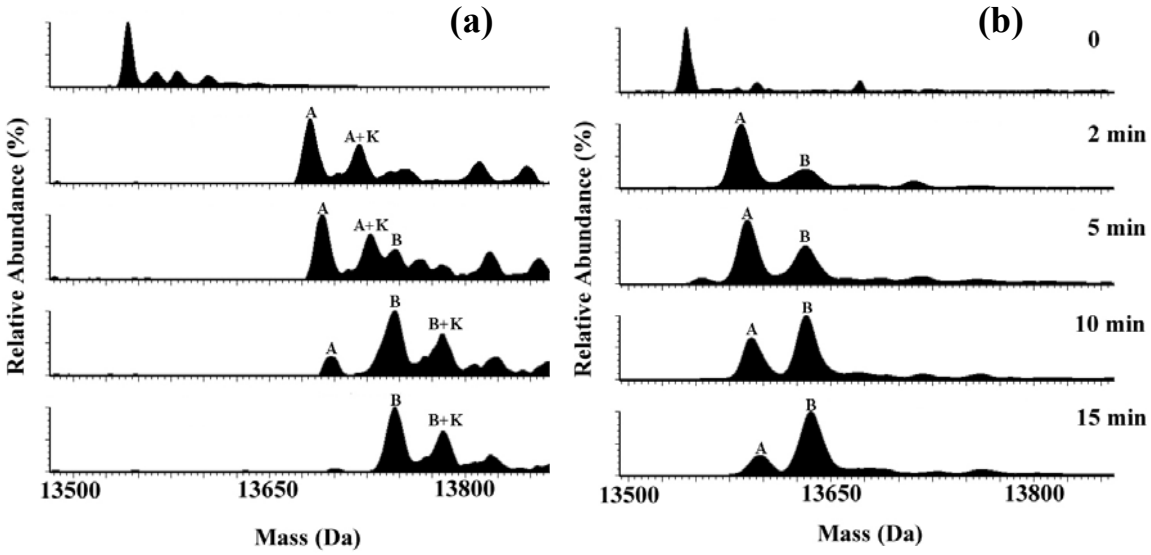


Figure 3.3 Hydrogen deuterium exchange of Cz-DnaB-N when analyzed by: (a) Direct injection at room temperature (~25 ^{o}C) and (b) quenching with acid at pH 2.1 following incubation in a water bath at 25 ^{o}C.

The figure compares results obtained by direct injection (figure 3.3 (a)) and by quenching with acid (figure 3.3 (b)). The spectra were transformed to a mass scale to simplify interpretation of results. Both methods show two populations of hydrogen/deuterium exchanged molecules (labelled A and B). The spectra obtained by the direct injection method were complicated by the presence of adducts at 39 mass units above the mass of the A or the B state, suggesting that these were potassium adducts. Careful investigation of buffers and Cz-DnaB-N stocks used in these experiments suggested that the source of this potassium was likely to be from the D_2O solution. In HDX experiments carried out in other laboratories using D_2O from the same source,[281] these adducts were not observed. DnaB-N used in the current experiments has an isoelectric point of 4.6[282] making this protein more vulnerable to adduction by cations.

Using the acid quenching method, these adducts were significantly reduced, since the D_2O solution was diluted 10-fold into acid for analysis, displacing potassium ions.

Following exchange in a solution of ~99% deuterium, the quenching method also showed that there were up to 69 and 108 amide protons exchanged for A and B states respectively; a larger mass shift was observed for the direct injection method, with 173 and 220 for A and B states, respectively. This is because in the quenching method the fast exchanging sites present at acidic and basic residues of the protein were reprotonated but in the quenching method all amide proton exchange was negligible over the time taken for analysis. The direct injection method allows all sites to remain deuterated, so that all exchangeable protons can be monitored.

The quenching method relies on the negligible rate of amide proton (or deuterons) exchange at low pH.[94] Amide exchange is base catalysed, with the rate increasing approximately 10-fold for every pH unit above 4. Below pH 2, the process is acid-catalysed.[79] Therefore, it is important to establish that under the experimental conditions, there is insignificant back exchange occurring. This can be ensured by using low temperatures for quenching and injecting into the mass spectrometer. Further, the conditions for quenching were tested using several different quenching conditions as outlined below.

To maintain low temperatures, aliquots of quenching buffer were kept on ice. Additionally, as described in the materials and methods (chapter 2), 10 μL aliquots of quenched protein were injected into the mass spectrometer using a Rheodyne injector. Quenching solution was used as the mobile phase which was passed through a 200 μL sample loop surrounded by ice containing NaCl. A flow rate of 50 μL/min was used to ensure samples were transported to the source of the instrument in a short period of time. These precautions were essential to minimize back-exchange.

The effect of pH of the quenching solution on the extent of back-exchange was also examined by determining deuterium incorporation at 10 ^{o}C by Cz-DnaB-N over the time period 30-210 minutes using quenching solutions at pH 2.1, 2.5, 3.0 and 3.5. Figure 3.4 shows a plot of these results when quenching solutions at pH 2.1 and 3.5 were used.

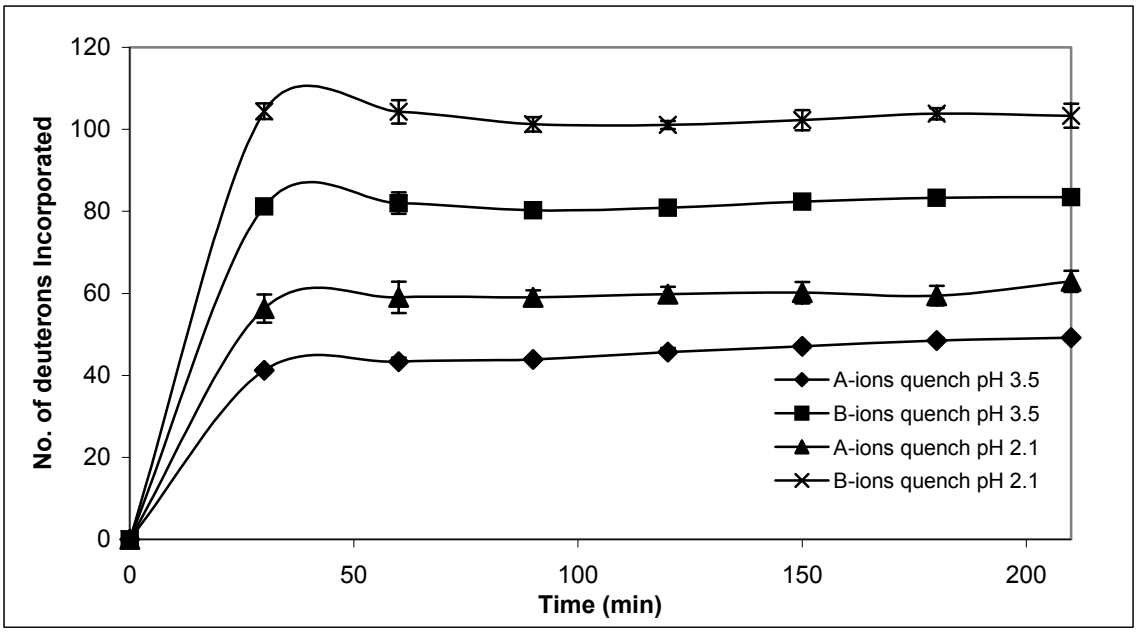


Figure 3.4. A comparison of quenching buffers at pH 2.1 and 3.5 when analysing the hydrogen-deuterium exchange characteristics of Cz-DnaB-N at a pH of 7.2 and 10 ^{o}C.

Using a quenching solution at pH 2.1, A-ions (▲) incorporated ~60 deuterons in the structure, and between ~102 deuterons in the B-ions (χ) over the time course of these reactions. Very similar deuterium incorporation levels were also observed when quenching solutions at pH 2.5 and 3.0 were used. When the pH was increased to 3.5, however, a significant drop in the number of deuterons incorporated into the structures A and B occurred with ~45 deuterons incorporated for the A-ions (♦), and ~82 deuterons for the B-ions (■). This is a drop of approximately 14 deuterons for the A-

state and a 20 deuterons for the B-state. This suggests that for both states there were a number of amide deuterons that were back-exchanging (replaced by a proton) when a pH of 3.5 was used for quenching. This observation shows that for the analysis of the DnaB-N mutants, the pH of the quenching solution should be between 2.1 and 3.0. For the analyses in the remainder of this chapter, pH was maintained at 2.1.

It has also been reported that back-exchange can occur in the source of the mass spectrometer as the increase in temperature may lead to exchange with atmospheric water.[281,283,284] Comparison of HDX results for DnaB-N using source desolvation temperatures of 40 and 120 ^{o}C showed that increasing the temperature to 120 ^{o}C did not result in any significant back exchange. The increase in the desolvation temperature assisted in the analysis of DnaB-N as resolution improved, however, most likely as the result of improved desolvation.

3.3.2 Unfolding Properties of DnaB-N Mutants

The experimental conditions determined above were used as the basis for experiments aimed at comparing the unfolding rates of Cz- and X-Lin DnaB-N. As discussed in chapter 1, protein folding is a series of events which may involve the sampling of intermediate states prior to the formation of final structure. Analyses of protein folding often begin using a protein in the presence of a denaturant, following dilution to decrease the concentration of the denaturant, allowing the protein to refold. This, however, does not necessarily present a true reflection of the folding processes that occur *in vivo* as folding begins from non-denaturing environments, and may affect the intermediates which are detected in subsequent folding events.

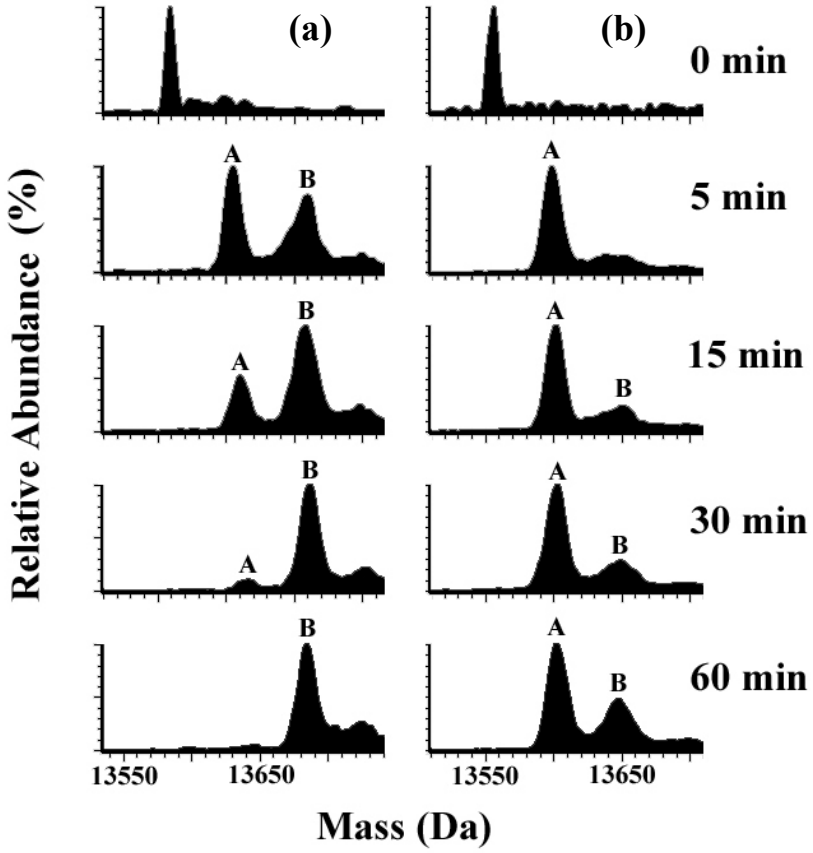


Figure 3.5 A comparison of the HDX exchange characteristics of: (a) X-Lin and (b) Cz-DnaB-N at a pH of 7.2 and 10 ^{o}C. A quenching method as described in section 2.4.2 was used in these experiments.

An alternative approach is to analyse transitions states sampled by proteins under native or near native conditions (i.e. with the addition of slight accesses of acid). These intermediate states are readily analyzed by hydrogen deuterium exchange (HDX) experiments. NMR and MS techniques have been shown to be complementary when examining HDX characteristics of proteins.[277] The site-specific nature of multidimensional NMR experiments allows the amides that exchange to be specifically identified. This is not possible from simple MS experiments. MS, however, has an advantage over NMR as it can readily identify populations of proteins in folded or unfolded states allowing rates of folding to be determined.[94,285] NMR experiments carried out by our collaborators showed that the average exchange rate for amide

protons was 10 ± 2 times faster for X-Lin DnaB-N than Cz-DnaB-N.[270] This difference in exchange rates was constant for the amide protons over the whole molecule. In the present work, experiments were carried out under conditions similar to those for NMR (pH 6.8-7.8 and at 10 ^{o}C). The extent of HDX measured by ESI-MS provided complementary information to the NMR data. In order to be compatible with ESI-MS, however, buffer conditions for exchange were different (10 mM NH_4OAc) compared to those for NMR (10 mM sodium phosphate with 0.02% (w/v) sodium azide). Figure 3.5 shows a time course of HDX (quenching method) for X-Lin-DnaB-N (figure 3.5 (a)) and Cz-DnaB-N (figure 3.5 (b)) at 10 ^{o}C and pH 7.2. Both exhibit a transition from an A state to a B state with both proteins having 50 ± 2 deuterons incorporated in the A-state and 100 ± 3 in the B-state. The observation of a distinct intermediate state in the unfolding of a protein under native conditions is unusual.[286]

HDX requires hydrogen bonding networks to be broken through structural changes in native protein structures. The following equation (3.1) describes the kinetics of exchange:[276,287]

$$\text{X-H}_{\text{close}} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \text{X-H}_{\text{open}} \xrightarrow{k_{ex}} \text{X-D} \qquad \textbf{(3.1)}$$

where k_{op} and k_{cl} are the opening and closing rates for the protein respectively, and k_{ex} is the chemical rate of HDX. This scheme suggests that native protein structures (X-H_{close}) are in equilibrium with an unfolded state (X-H_{open}). There are two extremes of behavior based on the relative magnitudes of k_{cl} and k_{ex}. When $k_{ex} >> k_{cl}$ then every time a protein unfolds, exchange will occur. This exchange mechanism is known as

EX1.[79,288] Under these conditions intermediate partially unfolded states can be detected by ESI-MS as several distinct populations of protein exist at different masses (as observed in figure 3.5). When $k_{cl} \geq k_{ex}$, the mass of a protein will increase gradually with time. This process is called an EX2 mechanism.[79,288]

Observation of an EX1 mechanism at neutral pH is very unusual. Commonly, to observe an EX1 mechanism, the pH of the solvent is either raised or lowered and the temperature of the solvent is raised to equal to or above the melting temperature[†] of the protein.[285,286,289,290] These denaturing conditions favour the sampling of intermediate or unfolded states since the extreme conditions disfavour refolding causing k_{cl} to decrease. This has been observed in ESI-MS studies of thioredoxin and some of its alkylated cysteine derivates,[285,289] a domain of ovomucoid protein,[290] and the 62-residue IgG binding domain of protein L.[286] We are observing here, however, the globular or subglobular transitions occurring under native conditions indicating the rate of protein closing to be quite slow. For a helical protein such as DnaB-N which has a low calculated contact order (the average separation in sequence between residues that are in contact with each other in the native structure) of 12.6,[270] this is very unusual as rates of folding of these proteins are commonly very fast when compared to the rates of proteins with a high contact order.[9] The nature of the α-helix assembly in DnaB-N is very unusual, only being observed for DnaB.[273] This may explain why folding rates are quite slow in comparison to a typical helical bundle such as the B domain of Staphylococcal protein A where it is believed that preorganisation of one or more secondary structures occurs in the unfolded protein enabling folding to be extremely fast.[291] Identification of

[†] Melting temperature is defined as the transition point where the ratio of native and denatured states is 1. Commonly this is determined using circular dichroism.

the structure of the unfolded form of DnaB-N and analysis of its transitions will be helpful in understanding this slow folding process.

To determine the rates of transition from the A-state to the B-state the relative populations of each state are required to be determined. This can be deduced from intensities of the A and B states, however, as seen in figure 3.5, peak shapes may vary during the course of the experiments. In figure 3.5 (a), close inspection of the B-state indicates that at 5 minutes this peak is skewed, however, as time increases the peak becomes more uniform in shape. Therefore, if peak intensities were used to determine relative populations of A and B states, this may not be a true reflection of abundances of ions. Using peak areas would provide a more accurate account of populations of molecules in these states.

There are a number of possible explanations for the skewed shape: First, salt adducts of the A state may be present in sufficient quantities partially overlapping B-state peaks. This is unlikely, however, as B-states were still uneven in shape when the intensity of the A-state decreased after 15 minutes. Secondly, ineffective quenching may have occurred causing some B-state molecules to revert to a B-state with slightly fewer deuterons incorporated. Finally, there may be two different forms of the B-state indicating an intermediate at slightly lower mass than the mass of the final B-state. Resolution in these experiments was not sufficient to unambiguously identify if this transition occurred, however, it does point out that peak shapes may be a sensitive indicator for populations of amide protons that undergo different rates of exchange.

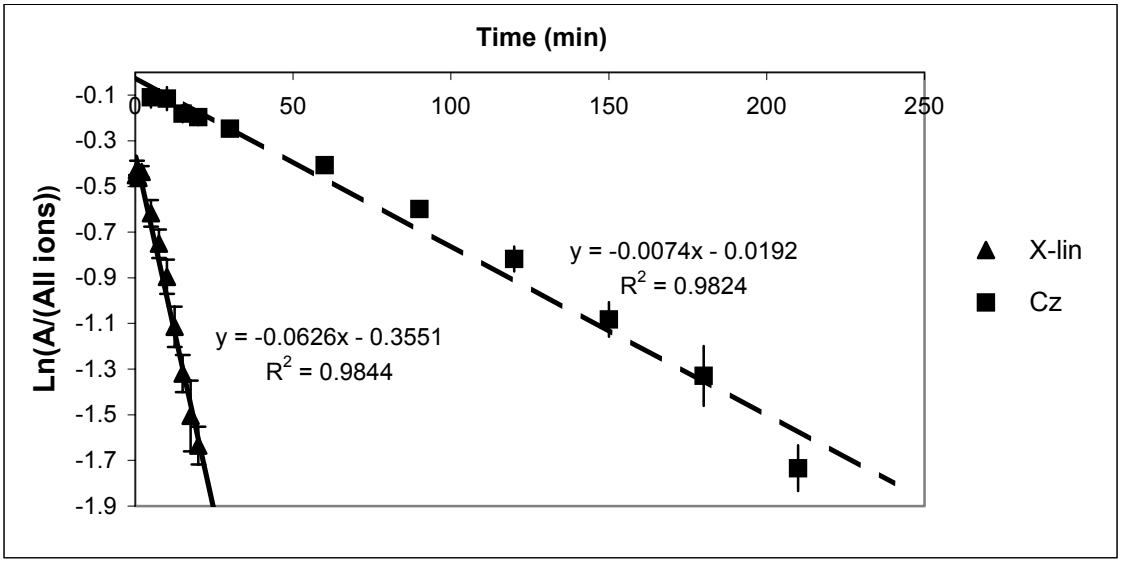


Figure 3.6 Pseudo first order ln plots of X-linear (X-Lin) DnaB-N and cyclized (Cz)-DnaB-N at a pH of 7.2 and 10 ^{o}C. Plots are the natural logarithm of the area of A-ions determined from mass spectra divided by the area of all ions against time. Each point represents the average of 5-9 separate experiments.

For analysis of HDX rates of X-Lin and Cz-DnaB-N, peak areas were used to calculate relative populations of A- and B-states as this allowed accommodation of peak broadness in the B-state. This study therefore assumed in the B-state there was only one population of molecules. By taking the area under the A-state peak and dividing it by the area under all peaks from all peaks in the spectrum at each time point, a first order plot for the loss of the A-state was obtained as shown in figure 3.6. The plots for both X-Lin- (▲) and Cz- (■) DnaB-N were obtained from an average of 5-9 separate experiments. The rate for Cz-DnaB-N was 0.0074 ± 0.0009 min^{-1} and the rate for X-Lin-DnaB-N mutant was found to be 0.063 ± 0.006 min^{-1}, approximately 10 times faster than the rate for the cyclized protein. As expected, the rates of appearance of the B-state were the same within experimental error. This shows that linking of the C- and N-termini reduced the rate of unfolding by a factor of ten. It appears that the restriction placed on the DnaB-N structure by cyclization of the C- and N- termini favours the folded state reducing its abilities to access the intermediate or unfolded state. This has

also been seen in the analysis of the N-terminal Src homology 3 domain where truncated forms of this structure were rescued from transition states by cyclisation.[292] In the current study, it is interesting to note that the cyclization did not fully stop DnaB-N from sampling these unfolding states suggesting that there was enough flexibility to allow unfolding to occur.

Comparing results obtained by ESI-MS with those obtained in the collaborative NMR study[270] revealed that a ten-fold increase in average exchange rates for Cz- compared with X-Lin DnaB-N was also observed in NMR amide exchange data. The nature of NMR experiments allowed exchange rates of individual amides to be determined, however, by determining the average exchange rate over all amide protons this ratio was also obtained. Although the ratio determined by ESI-MS and NMR was the same within experimental error, the absolute rates were different The difference may be a result of the different buffers required for the two experimental methods which have different ionic strength. DnaB-N has previously been shown to be sensitive to different buffers.[270]

NMR data also indicate for both proteins that helical segments exchanged significantly slower than linking residues. The six helical regions contain a total of approximately 48 residues which exchange slowly (see Williams *et al.*[270]). This number correlates closely to the number of amide protons that exchange during an unfolding event in ESI-MS spectra. This suggests that a globular unfolding process may be occurring allowing all amides to be exposed to the solvent for exchange.

3.3.3 Effect of Temperature on Exchange Rates

HDX rates for X-Lin- and Cz-DnaB-N were compared at 10 °C, 25°C and 35 °C at pH 7.2. These results are shown in table 3.1.

Table 3.1 The rates of unfolding of X-Linear and cyclized-DnaB-N at pH 7.2 at various temperatures.

	k_{op} (min^{-1})		
	X-Linear	**Cyclized**	$k_{op\ X\text{-}Lin}/k_{op\ Cz}$
10 °C	0.063 ± 0.006[a]	0.0074 ± 0.0009[a]	9 ± 2
25 °C	0.6 ± 0.2[a]	0.091[b]	6
35 °C	N/A[c]	0.76[b]	N/A

[a] Rates were determined from the average of 5-9 separate experiments. Errors are expressed as ± 1 standard deviation.
[b] Rates were determined from only one experiment. No standard deviation could be determined.
[c] Transition from the A-state to B-state was not observed over the time course of this experiment.

As expected, the rates of unfolding increased with temperature. At 35 °C, the rate for X-Lin-DnaB was too fast to be able to observe any A-state molecules. At all temperatures, the cyclized form was more stable than the linear form. Increased thermal stability appears to be a characteristic trait for cyclized proteins.[241,255,266,267,293]

3.3.4 Confirmation of EX1 Exchange Mechanism

EX1 and EX2 exchange can be differentiated, particularly where NMR has been used to measure HDX, by carrying out experiments at different pH values.[277] Since EX1 is the limiting case where $k_{ex} \geq k_{cl}$ (scheme 3.1), the rate is independent of the chemical rate of amide exchange. Therefore, an increase of k_{ex}, catalysed by base as the pH increases, does not change the observed rate of unfolding. In contrast, in the case of EX2, $k_{cl} >>$

k_{ex}, and the rate is dependent on k_{ex} which is pH-dependent. Table 3.2 shows the effect of pH on the rate of unfolding of X-Lin and Cz-DnaB-N. The pH values were between pH 6.8 and 7.8.

Table 3.2 The rates of unfolding of X-Lin- and Cz-DnaB-N at various pH values at 10 °C.

	k_{op} (min^{-1})		
	X-Linear	**Cyclized**	$k_{op\ X\text{-}Lin}/k_{op\ Cz}$
pH 7.8	0.042 ± 0.003^a	0.0045 ± 0.0004^a	9 ± 2
pH 7.2	0.063 ± 0.006^a	0.0074 ± 0.0009^a	9 ± 2
pH 6.8	0.14 ± 0.03^a	0.011 ± 0.002^a	13 ± 5

[a] Rates were determined from the average of 5-9 separate experiments. Errors are expressed as the average ± 1 standard deviation.

As the pH was increased from 6.8 to 7.8 the unfolding rate for X-Lin-DnaB-N decreased from 0.14 min^{-1} to 0.042 min^{-1} and for Cz-DnaB-N decreased from 0.0105 min^{-1} to 0.0045 min^{-1}. This is in disagreement with an EX2 mechanism of exchange where deuterium exchange rates would be expected to increase an order of magnitude for each pH unit.[79,294] It is very likely that an EX1 mechanism is being observed. The decrease in transition rate for both mutants is likely to a result of the stabilization of the structure upon increasing the pH. Interestingly, the ratio of X-Lin- to Cz- rates remained the same within experimental error, indicating the effect of pH is similar for both mutants.

3.4 Conclusions

Using a quenching method, HDX conditions were determined where linear and intein cyclized forms of DnaB-N could be analyzed by ESI-MS. Under native conditions (between pH 6.8 and 7.8) it was observed that HDX occurred via an EX1 mechanism for both proteins. The cyclization of the C- and N- residues by the formation of a peptide bond resulted in a ten-fold reduction in the rate of the opening (unfolding). This ten-fold reduction correlated well with the difference in the average exchange rates of amide hydrogens determined by NMR experiments. Interestingly, linking of C-and N-termini residues did not stop the protein from unfolding, suggesting unfolding may begin from different areas of the molecules, albeit at a reduced rate.

Observation of an EX1 exchange process at neutral pH is very unusual especially for a protein that has a low contact order. It is thought that this is a result of the unusual helical structure of DnaB-N.

Chapter 4

Thermal Stability of DnaB-N Mutants – comparison of positive and negative ion ESI-MS.

4.1 Introduction

In ESI mass spectra, denatured proteins (≤ 40,000) carry a greater number of charges than native proteins.[121,122,295] This is thought to be the result of increased solvent accessibility to acidic (negative ions) or basic sites (positive ions) on proteins[296] and has been the basis for the use of ESI mass spectrometry to be used to identify unfolded and transient populations of protein conformation.[297-299] Studies such as the hydrogen deuterium exchange (HDX) experiments of Eyles *et al*. have indicated that under mildly denaturing conditions more highly charged ions are in a more unfolded conformation than lower charged ions.[101] The application of ESI-MS to identify different conformational states of proteins is complicated as preparation of these states requires a change in solvent conditions, which in itself may affect the extent of ionization during the electrospray process.[300-304]

A number of papers have examined factors that might affect the extent of ionization of proteins in ESI mass spectra. These include the type of organic solvent[300,301] and the type of cation[302 ,303] or anion used as a co-solute.[304] These studies have found that when the solvent has a higher gas phase basicity, protein ions with fewer charges are observed. The lower charges are thought to result from ion pairing which occurs at the droplet surface between acidic and basic residues of the protein and solvent/co-solvent molecules.[305-307] Other studies have shown that the number of sites on proteins which are able to become charged[72,303,308,309] can affect the charges of the ions observed in the

ESI spectra (commonly referred to as charge state distributions or CSD). Therefore, a change in protein CSD may arise from these factors and not from structural changes in the protein.

In many studies, conformational changes of proteins have been effected by decreasing the pH and unfolding the protein.[123-125,297,299,310] It is difficult under these conditions to exclude the possibility that some of the additional positive charges occur because of an effect of solvent on the ionization process rather than unfolding of the protein. As seen in chapter 3, cyclized (Cz) and linear (X-Lin) forms of the N-terminal of DnaB (DnaB-N) have been produced. The Cz- and X-Lin-DnaB-N proteins have similar tertiary structures under native conditions, yet thermodynamic studies[241] and the HDX results reported in chapter 3, indicate Cz-DnaB-N is more stable than its linear counterpart. Since Cz-DnaB-N is expected to maintain its fold, a comparison of the CSDs of this protein and its linear form under different experimental conditions will isolate effects that are a function of changing solvent and source conditions from effects relating to unfolding (increased accessibility of ionizable residues). Cz-DnaB-N is more stable to thermal denaturation than X-Lin-DnaB-N,[241] and subsequently a smaller change in the CSD is expected with an increase in temperature.

4.2 Scope of Project

In this work, the effects of different solution (pH) and ionization source conditions (desolvation temperature) on the abundance of ions of different charge or CSDs in ESI mass spectra of the proteins, X-Lin-DnaB-N and Cz-DnaB-N were compared. The latter is stabilized against unfolding by the covalent linking of the N- and C- termini.

These results were compared with the effect of reduction of disulfide bonds of α-lactalbumin on ESI mass spectra. DnaB-N and α-lactalbumin are acidic proteins and these effects were probed by acquisition of positive and negative ion ESI mass spectra. For comparison, the effects of disulfide bond reduction on the CSD of the basic protein, ribonuclease A, were also examined.

4.3 Results and Discussion

4.3.1 Thermal Denaturing of X-Lin and Cz-DnaB-N

Thermal unfolding studies of proteins assayed by mass spectrometry have been reported by a number of groups.[71,239,298,310-312] In the study by Fenn, the rate of evaporation in the ion source was controlled by varying the flow rates and temperatures of gases used to desolvate ions.[71] A shift in CSD to more abundant highly charged ions was observed when experimental conditions caused the rate of evaporation of solvent to increase. In support of the IEM (see section 1.1.6.3), Fenn commented that the observed shift in CSD was due to a decrease in the time droplets spent in the low surface area-to-charge density region, inhibiting ions with low charge states from being evaporated from droplets. What this study did not consider was that during the more vigorous desolvation conditions, the protein (cytochrome c) may have unfolded, exposing more protonation sites and resulting in a shift in CSD. Cytochrome c has been shown to unfold as a result of thermal degradation as evidenced by CD and UV-vis studies[313] and mass spectrometry.[298] It is possible that both factors are significant.

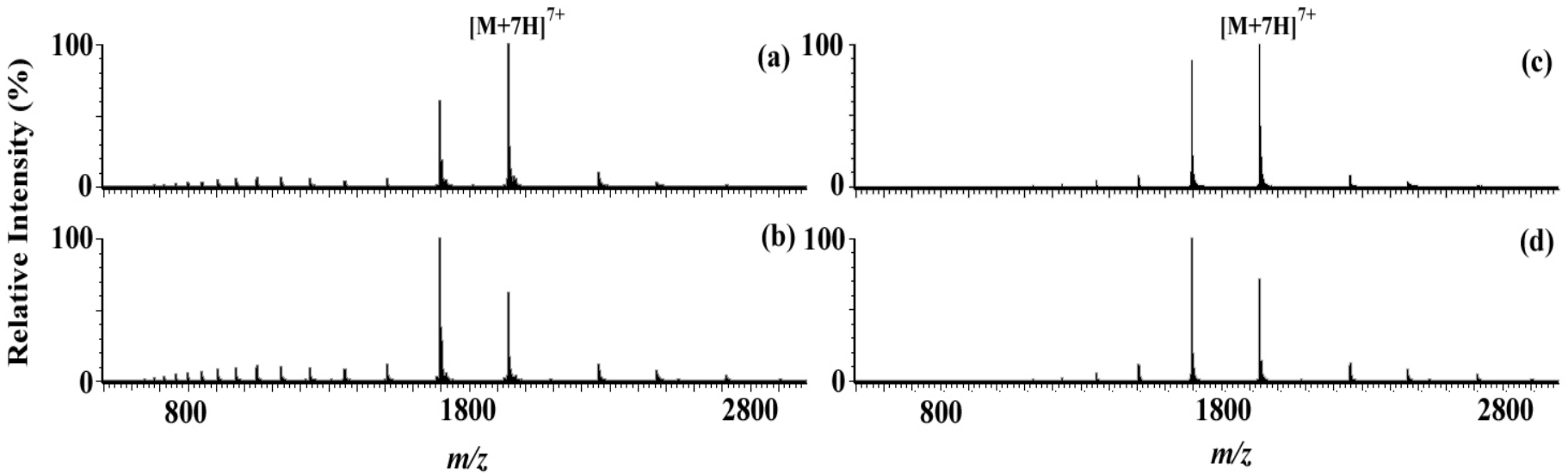


Figure 4.1. The positive ion ESI mass spectra of X-Lin- and Cz-DnaB-N (10 μM) when ionized from a solution of 10 mM NH_4OAc, pH 7.6. (a) X-Lin-DnaB-N at a desolvation temperature of 40 ^oC . (b) X-Lin-DnaB-N at a desolvation temperature of 240 ^oC (c) Cz-DnaB-N at a desolvation temperature of 40 ^oC. (d) Cz-DnaB-N at a desolvation temperature of 240 ^oC. All other instrument conditions were held constant.

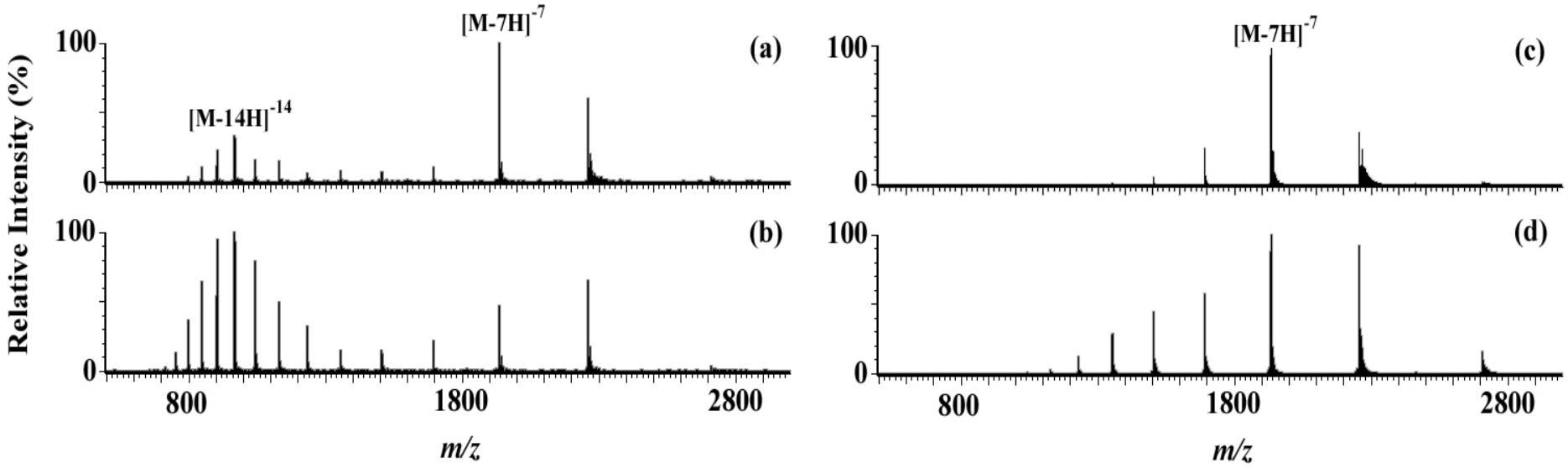


Figure 4.2. The negative ion ESI mass spectra of X-Lin- and Cz-DnaB-N (10 μM) when ionized from a solution of 10 mM NH_4OAc, pH 7.6. (a) X-Lin-DnaB-N at a desolvation temperature of 40 ^oC . (b) X-Lin-DnaB-N at a desolvation temperature of 240 ^oC (c) Cz DnaB-N at a desolvation temperature of 40 ^oC. (d) Cz-DnaB-N at a desolvation temperature of 240 ^oC. All other instrument conditions were held constant.

The melting temperatures of X-Lin and Cz-DnaB-N measured by CD were 39.8 and 53.4 ^{o}C for linear and cyclized forms, respectively.[241] In the QToF 2 mass spectrometer used in these studies, the desolvation temperature is the temperature applied to dry nitrogen which flows in the same direction as the ions formed in the electrospray source. Figure 4.1 shows the ESI positive ion mass spectra of Cz- and X-Lin-DnaB-N when the desolvation temperatures were 40 ^{o}C and 240 ^{o}C.

When X-Lin- and Cz-DnaB-N were ionized using a desolvation temperature of 40 ^{o}C (figure 4.1 (a) and (c), respectively) the spectra were very similar with the 7^{+} charge state (*m/z* 1938.0 for X-Lin and 1935.3 for Cz) the most abundant ion observed.$^{\nabla}$ The 8^{+} ion in the X-Lin spectrum (figure 4.1 (a)) was of relatively lower abundance when compared to Cz-DnaB-N (figure 4.1 (c)). Increasing the desolvation temperature to 240 ^{o}C caused only minor changes in the X-Lin and Cz-DnaB-N spectra (figures 4.1 (b) and (d), respectively) with the 8^{+} ion (*m/z* 1695.9 for X-Lin and 1693.6 for Cz) the most abundant ion in both spectra. A small increase in the abundance of ions from 9^{+} to 19^{+} was also observed in the X-Lin spectrum when the desolvation temperature was increased with the most abundant of these ions, the 13^{+} ion (*m/z* 1044.0), less than 15% of the base peak intensity.

X-Lin- and Cz-DnaB-N were analyzed under the same conditions except in negative ion mode. In the X-Lin spectrum (figure 4.2 (a)) the most abundant ion was 7^{-} (*m/z* 1936.1), however, a second less abundant distribution centred around the 14^{-} ion (*m/z* 967.5) was also present. The 14^{-} ion was approximately 30% of the base peak ion. In calorimetry studies, a percentage of the protein population was present in an unfolded

$^{\nabla}$ Here and throughout the remainder of this chapter, the $[M+7H]^{7+}$ ion is referred to as the 7^{+} ion or in negative ions the $[M-7H]^{7-}$ ion is referred to as the 7^{-} ion etc. The full description is shown in the respective figures.

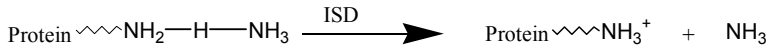
conformation,[241] consistent with this observation. A further decrease in desolvation temperature resulted in reduced abundance of the lower *m/z* series. These results were not highly reproducible, most likely because of the inability of the mass spectrometer to control desolvation temperatures when close to room temperature. When the desolvation temperature was increased to 240 ^{o}C, (figure 4.2 (b)) the ions centred around the 14^{-} ion were the most abundant in the spectrum of X-Lin-DnaB-N. A second less abundant series was still present, centred around the 8^{-} ion (*m/z* 2258.9). These negative ion results may suggest that thermal unfolding is occurring by heating of the X-Lin DnaB-N mutant, exposing more acidic residues and making them available for ionization. In contrast, this effect was not observed in the corresponding positive ion spectra (figure 4.1).

The negative ion Cz-spectra acquired using a desolvation temperature of 40 ^{o}C (figure 4.2 (c)) showed only one series centred around the 7^{-} ion (*m/z* 1933.4) at 40 ^{o}C consistent with a folded conformation. Increasing the desolvation temperature to 240 ^{o}C (figure 4.2 (d)) resulted in the appearance of more ions in this series with the 7^{-} ion remaining the most abundant. The ions observed ranged from 5^{-} to 12^{-} ions in comparison to X-Lin-DnaB-N which under the same conditions gave ions from 5^{-} to 19^{-}. Solvent accessibility is thought to be a major contributor to changes in the distribution of ions observed with more open protein structures having higher numbers of charges as steric hindrance and Coulombic repulsions will be minimised.[131,296] Using this principle the negative ion results suggest that X-Lin-DnaB-N unfolded to a greater extent than Cz-DnaB-N under these conditions, consistent with the enhanced thermal stability of the cyclized protein observed in solution studies.[241]

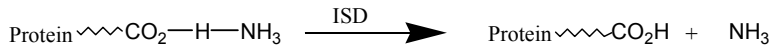
Circular dichroism studies of Williams *et al.* suggest that even at room temperature X-Lin-DnaB-N should be partially (reversibly) unfolded.[241] The negative ion ESI-MS spectra suggests that one major relative folded conformation is present. As expected, as the protein was heated (increase in desolvation temperature to 240 ^{o}C), the relative abundance of highly charged molecules increased. The measured desolvation temperature is not likely to be the exact temperature that the proteins inside an electrospray droplet experience. The lack of observable changes in ion distributions in positive ion ESI mass spectra stand in contrast to previous ESI mass spectrometry studies by Konermann and Douglas who reported that positive mode ionization could be used as a tool for identifying tertiary structural changes in proteins while results from negative ion spectra should be treated with care.[124] Their experiments used proteins with neutral or basic isoelectric points. A possible explanation for the differences observed between positive and negative ion mass spectra is discussed below. Another example, calmodulin and its complex with trifluoperazine, is discussed in chapter 5.

4.3.2 A Mechanism for the Opposite Polarity Ionization of DnaB-N

Including the linkers, the sequence of DnaB-N contains 16 basic residues (including the N-terminus which is not present in the Cz mutant and three histidine residues) and 23 acidic residues (including the C- terminus). This makes the pI of this protein acidic owing to the excess of acidic residues (sequence calculated pI = 4.67). Therefore at a pH of 7.6, as used in these experiments, the net charge on this protein is negative. Typically, ESI mass spectra of negatively charged proteins have been acquired in negative mode;[314] the acquisition of positive ion mass spectra for these proteins has been termed "wrong-way round" ionization.[306]



Scheme 4.1 (a)

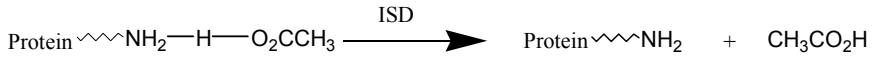


Scheme 4.1 (b)

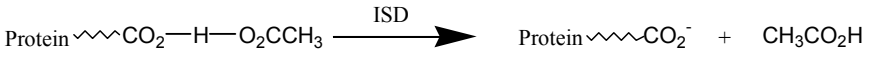
ISD = in source dissociation

In the positive ion mode the positive charge applied to the tip of the capillary causes cations which in the case of these experiments is NH_4^+ (as NH_4OAc is the buffer; some H_3O^+ can also be formed by this process however this may only be at concentrations of 1 μM as opposed to NH_4^+ which is in millimolar concentrations in solution and thus donates cations[303]), to enrich at the surface of the droplet. Some acetate anions may also be present in the droplet forming ion pairs. Conversely, in negative ion mode, acetate anions form at the droplet surface. Following desolvation of the solvent, ion pairs will be formed between available charged sites of the protein and ammonium or acetate ions as seen in schemes 4.1 and 4.2 for positive and negative ion modes, respectively. Evidence for the formation of these charged species has been observed (e.g. $[M+NH_4]^+$) using gentle collision conditions.[305-307] When accelerated through cone-skimmer regions of the mass spectrometer some in-source collisions with gas molecules will lead to dissociation of these pairs. Those species with the highest gas phase basicity will retain protons. The gas phase basicity of an alkylamine increases as the length of the chain increases. This suggests that in the tug-of-war between lysines (gas phase basicity 227 kcal/mol) and ammonia (gas phase basicity 196 kcal/mol), lysine amines will have the highest basicity and therefore retain the positive charge (scheme 4.1 (a)). For carboxylic anions, despite being acidic in solution, without the

stabilisation of anions by solvent, a highly exothermic neutralisation process occurs, removing the proton from ammonia (scheme 4.1 (b)).[315]



Scheme 4.2 (a)



Scheme 4.2 (b)

ISD = in source dissociation

In the negative ion mode (scheme 4.2) pairing between acetate and the available charged sites of the protein will again occur (schemes 4.2 (a) and (b)). As seen in positive ions, the acetate anion will not be stable in the gas phase so it will take protons from protonated basic sites of the protein. Unlike the amine derivatives, the acidities of carboxylate ions are not significantly changed by the addition of alkyl groups so charges are likely to remain where they were on the proton in solution unless affected by Coulombic repulsions.[133] For opposite polarity ionization, these situations mean that when an acidic protein is ionized from a neutral solution, in the positive ion mode, as seen for DnaB-N, the majority of the net negative charges will be neutralised, and basic residues may be able to adopt positive charges from ammonium cations allowing the polarity of the protein to be positive.

The effects these opposite polarity processes have on the ability to detect conformational changes has only been examined in a few studies.[124,316] Loo and co-workers mentioned that in the analysis of non-covalent complexes of proteins, changes in the polarity of the molecule may prevent successful detection of target non-covalent complexes.[316] Although this was referred to in the analysis of calcium binding proteins,

the same may be true for detecting conformational changes in other proteins. Konermann and Douglas have analyzed how positive and negative ionization affect the ability of ESI-MS to detect conformational changes in proteins.[124] In their analysis they examined the proteins cytochrome c, lysozyme and ubiquitin which have sequence calculated pIs of 9.52, 9.32 and 6.56, respectively. No strongly acidic proteins such as DnaB-N were examined. Konermann and Douglas commented that the changes in CSD in positive ion ESI mass spectra have been attributed to the increased accessibility of possible titration sites and the increased surface area of the unfolded state. Therefore it might be expected that these effects may influence the negative ion CSD of acidic proteins in an analogous way.[124]

Kebarle and co-workers have attempted to explain differences seen in positive and negative ion CSDs through the analysis of diamine and dicarboxylic acid systems.[132,133] By varying the length of the carbonyl chain linking the charged sites they were able to determine the effect of Coulombic repulsions on the gas phase basicities these groups. They found that provided positive charges were distanced approximately 10 Å apart, the gas phase basicity would be stronger than the basicity of ammonia leaving the diamines largely in a diprotonated state.[132] In negative ion mass spectra the distance between negative charges was required to be approximately 25 Å.[133] These observations were proposed to account for the common observation of seeing fewer charges in negative ions spectra of proteins sprayed from neutral solution as opposed to the number of positive charges.[124,317]

The question of why the positive ion ESI mass spectra of X-Lin-DnaB-N do not indicate unfolding is uncertain. A simple explanation for the inability to observe the

likely unfolding of X-Lin-DnaB-N with increasing desolvation temperature in positive ion ESI mass spectra and the ability to do this in negative ion ESI mass spectra is that the unfolding process exposes acidic residues (relative to the folded form) but does not expose more basic residues than in the folded form. That is, most of the basic residues that contribute to ESI mass spectra of the folded protein are on the surface whether the desolvation temperature is 40 ^{o}C or 240 ^{o}C. In order to test the effects of wrong-way round polarization for other acidic proteins α-lactalbumin, calmodulin (which will be looked at in chapter 5), and a basic protein ribonuclease A were also looked at using positive and negative ion ESI mass spectra.

4.3.3 Opposite Polarity Ionization of α-Lactalbumin

The milk protein α-lactalbumin is a small (14,179 Da), acidic protein containing 17 basic residues (including the N-terminus and 3 histidine residues) and 21 acidic residues (including the C-terminus) with a calculated pI of 4.80. α-Lactalbumin functions as one of the two components of lactose synthase which catalyses the final step in lactose biosynthesis[318] and has been shown to have one Ca^{2+} binding site.[319] A previous study of α-lactalbumin has shown that reduction of disulfide bonds by addition of DTT results in the production of a molten globular state which precipitated on aggregation.[240] Analysis of this protein by ESI-MS requires that unfolded forms of this protein remain in solution without precipitation. To allow the production of this form of α-lactalbumin, conditions were established to enable unfolding without precipitation.

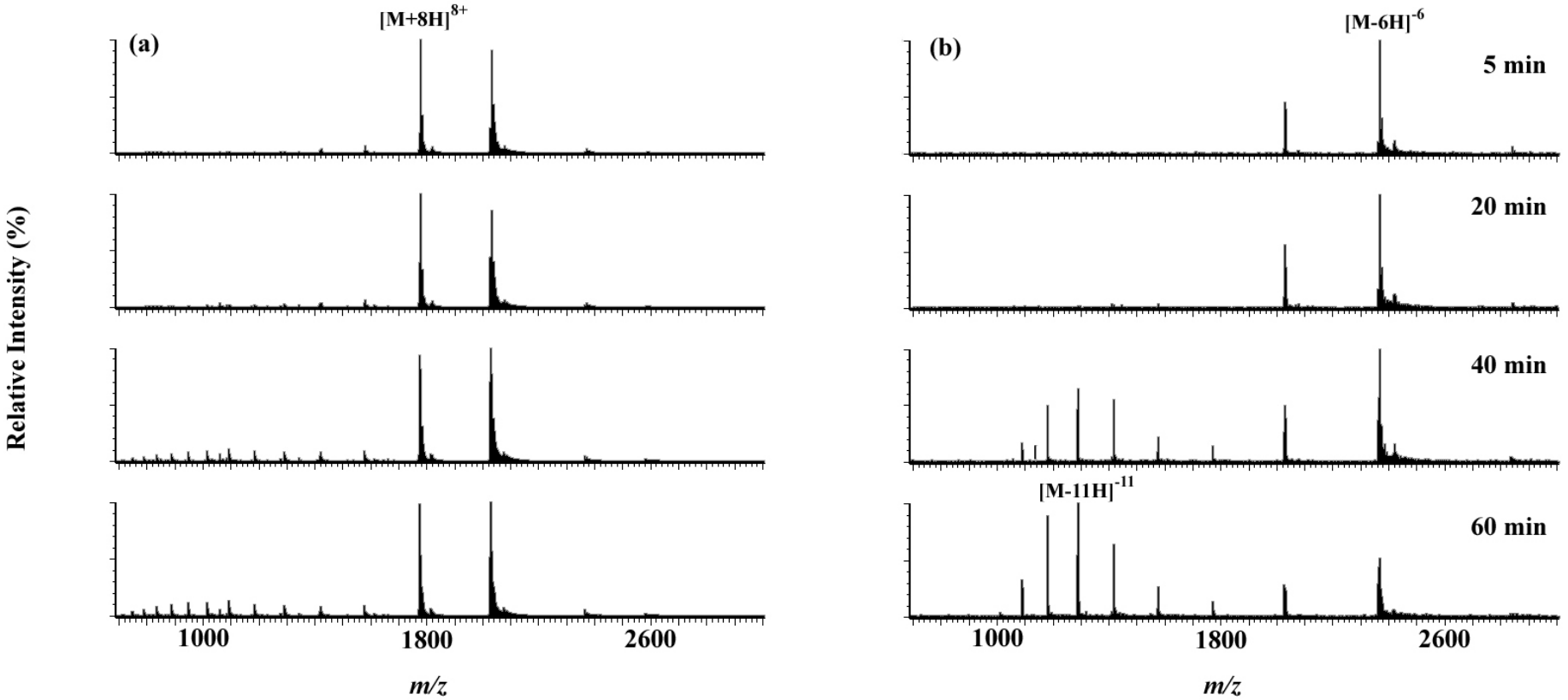


Figure 4.3 ESI mass spectra of α-lactalbumin (10 μM) in 10 mM NH_4OAc, pH 7.1, following treatment with a 560-fold excess of DTT. The effect on the charge distribution in: (a) positive and (b) negative ions is shown over a time course from 5-60 minutes.

In the analysis carried out by Carver and co-workers, a relatively high concentration of protein (140 μM) was treated with 20 mM DTT (140-fold excess over α-lactalbumin) in 50 mM phosphate buffer, pH 7.2. For ESI-MS analysis, substantially less α-lactalbumin (5 μM) can be used and NH_4OAc is required as a buffer since phosphate buffer interferes with the ESI process. These factors proved vital as the low concentration of α-lactalbumin and the presence of 10 mM NH_4OAc prevented precipitation over a range of DTT concentrations. An experiment showing the effect on ESI mass spectra of a 560-fold excess of DTT over α-lactalbumin can be seen in figure 4.3.

For both positive (figure 4.3 (a)) and negative ion (figure 4.3 (b)) spectra at an early time point (5 minutes), ions correspond largely to a species with a mass 38 units above the mass of the apoprotein. This indicates that the protein is in its native form (i.e. holoprotein), binding one calcium ion. Positive ion spectra are largely comprised of two ions the 8^+ and the 7^+ (*m/z* 1778.3 and 2032.8, respectively) with only one series observed. The most abundant ion in the negative ion spectra was the 6^- (*m/z* 2368.7), again with only one series. After 40 minutes treatment with DTT, a second distrubution centred around the 11^- ion (*m/z* 1288.5) is evident. After 60 minutes this second distribution is more abundant than that around the 6^- ion and persisted for at least 120 minutes. The positive ion spectrum did not change substantially over the same time.

Analysis of the masses of the two series of ions seen in the negative ion spectra after 40 and 60 minutes treatment with DTT revealed that the second more highly charged distrubution (lower *m/z*) corresponded to apoprotein (i.e. reduced protein that has lost Ca^{2+}), while the 7^- and 6^- ions were initially from holoprotein, with some 7^- and 6^- ions of the apoprotein observed at later time periods. Interestingly, 8^+ and 7^+ ions from the

apoprotein were observed in the positive ion spectra over time, however, a second ion series was not abundant. HDX experiments on the holo and apo forms of α-lactalbumin were used to show apo-α-lactalbumin was in a less structured conformation as judged by the extent of deuterium incorporation (see appendix figure A.1). Therefore, while there was some destabilisation of the structure over time, it did not lead to a substantial change in the ions observed in the positive ion spectra.

In a final test, the number of disulfide bonds reduced by DTT was measured by treating DTT-reduced α-lactalbumin with iodoacetamide at various time points. Iodoacetamide rapidly reacts with reduced cysteine residues allowing the number of disulfide bonds reduced to be determined from a shift in the mass of the protein (as utilised by Carver *et al*.[240]). It was found that from time points from 1 minute to 120 minutes one disulfide bond was reduced. Tryptic digestion could have been performed to identify which protein sites had been reduced, however, previous studies on α-lactalbumin have shown that reduction of the Cys6-Cys120 disulfide bond is extremely rapid (<one second).[240,320] In these studies, the loss of this disulfide bond did not cause a significant change in the tertiary structure,[240,320,321] however, after reduction of a second disulfide bond (Cys28-Cys111) the protein aggregated. In the experiments in this study, no additional disulfide bonds were reduced and no aggregates were observed either by monitoring the absorbance at 360 nm for 120 minutes and or from visual inspection over 48 hours. This may in part, be a consequence of the different conditions used in these experiments, however, mass spectrometry data and HDX results indicated that a conformational change had occurred with the reduction of one disulfide bond.

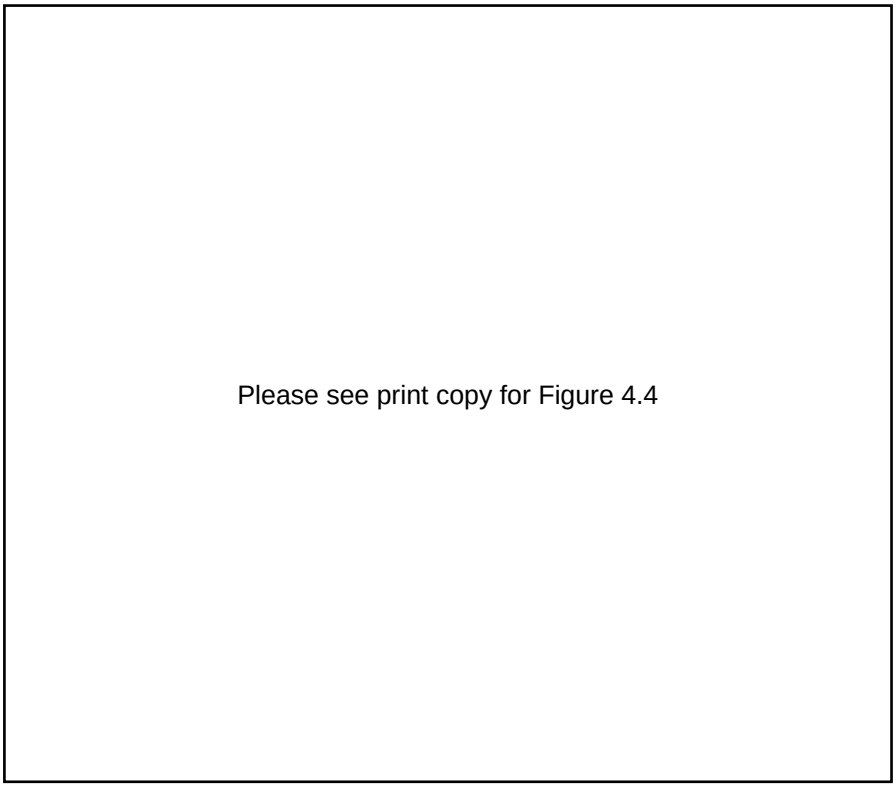


Figure 4.4 The crystal structure of α-lactalbumin (PDB 1HFX).[322] Cysteine residues are shown in red, basic resides in blue, acidic residues in green and the calcium ion (show in space fill) is shown in yellow.

Reference to the crystal structure of α-lactalbumin shown in figure 4.4 reveals that Cys6-Cys120 appears largely open to the solvent when compared to the other disulfide bonds. This would support the Cys6-Cys120 having the highest reduction rate as DTT can access it most readily. Reduction of this disulfide may allow a high degree of flexibility as relatively few tertiary structural contacts are available in this region. A number of acidic (8) and basic (6) residues are present in this region which may be able to initially stabilize the protein structure. Neutralization of these residues by formation of ion pairs with salts (NH_4^+ and CH_3COO^-) may destabilize this structure allowing for the progressive unfolding in this region.[323] This structural flexibility will continue

down to the Cys28-Cys111 disulfide bond. As a result of this increased flexibility, regions of the α-lactalbumin structure that bind calcium may also become destabilized releasing the Ca^{2+} ion.

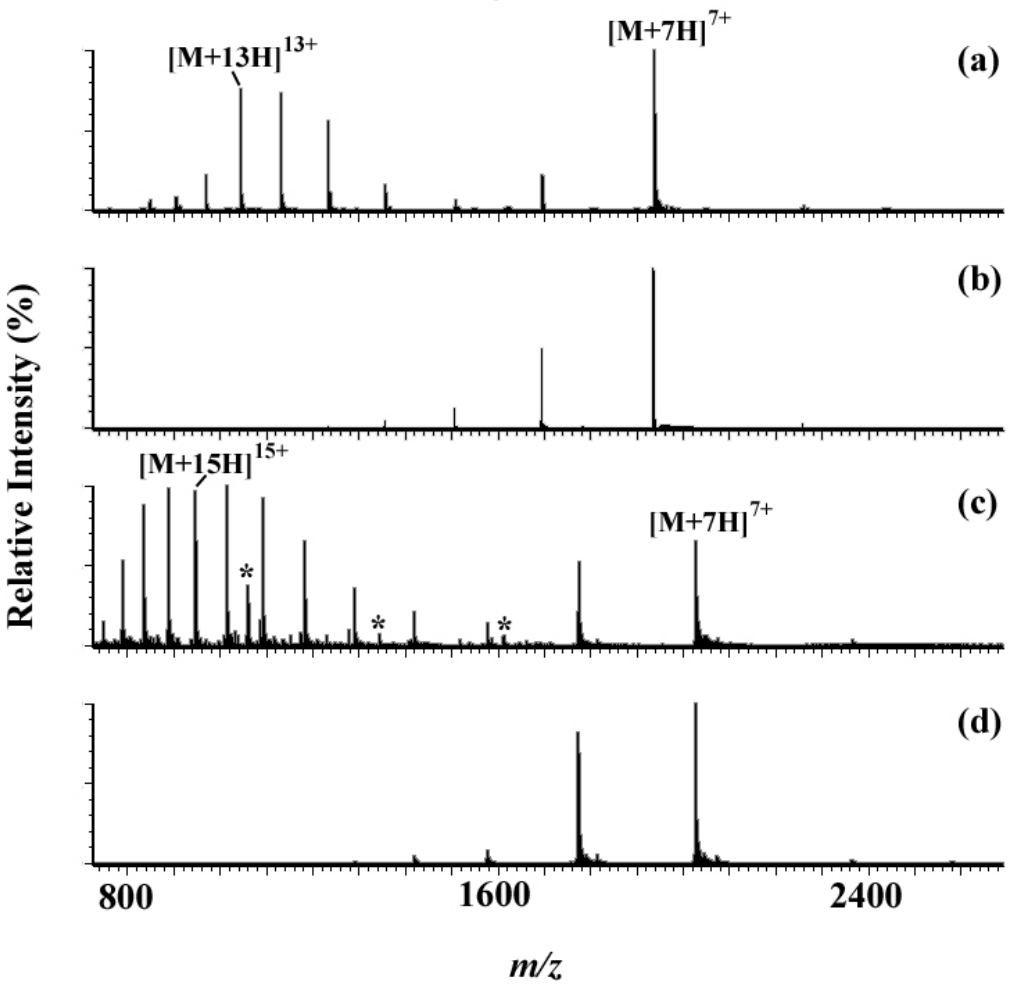


Figure 4.5 Positive ESI mass spectra of proteins at pH 2.0. (a) X-Lin-DnaB-N; (b) Cz-DnaB-N; (c) reduced α-lactalbumin, and (d) oxidised α-lactalbumin. DnaB-N spectra were acquired at a desolvation temperature of 40 °C and α-lactalbumin spectra were acquired at desolvation temperature of 120 °C. The symbol * indicates ions from an unknown species.

The loss of calcium from α-lactalbumin, the HDX incorporation rates, and the reduction of one disulfide bond suggest that a conformational change occurs under these conditions. As seen for X-Lin-DnaB-N, negative ion ESI mass spectra change in a way that is consistent with this. That is, there is a shift in the ions observed to lower *m/z*

(more highly charged ions). In contrast, no change was observed in positive ion ESI mass spectra. The maximum number of positive charges that can be observed for holo- and apo-α-lactalbumin was tested by obtaining positive ion ESI mass spectra from a solution at pH 2.0 (adjusted with formic acid). The spectra of Cz- and X-Lin-DnaB-N were also acquired under these conditions as shown in figure 4.5.

For apo-α-lactalbumin, the maximum observed charge state is the 19^+ ion (figure 4.5 (c)). This correlates closely to the number of basic residues (Lys + Arg + His + N-terminal = 17) present in the sequence of this protein. Additional protonation might occur at weakly basic sites such as Pro, Trp and Gln.[309] The spectrum of X-Lin-DnaB-N shows a maximum of 16^+ charges (figure 4.5 (a)), which correlates with the number of basic residues (Lys + Arg + His + N-terminal = 16) in this sequence. When these proteins were subjected to conditions that were thought to unfold (or partially unfold) them at neutral pH (high desolvation temperature for DnaB-N and reduction for α-lactalbumin), no changes in the positive ion ESI mass spectra were observed. One explanation for this is that at neutral pH basic amino acids are positively charged, while the acidic residues retain negative changes. Therefore, the maximum achievable net positive charge is lower. In contrast, at acidic pH, acidic residues should be largely protonated which may count for why excessive positive charges are observed under these conditions but not under neutral conditions. It must also be considered that the structures of negatively and positively charged molecules may be different since Coulombic repulsion between charges may destabilise some structures relative to others. Therefore, the ionization process itself may alter protein conformation.

Both Williams and co-workers[300,308,309,324-327] and Kebarle and co-workers[72,133,303,328] have examined the ability of proteins to adopt charges. These studies have largely involved positive ionization. In one study using native proteins, the gas phase basicites at individual sites of the protein were determined. The Rayleigh limit[329] was the major contributing factor for the maximum number of charges incorporated onto a protein in almost all cases examined.[72] An example when this was not the case was the analysis of the acidic protein pepsin (in positive ions) where the number of strong basic sites was less than the Rayleigh limit. Subsequently the maximum charge state observed was less than the Rayleigh limit. Although they did not examine denatured proteins, a similar argument may hold, although the numbers of charges will be different. As previously stated, at neutral pH both acidic and basic residues carry a proton, and may form ion pairs in native structures. Thus it could be speculated that if these pairs are still intact after unfolding or new pairs form in the unfolded conformations, the number of basic sites available to adopt charges will not change significantly. Therefore no change in the ions observed would be evident. It would then follow that in the right polarity ionization of X-Lin-DnaB-N and apo-α-lactalbumin (negative ions) this is not a problem because of the excess of acidic residues.

When proteins are restricted by cyclization or the presence of disulfide bonds as for Cz-DnaB-N (figure 4.5 (b)) and holo-α-lactalbumin (figure 4.5 (d)), a maximum of 10 positive charges were observed for Cz-DnaB-N and 11 for α-lactalbumin. This suggests that at pH 2, the proteins were not completely denatured and basic sites were not accessible for ionization or Coulombic repulsions limited the formation of some positive ions. Unrestricted forms of these proteins appear to unfold readily as seen from the large numbers of positive charges present in ESI mass spectra (figure 4.5 (a) and (c)

for DnaB-N and α-lactalbumin respectively). This suggests that for the analysis of acidic proteins, in positive ion ESI mass spectra, the polarity of the protein in solution should also be positive which can be achieved by the addition of acid. Conditions where this is not the case should be treated with extreme caution. The limitation here is of course, that the effects of acid treatment on the folds of different proteins (extent of denaturation) will vary.

4.3.4 Opposite Polarity Ionization of Ribonuclease A

Ribonuclease A is a basic, disulfide-bonded protein which functions *in vivo* as an endonuclease enzyme that specifically degrades single-stranded RNA at uracil and cytosine bases. In order to investigate further the effect of wrong-way round ionization of different conformational states of proteins, positive and negative ESI mass spectra were acquired. Unlike DnaB-N and α-lactalbumin, ribonuclease A is a basic protein containing 19 basic residues (including 4 histidines and the N- terminus) and 10 acidic residues (including the C-terminus). The sequence calculated pI is 8.64.[282] ESI mass spectra of ribonuclease A in positive and negative ion modes showed that the oxidised (folded) form of this protein was extremely stable, with no changes in the ions observed even after increasing the desolvation temperature to 240 ^{o}C, increasing voltages applied to sampling cones, or following the addition of up to 50% methanol (data not shown). Reduction of disulfide bonds was also found to be difficult when the protein was at 10 μM in 10 mM NH_4OAc irrespective of DTT concentration. In these experiments reduction of the protein occurred when 30% acetonitrile was included in reaction mixtures with a 100-fold excess of DTT (1 mM) and heating of the mixture at 45 ^{o}C. This method caused a change in the distribution of ions observed in both the positive and negative ion the ESI mass spectra of ribonuclease A as shown in figure 4.6.

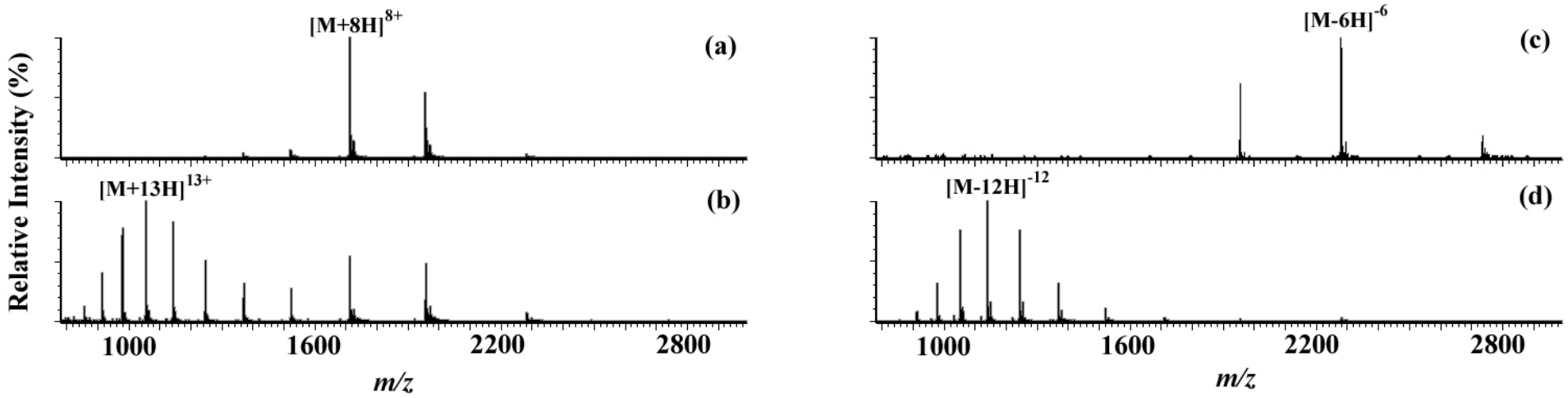


Figure 4.6 ESI mass spectra of ribonuclease A at 5 μM. (a) Positive ion spectrum produced from a 10 mM NH_4OAc solution at pH 7.1. (b) Positive ion spectrum produced from a 10 mM NH_4OAc, 1 mM DTT and 30% (v/v) CH_3CN at a of pH 7.1. (c) Negative ion spectrum produced from a 10 mM NH_4OAc solution at pH 7.1. (d) Negative ion spectrum produced from a 10 mM NH_4OAc, 1 mM DTT and 30% (v/v) CH_3CN at a of pH 7.1.

ESI mass spectra of untreated ribonuclease A (thought to be folded since disulfide bonds were intact) in positive and negative ions can be seen in figure 4.6 (a) and (c), respectively. In the positive ion ESI mass spectrum, one ion series centred around the 8^+ ion (*m/z* 1711.3) was observed. The negative ion spectrum also showed a single ion series centred around the 6^- ion (*m/z* 2279.7). The series of ions of lower charge in negative ions (which was also observed for α-lactalbumin (figure 4.3) before disulfide bond reduction) has been observed previously.[124,330] De La Mora explained this as being a consequence of the higher volatility of acetate ion when compared to ammonium ions, reducing the Rayleigh charge limit for negative ions.[135] When ribonuclease A was reduced, the positive ion spectrum (figure 4.9 (b)) shows two distributions, one series of highly charged ions centred around the 16^+ ion (*m/z* 1054.2), and the ions with the maximum abundance at the 8^+ ion (*m/z* 1956.9), present at approximately 50% of the base peak intensity (the 13^+ ion) observed in the spectrum of the unreduced protein. The negative ion spectrum (figure 4.6 (d)) is inconsistent with previous results where acidic (negative) proteins analyzed using wrong-way round polarity (positive) were not sensitive to changes in conformation. In this spectrum, distrubution was largley centred around the 12^- ion (*m/z* 1137.8), with only low abundance 6^- and 7^- ions observed. In both positive and negative ion modes, the mass of the reduced protein was 8 Da above that of the oxidised protein, corresponding to protonation of all 8 reduced cysteine residues (4 disulfide bonds).

The extent of reduction of disulfide bonds was tested by reaction with iodoacetamide (as above). The results of these experiments are shown in table A.1 of the appendix. After 30 minutes, 97% the protein had all 4 disulfides reduced which remained stable up to 60 minutes. Early time points indicate that reduction of disulfides is a progressive

process, however, these time points were not analyzed directly by MS as it was difficult to continually inject a sample a 45 ^{o}C.

Finally, to ensure a conformational change was occurring following reduction, HDX experiments were performed (see appendix figure A.2). Instead of using a quenching method as for DnaB-N and α-lactalbumin, a direct injection method was used because inefficient quenching was observed for reduced ribonuclease A, as evidenced by significant back-exchange. This suggested that a number of amide protons in the reduced state exchange very rapidly. No substantial salt adducts were observed for ribonuclease A when this method was employed in contrast to results for DnaB-N (chapter 3). In the oxidised (or native) state, a maximum of 183 sites had deuterons incorporated whereas in the reduced state, a maximum of 218 deuterons were incorporated. These results suggest that the solvent exposure of the reduced form is greater than that of the oxidised form supporting that there had been a conformational change on reduction.

The HDX results and the iodoacetamide reactions that confirm that reduction of the four disulfide bonds of ribonuclease A had occurred, support that the protein had unfolded, at least to some extent. In this case, both positive and negative ion mass spectra appear to support that this has occurred. It is therefore simplistic to suggest that wrong-way round ionization is insensitive to protein conformational changes.

Konermann and Douglas observed similar behaviour of reduced lysozyme (pI 9.32). The positive ion spectra indicated highly charged ions in the spectrum of the reduced form (10^{+} maximum abundance for the oxidised protein compared to 17^{+} maximum

abundance for the reduced/denatured protein). The negative ion mass spectra also exhibited higher charges for the reduced form although to a lesser extent (6^- maximum abundance for the oxidised protein to 9^- maximum abundance for the reduced protein). As discussed previously, the observation of fewer negative charges in comparison to positive charges has been attributed to the higher volatility of acetate anions.[75] De La Mora has attempted to explain those results in terms of the charged residue model (CRM) and the ion evaporation model (IEM). In unfolded proteins, generally the maximum observed charge state (z_{ob}) is greater then the Rayleigh limit (z_R).[75] For this to occur, De La Mora commented that the radius of the droplet containing the unfolded protein must extend to a size much larger than the radius of the native protein, a process which has been assumed to be a consequence of the more linear shape of a denatured protein. He speculated that this will allow more charge to be present in the droplet and subsequently $z_{ob} > z_R$ for a unfolded protein. If, however, the unfolded protein in the droplet has sufficient energy to be expelled via the IEM, before expansion on the droplet occurs, z_{ob} will be similar to the z_R, despite the unfolded state of the protein. Applying the same logic here, for ribonuclease A and lysozyme ($z_R = 10$ in positive and negative ions), this would suggest that in negative and positive ions species remain inside the droplet which stretches out allowing for a higher charge than z_R. In contrast, the results for X-Lin-DnaB-N and α-lactalbumin where in positive ions excess charges do not occur when the protein is denatured in a neutral solution, molecules may evaporate prior to droplet expansion resulting in the number of charges being governed by the Rayleigh limit. In negative ions this is not the case as $z_{ob} > z_R$. Based on this, it can be speculated that for some proteins, different mechanisms may dominate for ions sprayed under negative or positive ionization possibly as the result of differences in intrinsic properties or structures of the charged proteins. It may be possible to resolve

some of these issues by studing the same proteins using an ion mobility mass spectrometer,[416] where structural changes that may result from opposite polarity ionization might be detected.

4.4 Conclusion

Analysis of the thermal unfolding properties X-Lin- and Cz-DnaB-N by negative ion ESI-MS revealed information on the relative stability of these two proteins. Changes in solvent conditions were not required to effect unfolding allowing assignment of changes in the relative abundances of protein ions of different charges to be attributed to differences in the solvent accessibility of protein conformational states. Analyses in positive ions, however, did not reveal conformational changes in these proteins. This conflicts with previous ESI-MS studies,[124-126] that suggest conformational changes may be analyzed by ESI-MS. Similar observations were made for α-lactalbumin (another acidic protein) at neutral pH. Positive ion ESI mass spectra of reduced and oxidised forms of α-lactalbumin and X-Lin- and Cz-DnaB-N under acidic solution conditions (pH 2.0), distinguished between the folded and unfolded forms. These results suggested that when protein conformational changes are being analyzed, the polarity of ionization should be the same as the polarity of the protein in solution. The inability to detect conformational changes at neutral pH under opposite polarity conditions may possibly be attributed to ion pairing between charged acidic and basic residues, decreasing the availability of charging sites for denatured proteins, however as seen for ribonuclease A (a basic protein) this is not always the case. Therefore, this is an area that clearly merits further investigation.

A concern of this work is that when proteins are ionized in the opposite polarity to solution, structures and therefore functional properties may change inhibiting the ability to observe non-covalent complexes. Given the large volume of work currently using mass spectrometry as a tool to observe and characterise these complexes, chapter 5 will describe experiments designed to explore the ability of negative and positive ion ESI-MS to detect non-covalent complexes.

Chapter 5

Comparison of Positive and Negative Ionization ESI Mass Spectra of Calmodulin: Interactions with divalent metals and Antipsychotic Drugs

5.1 Introduction

The results in chapter 4 demonstrated that the extent to which ESI mass spectra of proteins are sensitive to changes in solution conditions and may differ for positive and negative ions. This may in part arise because the structure of proteins (conformation) may be affected as the polarity of ionization is reversed, but there may be additional factors involved. This chapter explores these factors further by focusing on the ability of ESI-MS to detect non-covalent complexes, using the well-characterized complexes of calmodulin with calcium and antipsychotic drugs as examples.

Calmodulin (CaM), has been studied extensively both by solution based methods[331-333] and by mass spectrometry.[180,214,279,314,316,334-336] CaM coordinates calcium flux within cells trigging a range of processes including DNA synthesis and smooth muscle contraction.[337] There is a high degree of sequence homology across species with the human and bovine sequences being identical. Figure 5.1 shows the structures of different forms of calmodulin and complexes.

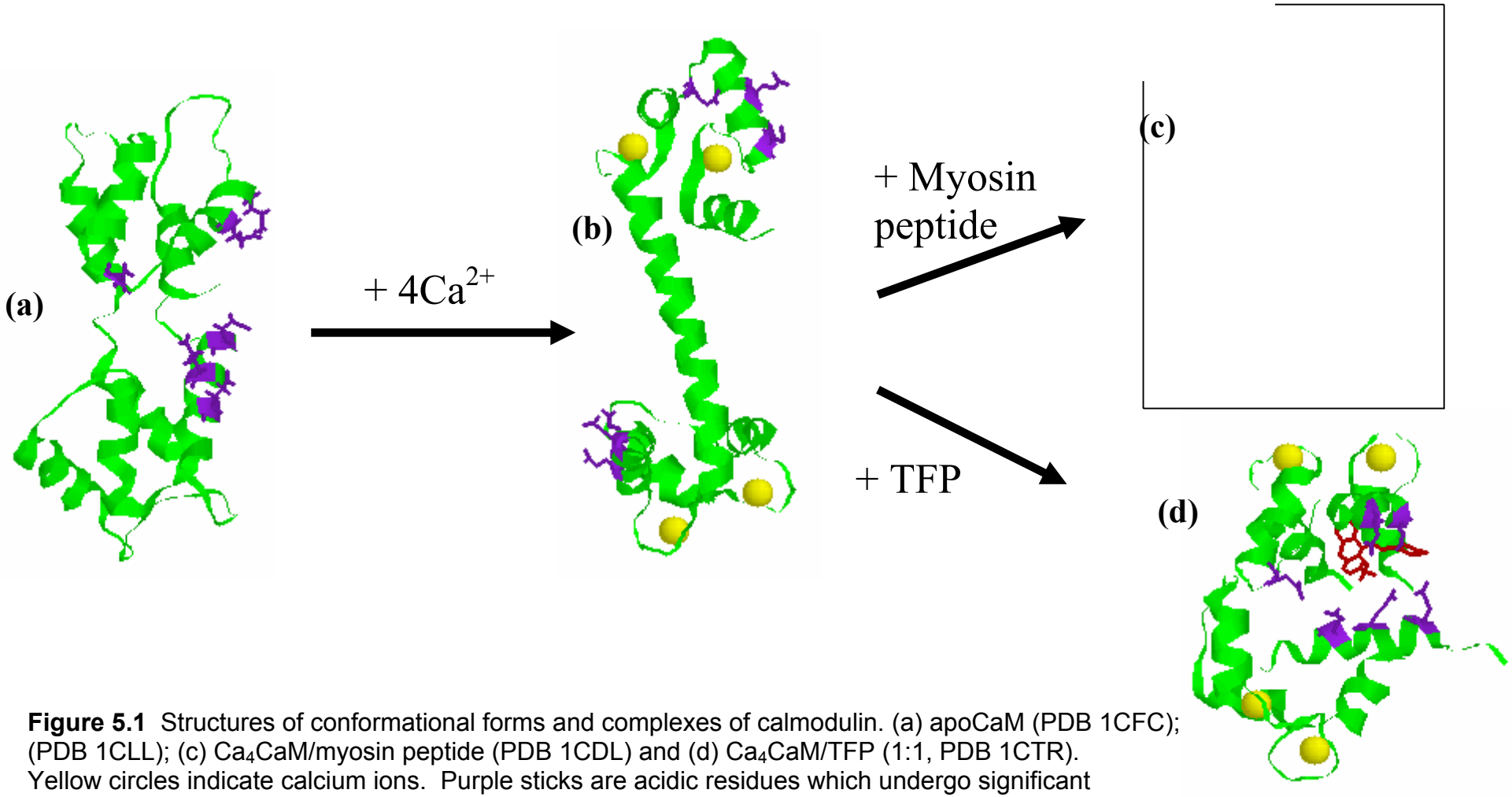


Figure 5.1 Structures of conformational forms and complexes of calmodulin. (a) apoCaM (PDB 1CFC); (PDB 1CLL); (c) Ca_4CaM/myosin peptide (PDB 1CDL) and (d) Ca_4CaM/TFP (1:1, PDB 1CTR). Yellow circles indicate calcium ions. Purple sticks are acidic residues which undergo significant movement on binding TFP or the myosin peptide. Myosin and TFP are shown in red in (c) and (d), respectively. Structure (a) is from NMR experiments, structures (b), (c) and (d) are from X-diffraction experiments.

The crystal structure of calcium saturated-calmodulin (Ca_4CaM) showing four calcium ions binding to separate EF-hand motifs in the CaM structure is shown in figure 5.1 (b).[338] These Ca^{2+} ions coordinate to 3 Asp, 1 Glu and 2 Thr residues causing a conformational change in the CaM structure exposing hydrophobic residues to a greater extent than when compared to calcium-free CaM (apoCaM; figure 5.1 (a)).[333,338]

Exposure of this hydrophobic patch of CaM allows interactions with a range of different proteins (or peptides) by forming non-covalent complexes.[339] A number of the peptide sequences which interact with Ca_4CaM have been investigated. These tend to be long sequences of hydrophobic and positively charged residues, arranged in an α-helix.[340] NMR[341] and X-ray crystallographic structures[342] for these complexes have been determined indicating the peptides are engulfed inside the hydrophobic cavity of Ca_4CaM. The C- and N- terminal EF-hand regions of Ca_4CaM move towards each other as they wrap around the peptide (see figure 5.1(c)).

Please see print copy for Figure 5.2

Figure 5.2 Structures of: (a) trifluoperazine (TFP), and (b) imipramine (IPA).

A number of the actions of calmodulin can be inhibited by binding to antipsychotic drugs such as trifluoperazine (TFP, figure 5.2 (a)) and imipramine (IPA, figure 5.2 (b)).[337] The crystal structures show TFP binds to hydrophobic patches in the EF-hand motifs of the C- and N-terminal domains (see figure 5.1 (d)).[343,344] The resulting collapse in the Ca_4CaM structure inhibits the actions of calmodulin. The conformational changes of CaM that occur on the binding of Ca^{2+} and as a result of the binding of antipsychotic drugs, makes CaM a very useful protein for studying the effect on ESI mass spectra of conformational changes of proteins that occur under native conditions (i.e. physiological pH range).

A number of ESI-MS studies have investigated aspects of the non-covalent complexes formed with CaM. Hu *et al.* used negative ion ESI-MS to confirm the observation that four Ca^{2+} ions bind to calmodulin.[314] Species containing 0-4 calcium molecules were observed. Using positive ion ESI-MS, additional auxiliary binding sites for calcium and other divalent metals such as Mg^{2+} have also been identified, which have been thought to provide a regulatory effect on the functions of CaM.[214] Veenstra *et al.* have used negative ion ESI-MS to compare spectra of apoCaM and CaM on addition of calcium, and found ions with lower charges in the spectra of Ca_4CaM which is consistent with a conformational change that decreased solvent accessibility to negatively charged residues.[180]

A number of studies have also been carried out examining the interactions between calmodulin and various peptides.[334,336,345-347] Spectra acquired in both positive and negative ion modes have indicated that these complexes can be detected regardless of the polarity of ionization with studies commonly highlighting the importance of calcium activation to allow interactions with peptides to occur. No significant differences were observed on binding of the peptides in the different ionization modes.

5.2 Scope of this Chapter

In an attempt to better understand the effect opposite polarity ionization has on the ability of ESI-MS to detect non-covalent complexes and indicate conformational changes in protein structure, this section analyses interactions between CaM and calcium and the binding of the antipsychotic drugs TFP and IPA. A comparison of positive and negative ion ESI mass spectra of these CaM complexes under native conditions was carried out and changes observed in ESI mass spectra are correlated with structural changes observed in NMR and crystallography studies. Hydrogen deuterium exchange (HDX) experiments were used to confirm that conformational changes occurred in solution under the experimental conditions. Finally, the ability of ESI-MS to maintain a "memory" of solution phase conformations was evaluated by performing in-source collisional activation of the Ca_4CaM-TFP complex.

5.3 Results and Discussion

5.3.1 Analysis of apoCaM

CaM is an acidic protein with an isoelectric point (pI) of ~4, so that at physiological pH CaM is negatively charged. Consequently, this protein has been analyzed previously in negative ion mode in an attempt to reduce the likelihood of observing non-specific metal ion attachment (e.g. Na^+).[314] Under these conditions, peaks in negative ion ESI mass spectra of apoCaM showed fewer adducts than those in positive ion ESI mass spectra of the same solution.[314]

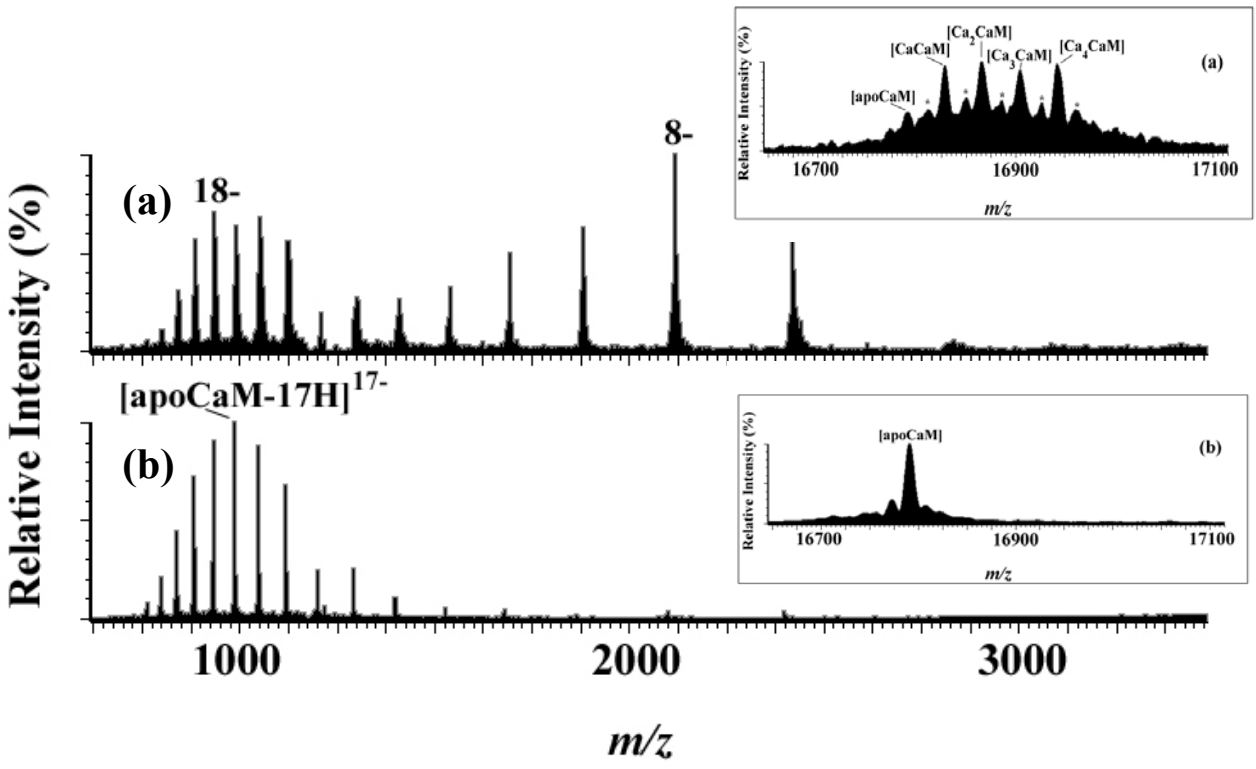


Figure 5.3 Negative ion ESI mass spectra of commercial CaM (a) prior to treatment with EDTA and (b) following treatment with EDTA. Insets show transformed spectra of the respective samples.

Figure 5.3 shows negative ion ESI mass spectra of commercial, "calcium-free" CaM before and after treatment with EDTA. The spectrum of the untreated sample showed ions of significant abundance from protein molecules with one to four bound calcium ions (figure 5.3 (a)). Sodium adducts were also be observed as denoted by the asterix (see inset of figure 5.3 (a)). ESI mass spectra were acquired with a desolvation temperature of 240 ^{o}C in an attempt remove non-specific adducts. This did not have a significant effect on the quality of the spectra. Therefore, a procedure was developed to remove adducts (and residual Ca^{2+}) by treatment with EDTA (as outlined in chapter 2). The spectrum of commercial calmodulin prepared using the EDTA treatment procedure is shown in figure 5.3 (b). The procedure successfully removed calcium and sodium adducts, as shown by the transformed spectrum (see inset of figure 5.3 (b)). The mass of the major species was found to be 16790.3 Da in agreement with the calculated mass of apoCaM.[314] A significant change in the charges of the ions was observed on the removal of calcium with the 17^{-} ion *m/z* 986.7 being most prominent under these conditions.[Ψ] In the presence of small amounts of calcium (figure 5.3(a)) at least two series appear to be present centred around the 18^{-} and 8^{-} ions, respectively. The higher abundance of ions in the series centred around the 8^{-} ion is consistent with the binding of calcium causing a conformational change of the protein.

[Ψ] As noted in the previous chapter the full descriptions for each ion are given on each figure, except where individual species were insufficiently resolved eg. figure 5.3 (a). In some cases the full description is required in the text for the purpose of discussion.

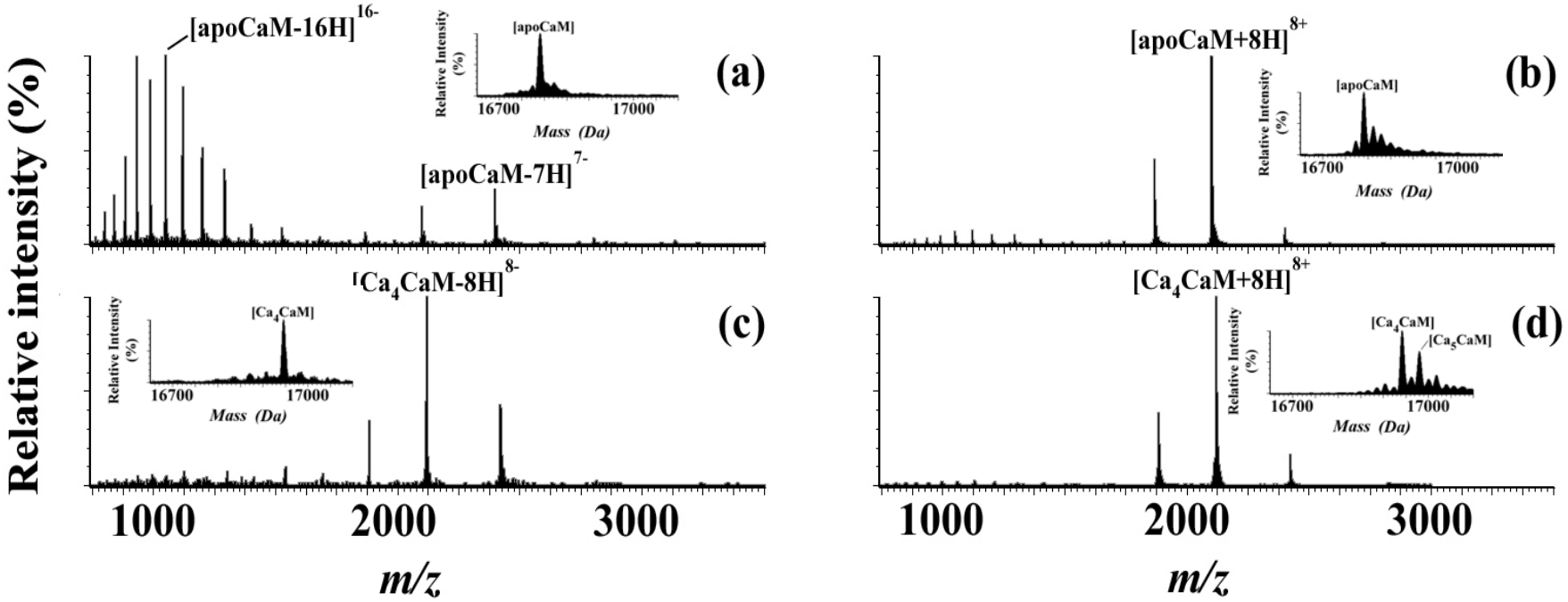


Figure 5.4 Negative and positive ion ESI mass spectra of apoCaM and Ca_4CaM (10 μM, pH 7.6). (a) Negative ion ESI mass spectrum of apoCaM. (b) Positive ion ESI mass spectrum of apoCaM. (c) Negative ion ESI mass spectrum of Ca_4CaM. (d) Positive ion ESI mass spectrum of Ca_4CaM. Inserts show the spectra transformed to a mass scale. In all experiments the desolvation temperature was set at 40 ^{o}C.

ApoCaM prepared as above was analyzed in the negative ion mode except this time with a desolvation temperature of 40 ^{o}C. Figure 5.4 shows this spectrum (figure 5.4 (a)) together with the corresponding positive ion spectrum (figure 5.4 (b)). The negative ion spectrum shows the presence of at least two series with the most abundant ions from each envelope at *m/z* 1048.4 and 2397.7, that corresponded to $[apoCaM-16H]^{16-}$, and $[apoCaM-7H]^{7-}$, respectively. The maximum charge observed was the $[apoCaM-21H]^{21-}$ ion. The presence of these series suggests that there were at least two protein conformations.[121] When compared to the spectrum when the desolvation temperature was increased to 240 °C (figure 5.3 (b)), ions from the $[apoCaM-8H]^{8-}$ and $[apoCaM-7H]^{7-}$ ions were of significantly lower abundance. This is consistent with an increase in desolvation temperature causing unfolding of proteins because the $[apoCaM-8H]^{8-}$ and $[apoCaM-7H]^{7-}$ ions arise from a more compact form of the protein that was sensitive to thermal unfolding.

The positive ion spectrum of apoCaM at a desolvation temperature of 40 ^{o}C is seen in figure 5.4 (b). In this spectrum, three ions centred at *m/z* 2099.8 ($[apoCaM+8H]^{8+}$) predominated, with ions of very low abundance from a charge envelope centred at *m/z* 1120.4 ($[apoCaM+15H]^{15+}$). Increasing the desolvation temperature to 240 ^{o}C resulted in a small increase in abundances of the envelope centred around the 15^{+} ion (data not shown).

The relative abundances of highly charged ions of apoCaM in positive ion ESI mass spectra are significantly lower than negative ion spectra under these solution conditions (pH 7.6) and is independent of desolvation temperatures under the conditions of these experiments. Applying the CRM[67,75] (see section 1.1.6.3), the maximum charge carried

by calmodulin should the same as the Rayleigh limit which has been previously calculated to be 10.[78] The positive ion spectrum of apoCaM (max charge 9^+) is therefore reasonably consistent with this calculation, however, the negative ion spectrum of the same calmodulin solution with a maximum of 21^- charges under these conditions is in far excess of the Rayleigh limit. Deviations from the Rayleigh limit are expected for denatured proteins which may suggest negative ion spectra of apoCaM are a more sensitive indicator of unstructured apoCaM than positive ion ESI mass spectra. CaM has 16 basic sites (six Lys, eight Arg, one His and the N-terminus). The positive ion ESI mass spectrum of apoCaM (figure 5.4 (b)) suggests that some of these sites were buried in the apoCaM structure and inaccessible to ionization because 9^+ is the maximum charge. When positive ion ESI mass spectra of apoCaM were acquired from a solution containing formic acid (pH 3.0; data not shown), the charge envelope centred at 8^+ remained, but a charge envelope centred at 14^+ (up to a maximum of 19^+, but at low abundance) was also present, consistent with the maximum number of basic sites on the protein, and suggesting that some of the protein molecules were denatured under these conditions, rendering more basic sites available for ionization. From these experiments at pH 7.6 it is not possible to establish if the observed maximum charge state is related to the Rayleigh limit and/or a reduction in the accessibility of ionizable sites in the folded form. These experiments were repeated on at least 5 occasions yielding the same results.

5.3.2 Analysis of Calcium-Binding Properties of CaM Using Positive and Negative Ion ESI Mass Spectra.

Figure 5.4 also shows the positive and negative ion spectra of Ca-bound calmodulin. The binding of divalent calcium ions displaces two protons to maintain charge balance,[314] resulting in a mass increase of 38 for each calcium bound. The negative ion ESI mass spectrum of Ca_4CaM (figure 5.4 (c)) gives a mass for Ca_4CaM of 16,942.3 Da, corresponding to a mass increase of 38 for each calcium ion as reported previously.[180,314] The ion series observed in the negative ion Ca_4CaM spectrum (figure 5.4 (c)) were significantly different to that apoCaM (figure 5.4(a)) which only had series with one charge envelope centred around the $[Ca_4CaM\text{-}8H]^{8-}$ ion (*m/z* 2117.2). The Ca_4CaM spectrum contains only one species of CaM bearing four calcium ions in contrast to previous studies where species containing 0-4 calcium ions could be detected,[180,314] making this one of the cleanest spectra reported to date. In preliminary experiments where smaller excesses of Ca^{2+} were added, CaM molecules with 0-4 calcium ion bound were observed in ESI mass spectra.

In the study of Loo and co-workers, ions were observed in two series, one comprising the 15^-, 14^-, and 13^- ions and corresponding to CaM molecules binding 0-4 calcium ions and the second comprising 10^-, 9^- and 8^- ions corresponding to Ca_4CaM only.[316] The authors interpreted these results as a conformational change occurring as a result of binding four calcium ions. This is very reasonable and X-ray structures show that there is a conformational change when CaM binds four calcium ions (see figure 5.1 (a) and (b)). Although a comparison of the apoCaM (figure 5.1 (a)) and Ca_4CaM (figure 5.1 (b)) structures[344,348] indicates binding of Ca^{2+} resulting in a more open structure, HDX experiments have shown that the amide backbone of apoCaM has a higher solvent accessibility than Ca_4CaM. This is consistent with greater flexibility of apoCaM in

solution when compared to Ca_4CaM.[279] The changes in negative ion mass spectra on the binding of Ca^{2+} to apoCaM can be interpreted in either of two ways. First, if the binding of calcium renders ionizable groups inaccessible to the surface as the result of a conformational change, a shift in the charge of the ions observed will occur. This has been explained by Hu *et al.*[316] and Veenstra *et al.*[180] as consistent with the results presented here. Alternatively, if calcium ion binding exposes more ionizable groups, a shift in mass spectrum to lower *m/z* ions (greater number of charges) will be observed.

As discussed above, ESI mass spectra suggest the binding of calcium ions causes the loss of two protons from the protein structure. A second explanation is therefore that the binding of each divalent calcium ion shields two negative charges that may otherwise have contributed to the negative ion apoCaM spectrum. These results (figures 5.4 (a) and figure 5.4 (b)) show the ion with the maximum intensity shifts from the 16^- ion to the 8^- ion on comparing the apoCaM and Ca_4CaM forms. In the absence of corroborating structural data, using the ESI-MS results alone, this shift in ion series cannot unequivocally be assigned as an effect of a conformational change. If calcium ions bind to sites which were not contributing to peaks in mass spectra of apoCaM, these sites would become deprotoned, balancing the charges of the calcium ions and the degree of charging on the protein would be dictated by the solvent accessibility and Coulombic repulsion of charges within the Ca_4CaM structure. Under this circumstance changes in the ions observed would correspond to changes in protein structure while in the former case they would not.

Veenstra *et al*. attempted to separate the effects of charge "shielding" and conformational changes that shift negative ion ESI mass spectra to higher *m/z* on

addition of metals to proteins by investigating the interaction between apoCaM and magnesium. Magnesium binds weakly to apoCaM at sites distinct from calcium and does not cause a conformational change.[214,349] Previously, Veenstra *et al*. reported no changes in the distribution of ions observed in positive (or negative) ion ESI mass spectra of CaM when between 1 and 6 magnesium ions were bound to the protein. Using these results the authors concluded that the changes observed in the observed ion abundances when calcium were a consequence of the conformational change of the protein.

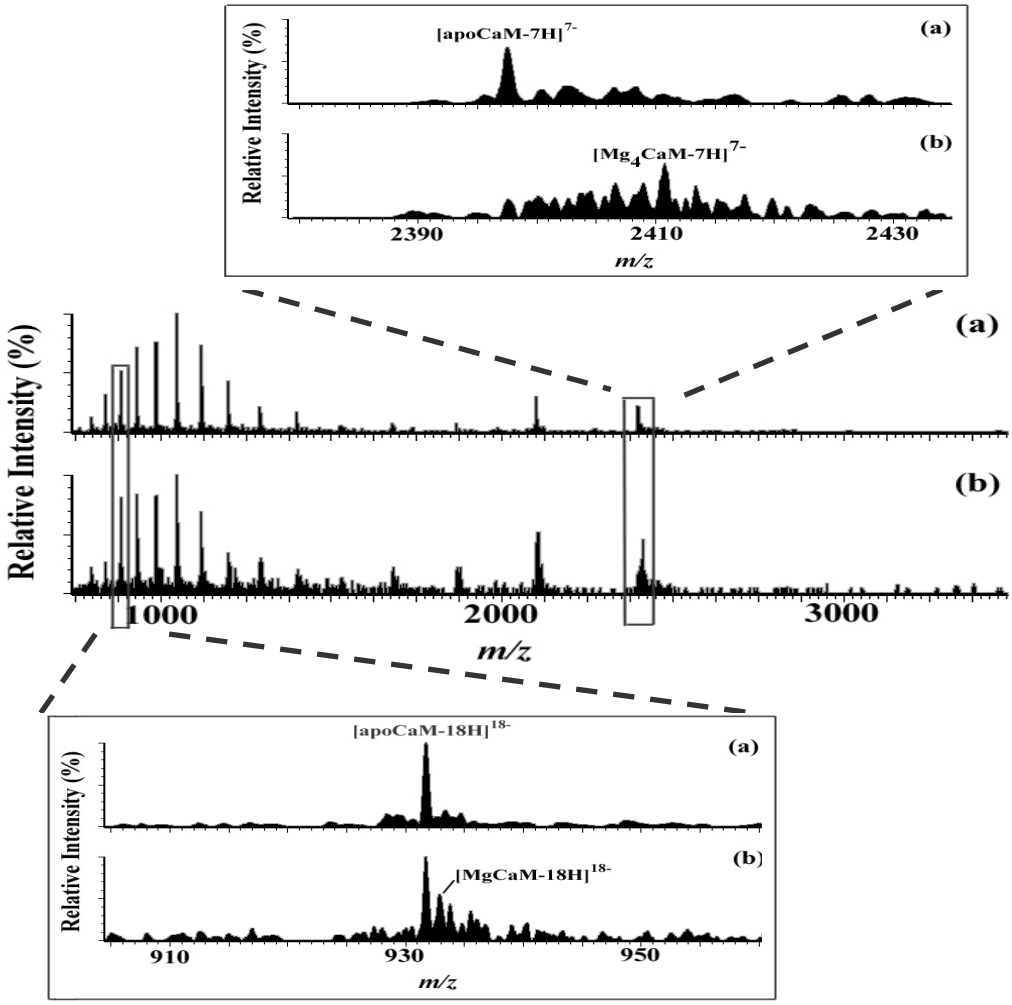


Figure 5.5 Negative ion ESI mass spectra of mixtures of apoCaM and Mg^{2+}. (a) 1:0.5, apoCaM: Mg^{2+}; (b) 1:10, apoCaM: Mg^{2+}. Regions of the spectra showing 7- and 18- ions are expanded.

In the present work, Mg^{2+} was added to apoCaM solutions in the ratios Mg^{2+}: apoCaM of 0.5:1 and 10:1. At higher concentrations of Mg^{2+}, the quality of the spectrum was poor. Figure 5.5 (a) shows a negative ion ESI mass spectrum of apoCaM: Mg^{2+} (1:0.5). This spectrum is similar to that of apoCaM (figure 5.4 (a)). Figure 5.5 (b) shows the negative ion ESI mass spectrum after addition of a 10-fold excess of Mg^{2+}. For each spectrum, the detail around the 18^- and 7^- ions is also shown. The spectra show that there was little, if any magnesium associated with calmodulin that gave rise to ions at lower values of *m/z* (higher numbers of charges). In contrast, calmodulin molecules with substantial amounts of bound Mg^{2+} gave ions only at higher *m/z* (fewer numbers of charges). These observations support that Mg^{2+} shields negative charges which contribute to negative ion MS of apoCaM. The conclusions from this work are therefore different from Veenstra *et al.* and suggest care should be taken when evaluating a change in protein structure from changes in distrubution of observed charges when metals bind.

On closer inspection of the spectra of Veenstra *et al.* however, it appears that they do not differ significantly from the spectra in figure 5.5. That is, the ions at low *m/z* (higher number of charges) appears to be predominantly from apoCaM. Therefore, the Veenstra *et al.* data also support that a shift in the observed ions occurs when Mg^{2+} is bound to CaM in the absence of a conformational change. It should be noted that in neither the work of Vennstra *et al.* or the current work, the lack of a conformational change on addition of Mg^{2+} was not independently verified.

On further consideration, it is highly questionable whether addition of a metal that binds at different (perhaps non-specific) sites can provide any information. To be useful in

discriminating between charge shielding and conformational change, the metal would have to bind at the same site as calcium but not cause a conformational change. It is unlikely that such an experiment will be possible.

The spectrum of Ca_4CaM in the positive ion mode (figure 5.4 (d)), is very similar to the apoCaM spectrum (figure 5.4 (b)). The $[Ca_4CaM+8H]^{8+}$ (*m/z* 2118.9) ion is most abundant and is flanked by the $[Ca_4CaM+9H]^{9+}$ (*m/z* 1883.6) and $[Ca_4CaM+7H]^{7+}$ (*m/z* 2421.5) ions which comprise the remainder of the charge series. As seen for other acidic proteins X-Lin and Lin-DnaB-N and α-lactalbumin (chapter 4), again it appears that positive ion mass spectra are not a reliable indicator of conformational changes of acidic proteins. Other factors including the number of strong basic sites and whether a conformational change affects the degree of solvent accessibility to acidic and basic sites may also be important factors that determine whether mass spectrometry can detect conformational changes of proteins.

5.3.3 Ca_4CaM-TFP/IPA Interactions

The binding of TFP to Ca_4CaM results in a profound conformational change collapsing the CaM structure as N- and C-terminal EF-hand motifs move towards each other (figure 5.1 (d)). It might be expected from the results described above that negative ion ESI-MS should be able to detect the conformational change as the Ca_4CaM-TFP complex is formed. Titration of TFP into the Ca_4CaM complex only resulted in the detection of this complex in the positive ion mode (discussed further below). No binding of TFP to Ca_4CaM was observed in negative ion ESI mass spectra as seen in figure 5.6.

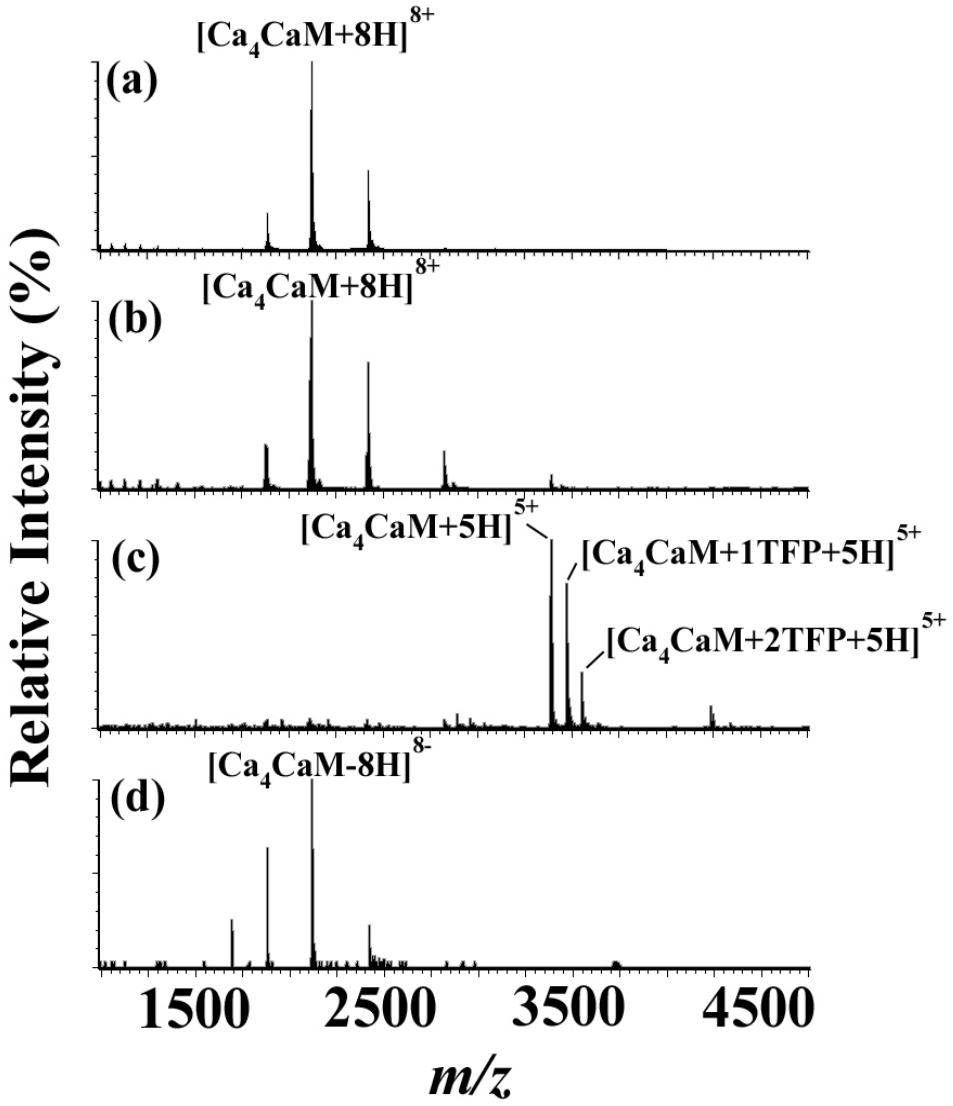


Figure 5.6 Comparison of ESI mass spectra of Ca_4CaM and mixtures with trifluoperazine or imipramine. (a) Positive ion ESI mass spectrum of Ca_4CaM. (b) Positive ion ESI mass spectrum of Ca_4CaM+IPA (1:5). (c) Positive ion ESI mass spectrum of Ca_4CaM+TFP (1:5). (d) Negative ion ESI mass spectrum of Ca_4CaM+TFP (1:5). The cone voltage in all these experiments was 35 V.

Figure 5.6 shows a comparison between the positive ion ESI spectrum of Ca_4CaM alone (figure 5.6 (a)), and that with TFP added in a 5 fold excess over Ca_4CaM (figure 5.6 (c)). The addition of TFP resulted in a shift from the 8^+ charge state to the 5^+ charge state. Complexes of one TFP molecule, ($[Ca_4CaM+1TFP+5H]^{5+}$, *m/z* 3470.9) and two TFP molecules ($Ca_4CaM+2TFP+5H]^{5+}$, *m/z* 3552.3) with Ca_4CaM in addition to unbound Ca_4CaM ($[Ca_4CaM+5H]^{5+}$, (*m/z* 3389.5) were observed under the

experimental conditions (cone voltage 35 V). Ratios up to 1:10 Ca_4CaM:TFP were attempted however, no additional binding was observed. Ions of low abundances corresponding to Ca_4CaM with the same numbers of TFP molecules were also observed for 6^+ and 4^+ charge states in figure 5.6 (c); these intensities were, however, were below 10% of the base peak intensity. The interaction of Ca_4CaM with multiple TFP molecules has previously been observed in crystallography studies.[343] TFP has a pK′a of 8.1, therefore at pH 7.6, it would be expected that ~50% of the TFP molecules would carry a positive charge. The binding of TFP can therefore only add positive charges, as a result, the observed change in CSD from 8^+ to 5^+ can more likely be accounted for by different degrees of solvent accessibility caused by a conformational change rather than the binding of TFP shielding positive charges.

Another antipsychotic drug, IPA, was added to Ca_4CaM (figure 5.6 (b)). Unlike TFP, no complex between Ca_4CaM and IPA could be observed despite solutions with ratios of Ca_4CaM to IPA up to 1:10, and irrespective of source conditions. This suggests IPA binds less tightly than TFP, which is supported by other studies where inhibition of calmodulin-induced activation of phophodiesterase, was more effective by TFP (IC_{50} = 28 μM) than IPA (IC_{50} = 125 μM).[337] Addition of IPA did cause a slight shift in the distribution of ions of Ca_4CaM, with a small increase in the intensities of the $[Ca_4CaM+7H]^{7+}$ (*m/z* 2421.4), $[Ca_4CaM+6H]^{6+}$ (*m/z* 2824.8) and $[Ca_4CaM+5H]^{5+}$ (*m/z* 3389.5) ions. The change in distribution of ions was not as significicant as observed for TFP (figure 5.6 (c)), however, it does suggest that some change in structure occurred when Ca_4CaM has been in contact with IPA, and possibly that IPA is bound in solution but is dissociated from Ca_4CaM in the gas phase.

To confirm the occurrence of conformational changes under the experimental conditions used in this work Ca_4CaM, Ca_4CaM/TFP and Ca_4CaM/IPA complexes were tested for solvent accessibility by HDX as judged by mass spectrometry. In these experiments complexes were prepared at the same ratios as calcium, TFP and IPA to CaM as for figure 5.6 (Ca_4CaM concentration = 100 μM), before incubation with deuterium oxide in a solution 10 mM NH_4OAc pH 7.6 (1 in 10 dilution, protein : D_2O solution) for various periods of time and quenching by acidification (as described in chapter 2). The results of these experiments are shown in figure 5.7.

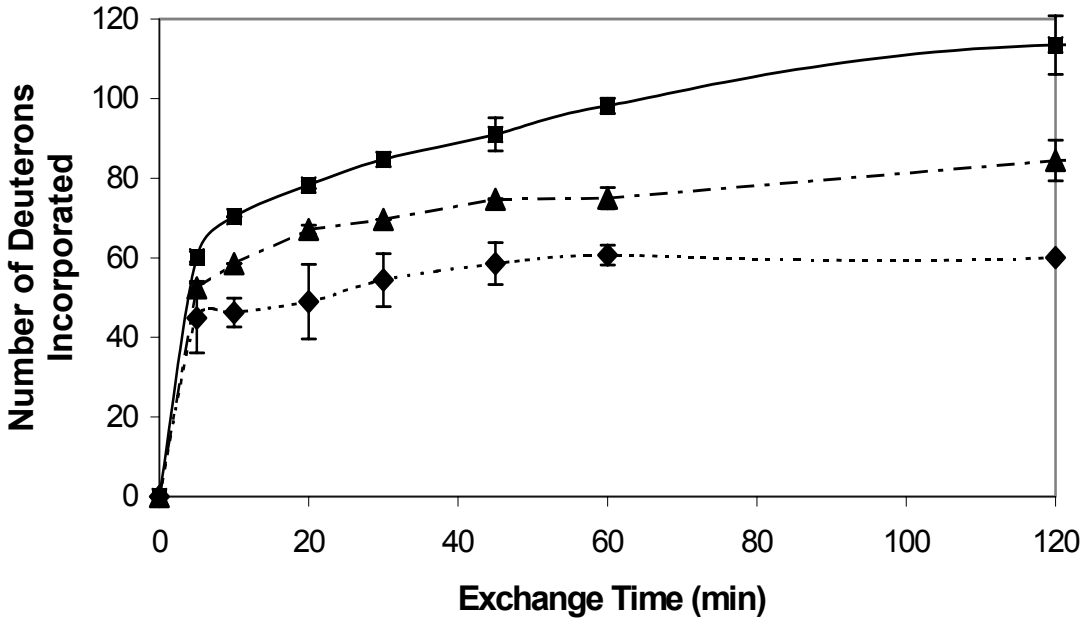


Figure 5.7 Extent of HDX with time for apoCaM, Ca_4CaM/IPA and Ca_4CaM/TFP. ■ apoCaM. ▲ Ca_4CaM/IPA. ◆ Ca_4CaM/TFP. The error bars are derived from 3 separate experiments. Experiments were performed by diluting 100 μM CaM complexes into a D_2O solution containing 10 mM NH_4OAc (pH 7.6) to give a concentration of 10 μM CaM. HDX was then quenched at time points between 5 and 120 minutes using the quenching method described in chapter 2.

After 120 minutes incubation with the D_2O solution, a maximum of 108, 60, and 84 amide protons had exchanged for Ca_4CaM, Ca_4CaM/TFP and Ca_4CaM/IPA complexes, respectively. This suggests that the relative solvent accessibility of amide protons is

Ca_4CaM > Ca_4CaM/IPA > Ca_4CaM/TFP supporting that complexes formed with TFP and IPA, but that the structures of the complexes were different, with different numbers (60 and 84 respectively) of accessible amide protons.

Figure 5.6 (d) shows a negative ion spectrum of the same Ca_4CaM/TFP solution used to obtain the positive ion spectrum in figure 5.6 (c). The negative ion spectrum does not indicate the presence of any complex between Ca_4CaM and TFP, and the charge envelope is still centred around the $[Ca_4CaM\text{-}8H]^{8-}$ (*m/z* 2117.2) ion as observed in the spectrum of Ca_4CaM (figure 5.4 (c)). No change in this spectrum was observed regardless of the source conditions (cone voltage, desolvation temperature) used. A possible explanation for these results is suggested by reference to the crystal structures of Ca_4CaM (figure 5.1 ((b)) and Ca_4CaM/TFP (figure 5.1 (d)).[344,348] The binding of TFP causes a number of acidic residues (seen in purple as sticks) to move towards each other. In the Ca_4CaM/TFP structure, Glu residues 114 and 14 move within 6.5 Å of each other, Glu123 and Glu11, 7.5 Å and Glu127 and Glu7, 6.8 Å. If these residues contribute to the negative ion ESI mass spectrum of Ca_4CaM; in the complex with TFP they may have been brought sufficiently close to one another to favor protonation of some of these residues, lowering the response factor for negative ion ESI mass spectra. Kebarle and coworkers have shown that there is a lower tolerance for proximity of negative charges than positive charges in the gas phase.[133] Alternatively, the negatively charged Ca_4CaM/TFP complex may be less stable (leading to gas phase dissociation) than the positively charged complex.

Interestingly, complexes of Ca_4CaM with peptides have been observed in a negative ion ESI-MS study.[345] This study analyzed the interaction between CaM and the peptide

melittin. Presently there is no high-resolution structure for this complex, however, a crystal structure between Ca_4CaM and a similar peptide from myosin has been determined (figure 5.1 (c)).[342] RasMol (version 2.6) was used to analyze the same glutamic acid residues discussed for the Ca_4CaM/TFP complex and revealed that these residues do not come as close together in this structure. Residues Glu114 and Glu14 are now separated by 8.5 Å, Glu123 and Glu11, 16.1 Å, and Glu127 and Glu7, 18.2 Å. Dissociation constants measured in solution for Ca_4CaM/TFP (1-8 μM[350-352]) and Ca_4CaM/melittin (3 nM[353]) indicate the binding of melittin is significantly stronger. The dissociation constants, however, have no bearing on the extent and site of electrospray ionization of a protein since the protein conformation (e.g. proximities of negative and positive charges) may be quite different.

5.3.4 MS/MS Studies of the Ca_4CaM/TFP Complex: Possible Memory Effects

HDX experiments (figure 5.7) confirmed that conformational changes occurred, limiting solvent accessibility to amide protons as Ca_4CaM bound TFP. Negative ion ESI spectra did not show ions from the complex consistent with dissociation of the complex in the gas phase and/or difficulty in forming negative ions of the complex. Although the Ca_4CaM/TFP complex is maintained (at least to some extent) in the gas phase as assessed by positive ion ESI spectra, it might be expected that dissociation of the complex in the early stages of the ionization process prior to complete evaporation of solvent may result in the Ca_4CaM formerly in a complex with TFP reverting back to a spectrum similar to that of Ca_4CaM as seen in figure 5.4 (c). Figure 5.8 shows the effect of increasing cone voltage on the Ca_4CaM/TFP complex.

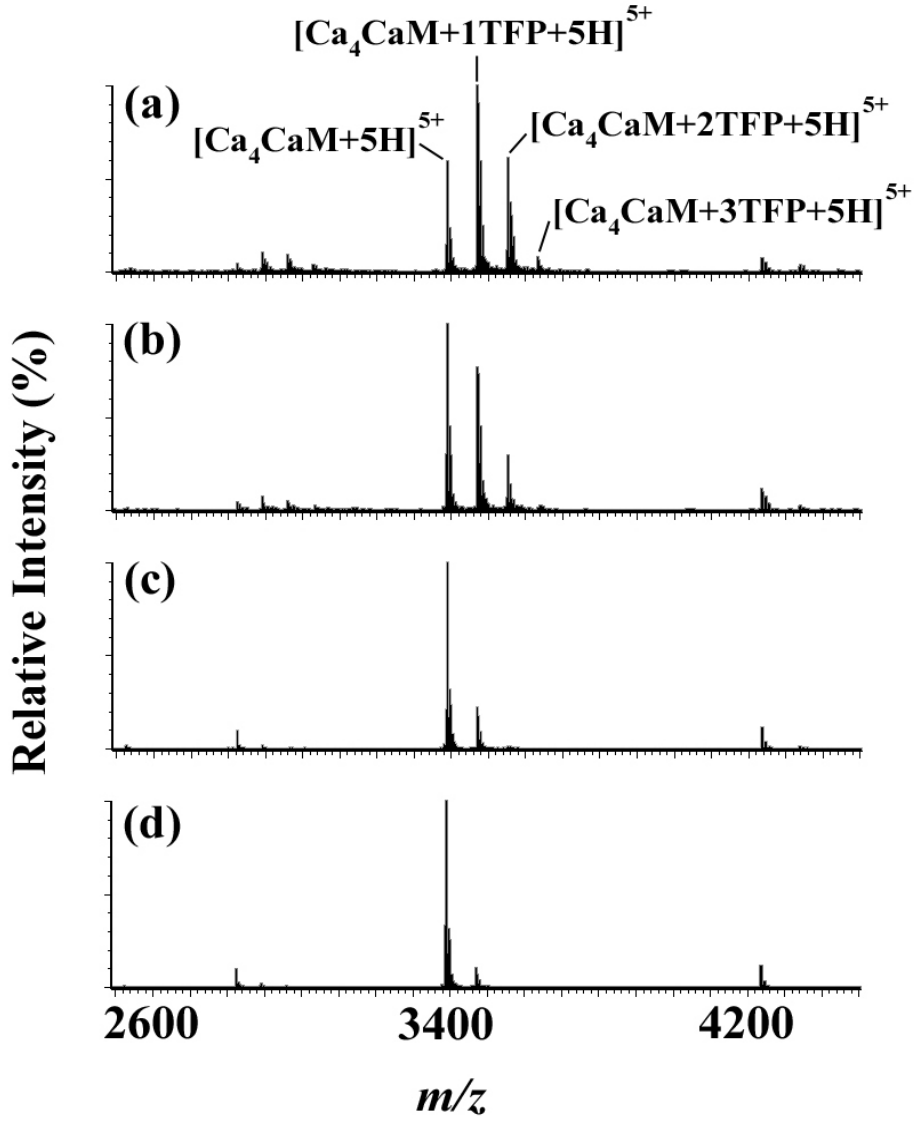


Figure 5.8 Effect of cone voltage on positive ion ESI mass spectra of Ca_4CaM/TFP complexes. (a) 30 V (b) 35 V (c) 40 V (d) 50 V

At 30 V (figure 5.8 (a)) the $[Ca_4CaM+1TFP+5H]^{5+}$ ion (*m/z* 3470.9) was the most abundant ion and ions from two ($[Ca_4CaM+2TFP+5H]^{5+}$, *m/z* 3552.3) and three ($[Ca_4CaM+3TFP+5H]^{5+}$, *m/z* 3633.6) TFP molecules bound were also observed. As the cone voltage was increased successively to 50 V (figure 5.8 (d)), the ions from 1:1, 1:2 and 1:3 complexes all diminished leaving only a species from Ca_4CaM which still had a charge envelope centred around the 5^+ ion. This series did not differ from when

TFP was bound to Ca_4CaM and indicates that in positive ion there may be a "memory" of the TFP bound conformation which was present in solution. This could not be tested with the instrumentation available in the laboratory.

The "memory effect" is thought to occur when there is insufficient time during ionization/desolvation to allow a conformational change to occur. Clemmer and co-workers used an ion mobility mass spectrometer to monitor structural transitions of cytochrome c and ubiquitin following trapping of ions in an ion trap over 10 to 200 ms and 10 ms to 30 seconds, respectively, for each protein.[354,355] In the case of cytochrome c, lower charged ions corresponding to more folded conformations did not appear to unfold for 30 ms, however after this time intensities from unfolded conformations of cytochrome c increased.[354] While the flight time through the Q-ToF instrument is in the low millisecond time frame which may be sufficent to allow a conformational change, if a conformational change occurs after evaporation of solvent, it would not be possible for a protein to pick up new charges except, perhaps, from the dissociating ligand. In the case of TFP, this is only one possible positive charge.

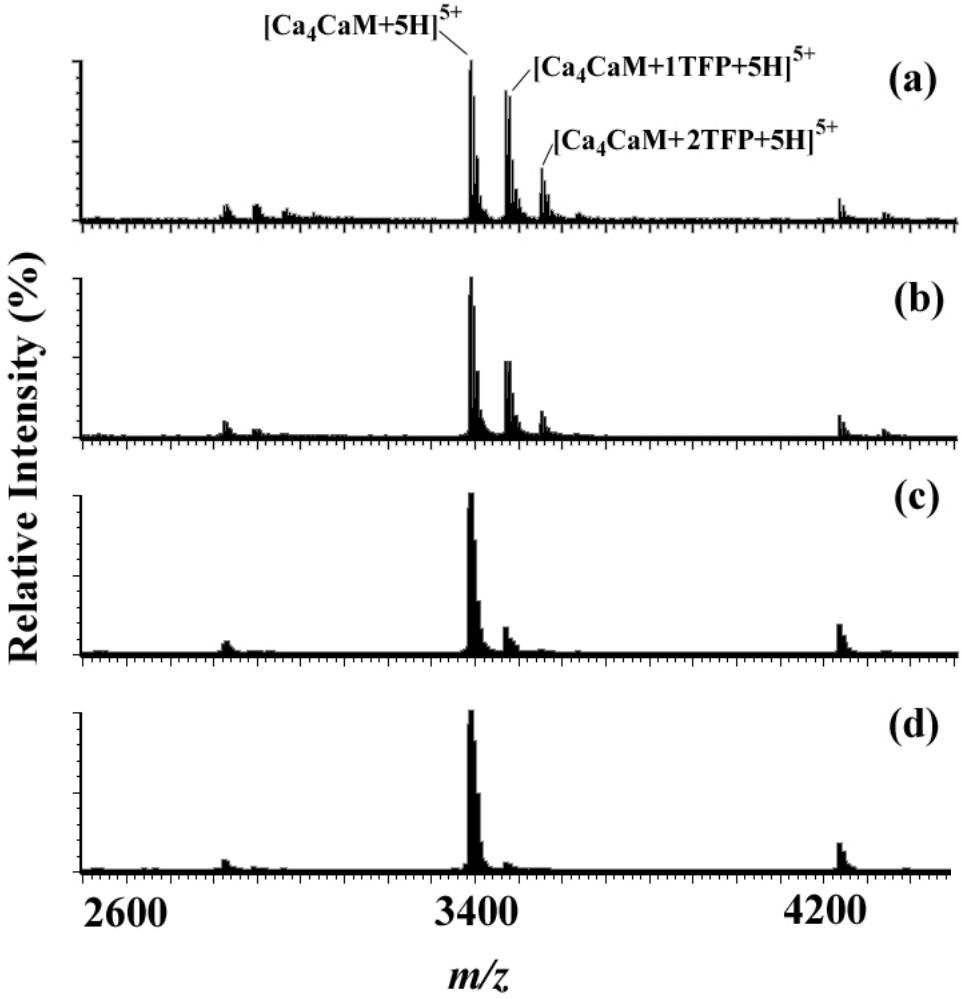


Figure 5.9 Effect of collision cell voltage on positive ion ESI mass spectra of Ca_4CaM/TFP complexes. (a) 4 V (b) 10 V (c) 15 V (d) 25 V. Cone voltage was held constant at 35 V.

The stability of the Ca4CaM/TFP complex was tested by carrying out collision-induced dissociation of the Ca_4CaM/TFP complex in the collision cell of the Q-ToF mass spectrometer. Figure 5.9 shows the effect of increasing the collision potential from 4 eV (figure 5.9 (a)) to 25 eV (figure 5.9 (d)) while maintaining the cone at 35 V. As seen in figure 5.8, a loss of the Ca_4CaM/TFP complex is observed, however, no change in the observed ion series occurs following dissociation as expected since there is no proton source other than TFP at this stage of analysis. To determine if structural changes occur

as a result of dissociation either ion mobility[416] or HDX experiments in the gas phase[417] would need to be carried out.

5.4 Conclusion

Calmodulin is a useful protein for observing the effect(s) opposite polarity ionization has on: (1) the ability to detect by ESI-MS non-covalent complexes and, (2) identifying structural changes from changes in observed ion series. From the interaction of calcium with CaM complexes, four Ca^{2+} ions were observed to bind in both positive and negative ion ESI mass spectra. This indicated that the polarity of ionization did not substantially affect the interactions between CaM and Ca^{2+} ions. In correlation with DnaB-N and α-lactalbumin studies, analysing the protein using ESI-MS in the same polarity as the net charge in solution allowed changes in ions observed (from 16^- to 8^-) to be detected on the addition of calcium. In positive ion ESI-MS, regardless of the addition of calcium, the charge envelope did not shift, with the 8^+ ion the most abundant. This is most likely because there were few accessible sites for protonation or because the Rayleigh limit had been reached for those structural forms. However, when looking at metal binding in negative mode, interpretation of changes observed in ion series need to be carefully examined as charge neutralisation may also cause changes in the ions detected rather than changes in solvent accessibility brought about as a result of conformational change.

The analysis of Ca_4CaM interactions with TFP was not possible using negative ion ESI-MS. X-ray crystallography studies indicate the structure of Ca_4CaM undergoes a profound conformational change when it interacts with TFP causing its structure to become relatively compact. The relative low tolerance for the proximity of negative

charges may have precluded formation of a negatively charged complex, alternatively, the negatively charged complex may have been unstable in the gas phase. It should be noted that care should be taken when comparing X-ray structures (solid phase with water present) and ESI-MS (gas phase) data. In positive ion ESI-MS the 5^+ ion was now the most abundant ion in the spectrum of Ca_4CaM once TFP was bound (c.f. 8^+ for Ca_4CaM). Complexes containing 1-3 TFP molecules were observed, consistent with the occurrence of a conformational change decreasing solvent accessibility to positively charged sties. Following dissociation of this complex using higher cone or collision voltages, no change in the ion series occurred. To establish if there was a memory of solution structures ion mobility and/or gas phase HDX experiments will be required.

Chapter 6
DnaB-DnaC Interactions

6.1 Introduction

Replication of DNA requires a functionally conserved set of components. The analogous protein components from different organisms often have no sequence similarity, but have similar mechanisms of replicating DNA. The functions involved in replication include: initiation at origins of replication, DNA synthesis at replication forks, and termination of replication. In addition cooperative interactions between multiple protein-protein and protein-DNA complexes are an integral part of the preparation for DNA synthesis.[356] Elucidation of the mechanisms of replication may require purification of an intact complex (referred to as the replisome) to allow aspects of its function to be easily probed. Presently, the intact replisome has not been purified and structural studies and a variety of assays that have provided details on the associations, functions and mechanisms of actions of subassemblies of the replisome have been made possible by over-expression and purification of the various components.

6.1.1 The *E. coli* replisome

The ability of bacteria to replicate rapidly and the relative ease of manipulating the expression of proteins has enabled various components of the replisomes of *Escherichia coli*, *Bacillus subtilis* and bacteriophages T4 and T7 to be studied extensively. The *E. coli* replisome is perhaps the most studied of these,[356-358] with an assembly of up to 30 or so different proteins identified as being involved in replication. (See figure 6.1)

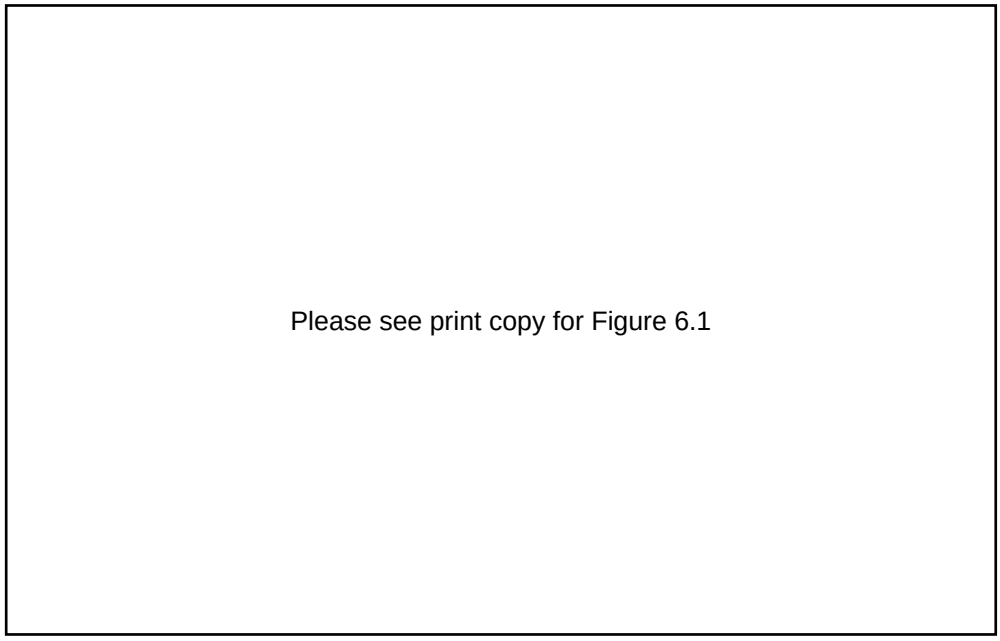


Figure 6.1 Components of the E. coli replisome. (a) The replicase with DNA polymerase III encircled and (b) the DNA polymerase III holoenzyme. Taken from Schaeffer *et al*..[356]

Replication begins from a specific region of the bacterial DNA known as the origin of replication (oriC). This is a sequence of double-stranded (ds) DNA that must be separated into a fork so that replisome components can be loaded onto single-stranded (ss) DNA. Binding of the protein, DnaA, and the cofactors Hu and/or IHF (histone-like proteins that stabilise DNA) to five copies of a specific A-T rich base pair sequences in oriC causes a bulge in the DNA.[359,360] This allows the primary *E. coli* helicase, DnaB, to be loaded onto ssDNA. Further details of this protein will be discussed below, however, the primary function of DnaB is to unwind dsDNA. Once loaded onto ssDNA, a process requiring the actions of its loading partner, DnaC (discussed below), DnaB can begin to unwind duplex DNA following priming by DnaG, the *E. coli* primase (RNA polymerase). DnaG functions to place RNA primers on the ssDNA.[361]

The cooperative actions of these components allows additional components of the replisome to be loaded onto ssDNA at the replication fork.

Additional proteins which have been identified to play important roles in replication include components of the DNA polymerase III holoenzyme (figure 6.1 (b)). DNA polymerase III is present as asymmetric dimer of the polymerase active core (α, ε, θ-subunits of DNA polymerase III),[358,362,363] the DnaX hexameric complex (δ', γ, τ_2, δ, ψ, and χ) and the clamp (β_2; see figure 6.1).[357,364] DNA polymerase III functions to polymerise new DNA strands by continuous synthesis of a leading strand and discontinuous synthesis on the lagging strand (Okazaki fragments).[365] The core contains the α-subunit, a 130 kDa protein with a conserved set of carboxylates and other polar residues which interact with divalent ions catalysing polymerisation of DNA strands.[366] The α-subunit associates with the ε-subunit, a 3′-5′ exonuclease enzyme that performs proofreading functions, and the θ-subunit which may function to stabilize ε.[213]

Separate polymerase cores are held together by a heptamer of proteins called the DnaX complex which contains the δ', γ, τ_2, δ, ψ, and χ subunits (figure 6.1).[357] The two τ-subunits allows interaction with separate polymerase α-subunits.[367,368] This is believed to be an important interaction as it allows for communication between leading and lagging strand replication.[367-369] The DnaX complex is also referred to as the clamp loader. The clamp (dimers of β-subunits, β_2) can be seen in figure 6.1 (b) below the core α, ε and θ polymerase complex which encircles dsDNA at the primer terminus, but does not contact the DNA.[364,370] The β_2 clamp marks primed sites on leading and lagging DNA sequences for tethering of polymerases to the DNA template and helps recombination of newly synthesized strands. The α-θ-ε polymerase core, the DnaX

complex and clamp all interact with each other and other components of the replisome to regulate DNA replication.

6.1.2 Helicases

Helicases are enzymes that generate ssDNA or RNA devoid of any secondary structure and play important roles in replication, repair and recombination.[371-375] Generally, they are classified according to their direction of unwinding (5′-3′ or 3′-5′), DNA or RNA, primary sequence, quaternary or oligomeric structure and the presence of a ring-shape. Analysis of proteins that display helicase activities has lead to the classification of three major superfamilies, SF1, SF2 and SF3, as well as several separate families including hexameric, replicative, primase-related DNA helicases (DnaB-like family) and helicases that show high sequence similarity with non-helicase ATPases.[376,377]

The DnaB-like family shares five conserved motifs and assemble into a hexameric ring with a central channel and unwind dsDNA in a 5′-3′ direction. High resolution structures of fragments of the bacteriophage T7 gene 4 primase-helicase protein (T7gp4)[378] and the RepA protein encoded by the RSF1010 plasmid[379] indicate similarities to the RecA protein core containing Walker A and Walker B motifs which are involved in MgATP/MgADP binding and hydrolysis. The ability of helicases to bind and hydrolyse nucleotides like ATP appears to be a common feature of helicases and provides the energy for DNA/RNA unwinding activities.[375,380]

6.1.3 DnaB

Presently, there is no high-resolution X-ray crystal structure of DnaB available. Sedimentation (analytical ultracentrifuge studies) experiments have indicated that in the presence of magnesium, DnaB exists as a hexamer ($(DnaB)_6$) of 52 kDa subunits over a

protein concentration range of $\sim 10^{-7}$ to 10^{-5} M (hexamer).[381] Using fluorescent purine nucleotide analogues, three high affinity and three low affinity nucleoside binding sites were observed at 10 ^{o}C.[382,383] Further, cryoelectron micrograph (cryoEM) images have indicated that the $(DnaB)_6$ structure appears to exist in either C_3 and C_6 symmetry depending on the experimental conditions (see figure 6.2).[384-387]

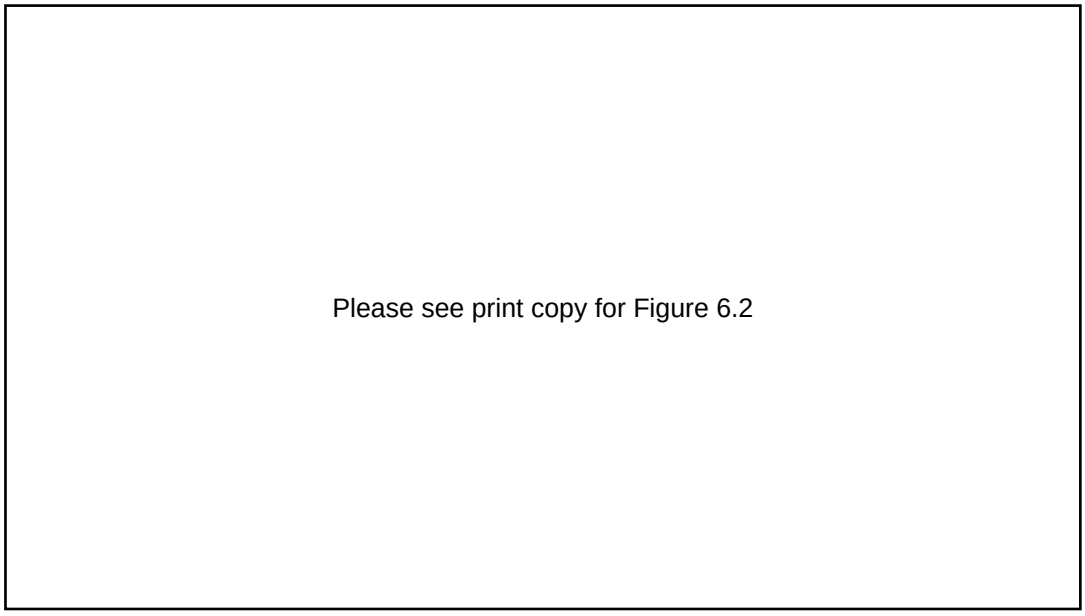


Figure 6.2 Cryoelectron microscopy images of DnaB assemblies at pH 8.1, 7.6, 7.2 and 6.5. Top row indicates conditions where C_6 symmetry was observed. Bottom row indicates conditions which allowed identification of C_3 symmetries. At a pH ≥ 7.6 DnaB was exclusively in C_3 symmetry. At pH 7.2 and 6.5 50% of molecules were in C_3 symmetry and 50% in C_6 symmetry. Adapted from Donate *et al*.[387]

A study by Donate *et al.* has indicated that pH changes can reversibly alter the quaternary structure allowing transitions from C_3 to C_6 symmetry to occur as shown in figure 6.2.[387] At pH ≥ 7.6, DnaB was exclusively in C_3 symmetry and from pH 6.5-7.2 a 50% mix of C_3 and C_6 symmetries were observed. Similar symmetry arrangements have also been observed for the *B. subtilis* bacteriophage SPP1 replicative helicase G40P[388] and the RepA helicase encoded from plasmid RSF1010,[389] suggesting this

transition may be common and may have implications in the mechanism of dsDNA unwinding.

DnaB has been shown to interact with ssDNA, dsDNA, and a number of proteins including the *E. coli* primase DnaG, the initiation protein, DnaA, the τ subunit of the clamp loader complex and the DnaB loading partner, DnaC.[356] DnaC is a 28 kDa protein that loads DnaB onto ssDNA at oriC. CryoEM images indicated that DnaC interacts with the $(DnaB)_6$ in a 6:1 ratio.[386,390] Bárcenta *et al.* indicated that DnaC appears to dimerize as it interacts with the $(DnaB)_6$ complex locking this complex into a C_3 architecture. This interaction also closes the inner channel of DnaB sufficiently to inhibit ssDNA entering.[390] DnaC is a member of the AAA+ family of ATPases (ATPases associated with a variety of cellular activities). The hydrolysis of ATP and interaction with ssDNA appear to play an important role in the dissociation of the $(DnaB)_6(DnaC)_6$ complex and loading of $(DnaB)_6$ on to ssDNA.[391-393]

6.2 Scope of this Chapter

To date the exact mechanisms of assembly and disassembly of the $(DnaB)_6$ and $(DnaB)_6(DnaC)_6$ complexes have not been elucidated. To understand completely the role of DnaB in the replisome, characterization of the interactions between the subunits and therein other proteins including, DnaC is required.

This chapter describes experiments in which nanoESI-MS has been employed to complement earlier work and enable further understanding of the associations and complexes of DnaB, particularly in its interaction with DnaC. As the $(DnaB)_6$ and $(DnaB)_6(DnaC)_6$ complexes have theoretical masses of 314 and 480 kDa, respectively

(without the addition of magnesium and nucleotides), analysis using commonly-available, commercial ESI mass spectrometers may not be feasible. Therefore to allow analysis of this complex, a factory-modified QToF instrument was used. The instrument is shown schematically in figure 6.3.

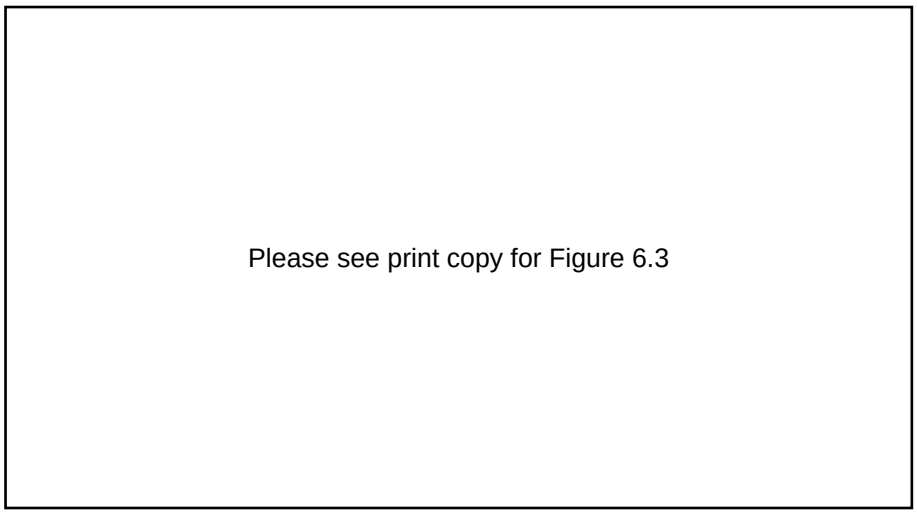


Figure 6.3 A schematic of the Waters QToF Ultima™ used in these experiments. Adopted from the Waters website (http://www.waters.com/).

Modifications to this instrument were based on the instrument developed by Micromass (now Waters) for Robinson and coworkers.[200] These include a resolving quadrupole of *m/z* 32,000, an increased power supply to the collision cell to allow tandem mass spectra of very large macromolecular complexes to be performed, and an additional supply of argon to the ion optics region near the source to assist in collisional cooling of large macromolecules.[199]

Using this instrument, conditions were established to analyze the DnaB complex. Following optimisation of spraying and instrument conditions, the optimal ammonium acetate concentration for achieving resolution sufficiently to resolve two

macromolecular species, a hexamer of DnaB and a heptamer, were determined. The heptamer has not been detected previously by other techniques. Conditions for favouring the heptamer were established.

The complex between DnaB and DnaC was also prepared, and a cooperative interaction was indicated since 5 DnaC monomers bound to the $(DnaB)_6$ before the $(DnaB)_6(DnaC)_6$ complex was formed with no intermediate forms containing fewer than 5 DnaC molecules. Tandem (MS/MS) studies of $(DnaB)_6(DnaC)_6$ also indicated loss of DnaC monomers. Properties of DnaC monomers in the presence of ATP and ADP were also examined and will be discussed in relation to the properties of the $(DnaB)_6(DnaC)_6$ complex.

6.3 Results and Discussion

6.3.1 Establishment of Conditions for the Analysis of DnaB

The analysis of protein samples by ESI-MS has some limitations on the components of solutions used to dissolve the analyte. Commonly, this is limited to volatile salts, such as ammonium acetate or ammonium carbonate, however, the presence of small quantities of non-volatile components can be acceptable, depending on the experimental conditions. This poses a problem for the analysis of DnaB and the DnaB/DnaC complexes as studies using solution techniques commonly employ up to 10 mM $Mg(OAc)_2$, 200 mM NaCl, 1 mM ATP (or ADP or AMP-PNP or ATPγS) in 50 mM Tris.HCl, pH 7.6,[384,386,387,390] all of which are largely not compatible for MS studies. To overcome this problem, solutions of DnaB were exchanged into 50 mM NH_4OAc (pH 7.6), 1 mM $Mg(OAc)_2$ and 0.1 mM ATP (or ADP). The magnesium and nucleotides

were included in this solution as they have been shown to be important components in the stability of the complex (i.e. Mg^{2+}) and in the mechanism of action of the helicase (Mg/ATP or Mg/ADP). The concentrations of these components were reduced to help decrease the volume of ions produced by ESI and increase resolution by reducing non-specific adducts.

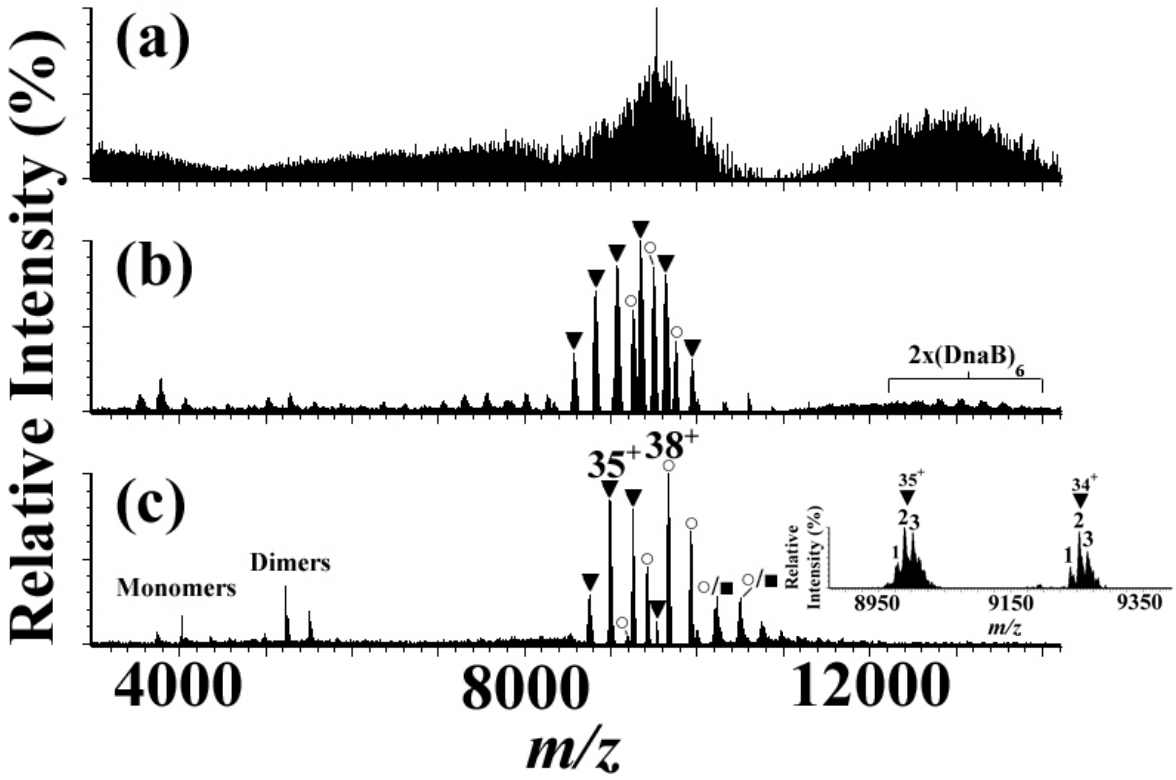


Figure 6.4 ESI mass spectra of the DnaB complex acquired from a solution of: (a) 50 mM NH_4OAc (pH 7.6), 1 mM $Mg(OAc)_2$ and 0.1 mM ATP; (b) 200 mM NH_4OAc (pH 7.6), 1 mM $Mg(OAc)_2$ and 0.1 mM ATP, and (c) 1.5 M NH_4OAc (pH 7.6), 1 mM $Mg(OAc)_2$ and 0.1 mM ATP. Cone and RF lens 1 voltages were set at 430 V and 120 eV in (a) and increased to 550 V and 190 eV, respectively for (b) and (c). ▼ = $(DnaB)_6$, ○ = $(DnaB)_7$, ■ = $(DnaB)_8$, monomers, dimers and dimers of hexamers ($2x(DnaB)_6$) of DnaB species are marker in their respective spectra. The numbers in the insert corresponds to possible numbers of nucleotide molecules bound.

Figure 6.4 shows the ESI mass spectra of DnaB in different solutions. In the spectrum of DnaB when sprayed from a solution 50 mM NH_4OAc (pH 7.6), 1 mM $Mg(OAc)_2$ and 0.1 mM ATP (figure 6.4 (a)) the addition of these non-volatile components produced a

spectrum with two envelopes of ions, one from *m/z* 8500 to 10500 and a second from *m/z* 11500 to 14000. The ions in these envelopes were not sufficiently resolved to allow assignment of the charge on the ions. By increasing the concentration of ammonium acetate to 200 mM, increasing sample cone (or sample orifice) voltages, the potential across the ion optics region (RF lens 1; see figure 6.3) and adjusting the nanospray so that it was moved away from the sample cone, resolution of these species could be improved to allow some charge states to be resolved (figure 6.4 (b)). The increase in the ammonium acetate concentration may have assisted in displacing magnesium adducts from these ions. By increasing the energy of ions through the initial stages of the mass spectrometer, "collisional cleaning" (removal of weakly bound solvent molecules) of ions was achieved owing to collisions with argon molecules bled into the mass spectrometer in the ion optics and collision cell regions.

Varying these conditions led to a significant drop in the intensity of the ions from *m/z* 11500 to 14000, however, some poorly resolved ions could be detected and were possibly from a $2x(DnaB)_6$ complex (figure 6.4 (b)). From *m/z* 8500 to 10500 a significant increase in resolution was observed allowing identification of what appeared to be at least two separate species. Assignment of these species was difficult as the peak width at half-hight was *m/z* 50, however, following a further increase in ammonium acetate concentration to 1.5 M (figure 6.4 (c)), resolution was sufficient to assign these species as the hexamer of DnaB (▼; mass (Da) = 314,130 ± 20) and a previously unknown heptamer (○; mass (Da) = 366,406 ± 20). An octamer (■) may also have been present, however, the resolution was insufficient to distinguish this from the heptamer species. No ions from the $2x(DnaB)_6$ complex could be detected under

these conditions, suggesting this complex occurred as a result of electrostatic interactions that were only present at lower ammonium acetate concentrations.

Monomers of DnaB observed in nanoESI mass spectra of solutions essentially free of Mg^{2+} and nucleotides had a calculated average mass of 52,278.1 ± 0.3 Da. The Expasy calculated mass of DnaB is 52,390, the difference being from a protease cleavage. The mass of $(DnaB)_6$ free from Mg^{2+} and ATP would therefore be expected to be at 313,668 Da, 462 Da below the observed mass. This would suggest that the experimentally observed mass of $(DnaB)_6$ has at least one nucleotide bound (ADP) and possibly $2Mg^{2+}$ ions under these conditions. The resolution of the ions, however, makes it difficult to determine accurately the stoichiometry of ligand binding. In future experiments, DnaB could be expressed in the *E. coli* expression system used for its production in media comprised of pure C^{12}, N^{14} and O^{16} isotopes. This would prevent less abundant isotopes broadening spectra and improve sensitivity as previously reported by Charlebois *et al.*.[110] This experiment would be, however, quite expensive to carry out.

The expansion of the 35^+ and 34^+ charge states of $(DnaB)_6$ (see figure 6.4 (c)) indicates the presence of ions corresponding to species that differ in mass by ~480 Da. These ions are not resolved well enough to distinguish between ATP (+ 505 Da) or ADP (+ 425 Da) binding, however, as DnaB does have some ATPase activity it is likely that both species are present.[271,394] Three high affinity sites for nucleotide binding have been reported previously by Biswas *et al.*.[382] These may also be observed here with three peaks observed for the 35^+ and 34^+ charge states indicating the binding of one, two and three nucleotides, respectively. Interestingly, the mass difference from $(DnaB)_6$ to

$(DnaB)_7$ is 52,276 Da, one DnaB monomer, suggesting that the binding of the seventh DnaB molecule does not require Mg or nucleotides to interact. Possibly this is a non-specific interaction occurring as the result of association of $(DnaB)_6$ and DnaB in solution and/or during ionization.

The observation of a $(DnaB)_7$ species is curious, as previous structural studies (cryoEM) have not detected its presence. The solution conditions used in these experiments may favor formation of the $(DnaB)_7$ species. Secondly, in the current experiments, the concentration of $(DnaB)_6$ was 10 μM, as opposed to cryoEM studies where the concentration of $(DnaB)_6$ was 0.7 μM.[387] ESI-MS studies of insulin at high concentrations (2 mM) have revealed the presence of additional oligomeric states as a result of high protein concentrations.[225] This may suggest the 10-fold increase in concentration used for ESI-MS experiments leads to some aggregation of DnaB not observed in cryoEM studies. A final possibility is that the heptameric species is a biologically relevant species and was simply not able to be detected by other techniques. Native gels and other hydrodynamic measurements may have variations of up to 20% which would not be able to resolve hexameric and heptameric assemblies of DnaB.[395]

There have been some observations made for related proteins which support that $(DnaB)_7$ may have some biological relevance. The human Rad52 helicase,[396] *Thermus thermophilus* RuvB helicase,[397] *Methanabacterium thermoautotrophicum* MCM helicase[398] and the DnaB-like primase-helicase protein from the bacteriophage T7[395] have also been observed to form heptamers. For the bacteriophage T7, helicase and primase activities are present on the one-polypeptide chain. Crystal structures of both

hexamers and heptamers have been determined, indicating subunit packing being similar for both oligomeric forms, however, one of the most noticeable features is the increase in the size of the central channel which varies from 24 to 28 Å for the heptamer.[395] This is large enough to accommodate dsDNA in the heptameric channel. The hexamer, however, can only accommodate ssDNA. The authors of this paper[395] have suggested that the heptamer may play roles in DNA remodelling activity.[399] That is Kaplan and O'Donnell have shown that the helicase can encircle dsDNA and then actively translocate it, displacing bound proteins as it moves.[399] It may be that the $(DnaB)_7$ observed in the current nanoESI mass spectra is the active species in this activity. The cryoEM studies by San Martin *et al.*, however, suggested the central channel of DnaB (C_3 symmetry) is approximately 30 Å in diameter, sufficient in size to accommodate dsDNA into the channel.[386] This makes it uncertain which oligomeric form of DnaB is active in this process, and further studies aimed at purifying $(DnaB)_6$ and $(DnaB)_7$ to test activity and/or determine structures in the presence of ssDNA and dsDNA will be required to resolve these issues.[×]

6.3.2 Conditions for Favouring $(DnaB)_6$ or $(DnaB)_7$

Conditions that favour formation and detection of $(DnaB)_7$ were obtained by addition of methanol to a final concentration of 30% (*v/v*) to a solution of DnaB in 200 mM NH_4OAc (pH 7.6), 1 mM $Mg(OAc)_2$, 0.1 mM ATP (see figure 6.5 (a)).

[×] Preliminary ESI-MS studies suggested that addition of dsDNA to a solution of DnaB oligomers favours $(DnaB)_7$ (data not shown). Mass resolution of these species was insufficient to identify if dsDNA was bound to the heptamer.

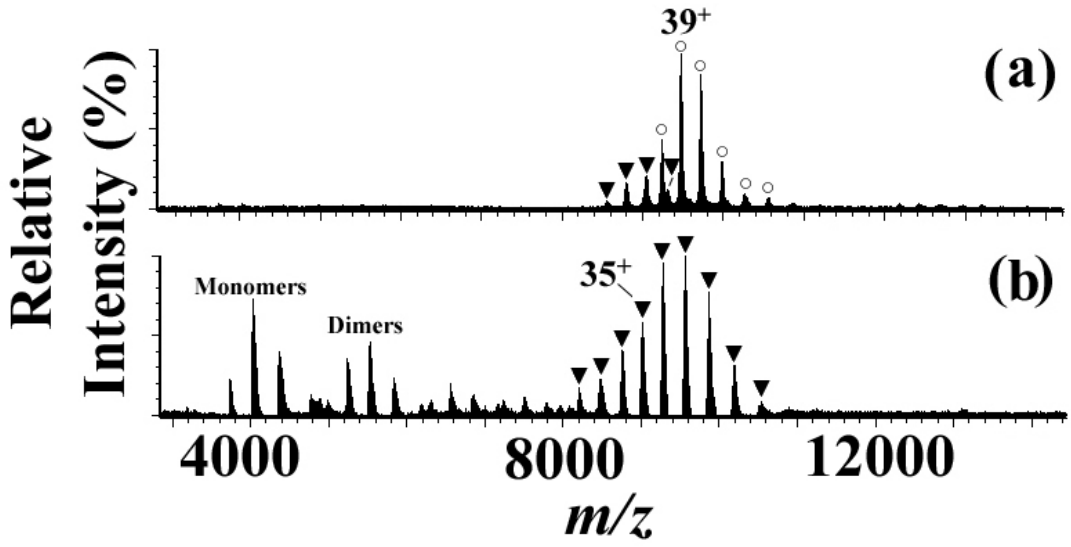


Figure 6.5 NanoESI mass spectra of DnaB complexes obtained acquired from a solution of 200 mM NH_4OAc (pH 7.6), and (a) 1 mM $Mg(OAc)_2$, 0.1 mM ATP plus 30% (*v/v*) methanol, and (b) no additivies. ▼ = $(DnaB)_6$, ○ = $(DnaB)_7$.

Under these conditions the intensities of the ions from $(DnaB)_7$ (○) were greater than for $(DnaB)_6$ ions (▼). It was initially hypothesised in performing the methanol titration experiment, that if the $(DnaB)_7$ complex was stabilized by hydrophobic interactions, dissociation of the $(DnaB)_7$ would occur. Clearly, this is not the case. Recent ESI-MS studies by Gupta *et al.* of the complex between θ- and ε-subunits of the *E. coli* holoenzyme (a complex believed to be maintained by hydrophobic interactions), have shown that it dissociates in the presence of alcohols.[213] In the present work, $(DnaB)_7$ was favoured by the addition of methanol. Complexes involving electrostatic interactions like the Tus-Ter, protein-DNA complex have been able to be dissociated in the presence of high concentrations of ammonium acetate.[323] As seen in figure 6.4 (c), increasing the concentration of ammonium acetate up to 1.5 M did not result in the dissociation of $(DnaB)_6$ or $(DnaB)_7$ species. Ammonium acetate concentrations were further increased, however, this resulted in precipitation of DnaB precluding further analysis. These results do indicate a strong interaction between DnaB subunits in both

oligomeric forms, however $(DnaB)_7$ can be favoured by addition of 30% methanol. The conditions developed here for favouring different oligomeric forms of DnaB, might be used to inform the choice of solvent used in other structural studies, such as in attempts at crystallization.

On the removal of Mg^{2+} by buffer exchange with 3 x 400 µL aliquots of 50 mM NH_4OAc (pH 7.6) before making the concentration of ammonium acetate up to 200 mM, loss of $(DnaB)_7$ was no longer detected (figure 6.5 (b)). An increase in the abundance of monomers and dimers was also observed in the resulting spectra. These results suggest that there is a requirement for magnesium to be present in sufficient quantities to allow the $(DnaB)_7$ complex to form. This is interesting as no further addition of Mg^{2+} was detected in the mass shift from the $(DnaB)_6$ to the $(DnaB)_7$. When this experiment was repeated, but with an additional washing step with 50 mM NH_4OAc (pH 7.6), close to total dissociation of the $(DnaB)_6$ complex occurred leaving ions of $(DnaB)_1$, $(DnaB)_2$, $(DnaB)_3$, $(DnaB)_4$, $(DnaB)_5$ and $(DnaB)_6$ in decreasing abundance. This suggests that under the conditions used for the spectrum in figure 6.5 (b) some residual Mg^{2+} was likely to have been present.

6.3.3 Dissociation of $(DnaB)_6$ and $(DnaB)_7$ by nanoESI-MS/MS

Tandem mass spectrometry (MS/MS) of the $(DnaB)_6$ and $(DnaB)_7$ complexes were also attempted to study the dissociation behaviour of these complexes. $(DnaB)_6$ was not able to be dissociated (when selecting 36^+-34^+ charge states as precursor ions) despite increasing the energy of mass selected ions concomitant with maximising sample cone voltages, reducing pressure in the ion optics region (see figure 6.3) to decrease collisional cooling, and increasing the pressure in and voltage across the collision cell. DnaB does not contain any cysteine residues so the likelihood of covalent linkages

forming in the complex is low. As DnaB is a large protein, however, it is likely that the large number of non-covalent interactions maintain its structure making the complex very stable. In contrast, the $(DnaB)_7$ showed some interesting dissociation behaviour.

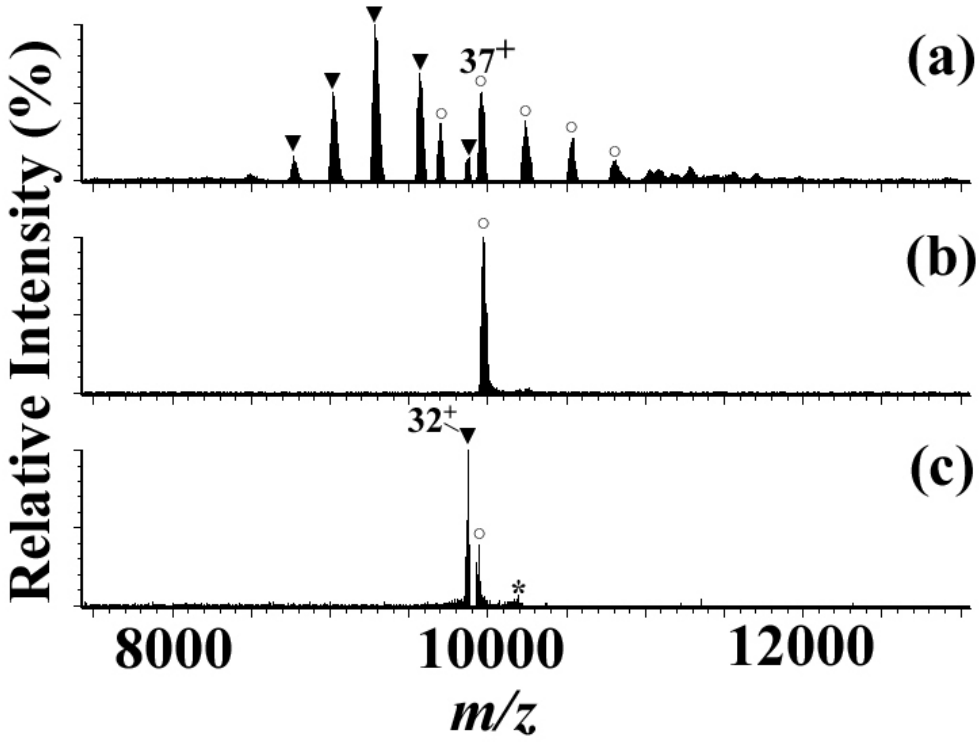


Figure 6.6 MS/MS analysis of $(DnaB)_7$: (a) Mass spectrum of DnaB nanosprayed from a solution of 1 M NH_4OAc (pH 6.5), 1 mM $Mg(OAc)_2$ and 0.1 mM ADP. (b) A mass spectrum of the selected precursor ion (*m/z* 9960; 37^+ charge state of $(DnaB)_6$) and, (c) the resulting MS/MS spectrum using a collision voltage of 90 V. See text for additional changes to the mass spectrometer conditions. ▼ = $(DnaB)_6$, ○ = $(DnaB)_7$ and * = $(DnaB)_1$ fragment ion (residues 14-471 or 13-470; 5^+).

Tandem mass spectrometry of the $(DnaB)_7$ complex using a variety of solution and experimental conditions was attempted. Shown in figure 6.6 (a) is a mass spectrum acquired from a solution of 1 M NH_4OAc (pH 6.5), 1 mM $Mg(OAc)_2$ and 0.1 M ADP. It should be noted that despite the addition of ADP to this sample, both hexamers and heptamers could be observed. When the 37^+ charge state of $(DnaB)_7$ was mass selected

and analysed using relatively energenic conditions,[ω] a decrease in intensity of the 37^+ charge state of $(DnaB)_7$ and an increase in an ion at *m/z* 9879 which is consistent with the 32^+ charge state of the $(DnaB)_6$ was observed (figure 6.6 (c)). An ion of low abundance (denoted by *) was observed at *m/z* 10190. To maintain the charge and mass balance, a monomer of DnaB with a 5^+ charge should be observed. Based on the experimentally determined mass of DnaB, this would give an ion at *m/z* 10457. This is higher than the observed mass (*m/z* 10190) and may suggest that the ion observed is a fragment of DnaB carrying 5 charges. Sequence analysis of DnaB indicated that the sequences 14-471 or 13-470 have a mass of 50945 (5^+ = 10190) supporting the presence of a DnaB fragment. Commonly, MS/MS of non-covalent macromolecular complexes show, monomers dissociated in an asymmetric separation of charges (i.e. a disproportionate amount of charge is carried by monomers relative to the number of protein subunits which make-up the complex).[183,239,400-402] Here a more symmetric dissociation is apparently occurring presumably due to fragmentation of the polypeptide chain. Asymmetric dissociations have been reported to occur as a result of unfolding of subunits.[403,404] Here the interactions between of DnaB monomers in $(DnaB)_7$ appear strong enough to fragment rather than unfold maintaining the charge balance of monomers. Studies by Feng and Yin *et al.* analysing the gas phase collision-induced dissociation of non-covalent protein-ligand complexes maintained by electrostatic interactions, have noticed that despite having a weak solution dissociation constant in the millimolar range, strengthening of interactions in the gas phase can be sufficient to cause fragmentation of polypeptide chains prior to dissociation of the non-covalent complex.[215,405]

[ω] i.e. high source cone voltage, lower pressure in collisional cooling region and applied collision cell voltage of 90 V.

6.3.4 Detecting Quaternary Structural Changes of DnaB

Quaternary structures of DnaB have been shown to be sensitive to the pH of the solvent.[387] When pH ≥ 7.6 was used for cryoEM studies $(DnaB)_6$ existed in C_3 symmetry. When, however, the pH was dropped below 7.2, both C_3 and C_6 symmetries were observed (see figure 6.2).[387] These changes may alter the solvent accessibility of the positively charged sites allowing the different conformations (as judged by changes in ions observed) to be probed by ESI-MS. Figure 6.7 shows the results of experiments carried out with this aim.

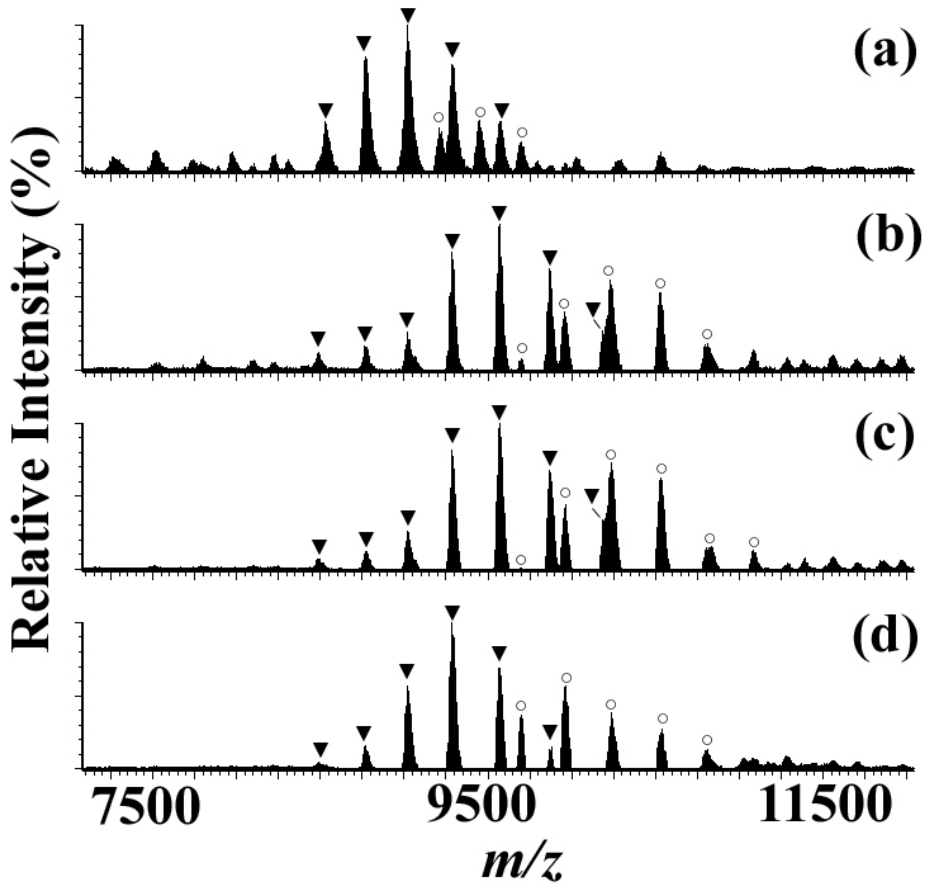


Figure 6.7 The effect pH has on the ESI mass spectra of DnaB complexes in a solution of 1 M NH_4OAc, 1 mM $Mg(OAc)_2$ and 0.1 mM ADP . The pHs of these solutions were: (a) 8.1, (b) 7.6, (c) 7.1 and (d) 6.5. ▼ = $(DnaB)_6$, ○ = $(DnaB)_7$. ESI-MS conditions were as described in the Materials and Methods (chapter 2).

Figure 6.7 (a) shows a spectrum of DnaB acquired from a solution of 1 M NH_4OAc (pH 8.1), 1 mM MgOAc and 0.1 mM ADP. Under these solution conditions the protein is thought to have C_3 symmetry. The spectrum is comprised largely of ions from $(DnaB)_6$ (▼), centred around the 35^+ ion. At this pH, ions from $(DnaB)_7$ ions (○) are not abundant. When the pH was adjusted to 7.6 (figure 6.7 (b)), the most abundant ion is now the 33^+ ion, and the ions from $(DnaB)_7$ have increased in intensity with the 36^+ ion of $(DnaB)_7$ the most abundant. Spectra acquired at pH 7.2 (figure 6.7 (c)) are very similar to those at pH 7.6, however, when the pH was dropped to 6.5 (figure 6.7 (d)), the most abundant ions for $(DnaB)_6$ and $(DnaB)_7$ were 34^+ and 37^+, respectively, i.e. ions with 1 additional charge in each case. Both cryoEM and ESI-MS have a number of artefacts that may affect the observed results for each technique. This makes comparisons between the two techniques difficult. It is interesting to note however from the ESI-MS spectra that as the pH was decreased, approaching conditions where both C_3 and C_6 symmetries have been observed by cryoEM, the intensities of the $(DnaB)_7$ ions increased. It should also be noted that DnaB is an acidic protein (calculated pI of 4.95[282]), and as seen in chapter 4 the ability of ESI-MS to detect changes in protein conformation when ionized under opposite polarity ionization, as seen here, may not allow for these changes to be observed. Experiments in the negative ion mode were also attempted, however, owing to the high Mg^{2+} concentration, ions from DnaB complexes were unable to be detected.

6.3.5 Associations of the DnaB/DnaC Complex

Analysis of the DnaB/DnaC complex has largely centred around the interactions, symmetry and the dissociation processes of the complex. These studies have indicated that in the presence of ATP, DnaB hexamers form complexes with six DnaC molecules.[406,407] A combination of mutational (DnaC binding properties of mutants

were analyzed by protein overlay methods and helicase DNA replication assays)[408] and cryoEM images[390] have indicated N-terminal residues of DnaC appear to interact with the C-terminal domains of DnaB in $(DnaB)_6$, fixing the complex in C_3 symmetry. Dissociation of this complex appears to require the hydrolysis of ATP and interaction with ssDNA so that the $(DnaB)_6$ can be loaded onto ssDNA.[409,410] The order of assembly has not been extensively studied. Here we have used ESI-MS to probe the interactions of DnaB with DnaC complex (see figure 6.8).

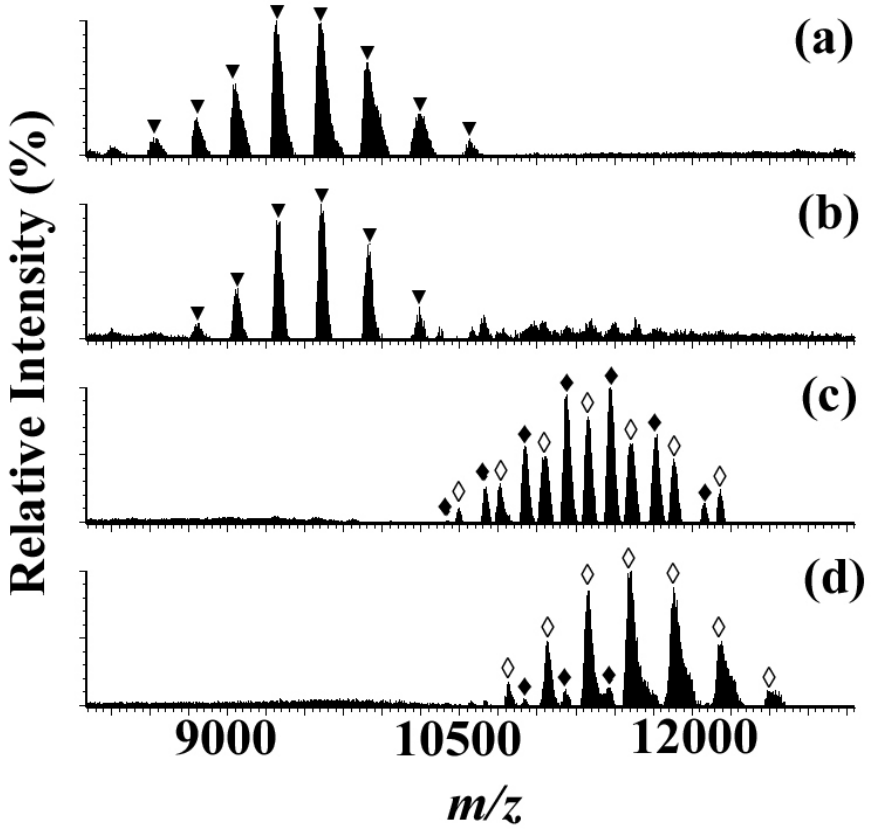


Figure 6.8 NanoESI mass spectra of the titration of DnaC into a solution of DnaB. Concentration of $(DnaB)_6$ was maintained at 10 μM and DnaC monomers were: (a) 10 μM (b) 30 μM (c) 60 μM and (d) 80 μM. A solvent of 200 mM NH_4OAc (pH 7.6), 0.1 mM $Mg(OAc)_2$ and 0.1 mM ATP was used in this experiment. ▼ = $(DnaB)_6$, ♦ = $(DnaB)_6(DnaC)_5$, ◊ = $(DnaB)_6(DnaC)_6$.

As it was found that the DnaB/DnaC complex was prone to precipitation at concentrations of ammonium acetate > 500 mM, solution conditions were kept at 200 mM NH_4OAc (pH 7.6), and 0.1 mM ATP, however, the concentration of $Mg(OAc)_2$ was dropped to 0.1 mM in an attempt to improve resolution. Figure 6.8 (a) shows a spectrum of a solution containing 60 μM DnaB monomers and 10 μM DnaC monomers (the concentration of DnaB monomers was held constant over these experiments). No DnaB/DnaC complexes could be detected. Despite the poor resolution of the ions there does not seem to be any heptamer present under these conditions. $(DnaB)_7$ was observed when sprayed from a solution containing the same concentration of Mg^{2+} used here, without the addition of DnaC. This may suggest that the $(DnaB)_7$ species may only have functional roles in the absence of DnaC or perhaps heptamer formation is a way DnaB regulates its own activities in the absence of DnaC. Increasing the concentration of DnaC monomers to 30 μM (figure 6.8 (b)) did not result in the detection of any DnaB/DnaC complex. When the concentration of DnaC was further increased to a concentration of 60 μM (figure 6.8 (c)), the presence of two new complexes could be detected: one with a calculated mass of 458,790 ± 50 Da corresponding to a complex of $(DnaB)_6$ with 5 DnaC monomers ($(DnaB)_6(DnaC)_5$; ♦), and a second complex with a mass of 486, 745 ± 60 corresponding $(DnaB)_6$ with 6 DnaC monomers ($(DnaB)_6(DnaC)_6$; ◊). As the concentration of DnaC monomers was further increased to 80 μM (figure 6.8 (d)), the $(DnaB)_6(DnaC)_6$ complex was favoured.

These results suggest that the formation of a complex of DnaC with $(DnaB)_6$ is a cooperative process, since no intermediate complexes of $(DnaB)_6$ with 1, 2, 3 or 4 molecules of DnaC were observed. The presence of a $((DnaB)_6(DnaC)_5$ complex is unexpected as cryoEM studies have indicated that DnaC appears to dimerise as it

interacts with DnaB.[390] This species may be particular to the gas phase. No dimers of DnaC have been identified previously in solution, suggesting that this is an interaction which occurs as a result of interactions with $(DnaB)_6$.

6.3.6 Collision-Induced Dissociation of the $(DnaB)_6(DnaC)_6$ Complex

NanoESI-MS/MS was used to examine the stability of the $(DnaB)_6(DnaC)_6$ complex (See figure 6.9).

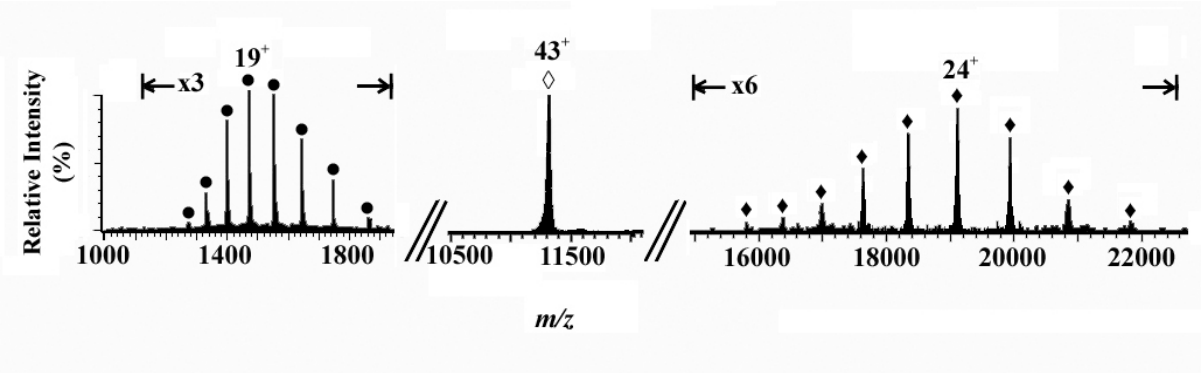


Figure 6.9 MS/MS spectrum of the 43^+ ion of the $(DnaB)_6(DnaC)_6$ complex. The collision cell voltage was increased to 100 V. Solution conditions were 200 mM NH_4OAc (pH 7.6), 0.1 mM $Mg(OAc)_2$ and 0.1 mM ATP. ♦ = $(DnaB)_6(DnaC)_5$, ● = DnaC monomers.

The 43^+ ion of the $(DnaB)_6(DnaC)_6$ complex was selected as precursor ion and MS/MS experiments were performed using similar conditions to those described for the dissociation of the DnaB complexes. When the voltage of the collision cell was increased to 100 V, two products were observed. At low *m/z*, ions from DnaC monomers were detected (●) with the 19^+ ion the most abundant. At high *m/z*, ions from $(DnaB)_6(DnaC)_5$ were observed with the 24^+ ion most abundant. These results suggest that the $(DnaB)_6(DnaC)_6$ complex is very stable. In the current ESI-MS/MS experiments, the loss of monomers seems to correlate with other reports in mass spectrometry literature where dissociation of macromolecular non-covalent complexes in the gas phase occurs by loss of monomers in an asymmetric dissociation.[183,239,400-402]

A similar dissociation mechanism occurred here, with DnaC monomers taking 19 of the 43 available charges, a disproportionate amount considering the number of monomeric species (12) involved in the complex. Studies by Felitsyn *et al*. and Jurchen and Williams have indicated that asymmetric processes of dissociation are affected by the charge of the ion selected, and temperature of the ion, however, it also appears that some unfolding of protein conformation leads to asymmetric dissociation.[403,404,411]

6.3.7 DnaC-ATP/ADP Interactions

Once in the presence of ssDNA, ATP hydrolysis by DnaC has been shown to result in the dissociation of DnaC from the $(DnaB)_6(DnaC)_6$ complex.[410] This has led to suggestions that DnaC is an ATP/ADP switch protein which has different affinities for DnaB when binding the respective nucleotides. Here we have used ESI-MS to probe the interactions of DnaC with ATP and ADP. Figure 6.10 shows spectra of DnaC alone (figure 6.10 (a)), and in the presence of ATP (figure 6.10 (b)) and ADP (figure 6.10 (c)).

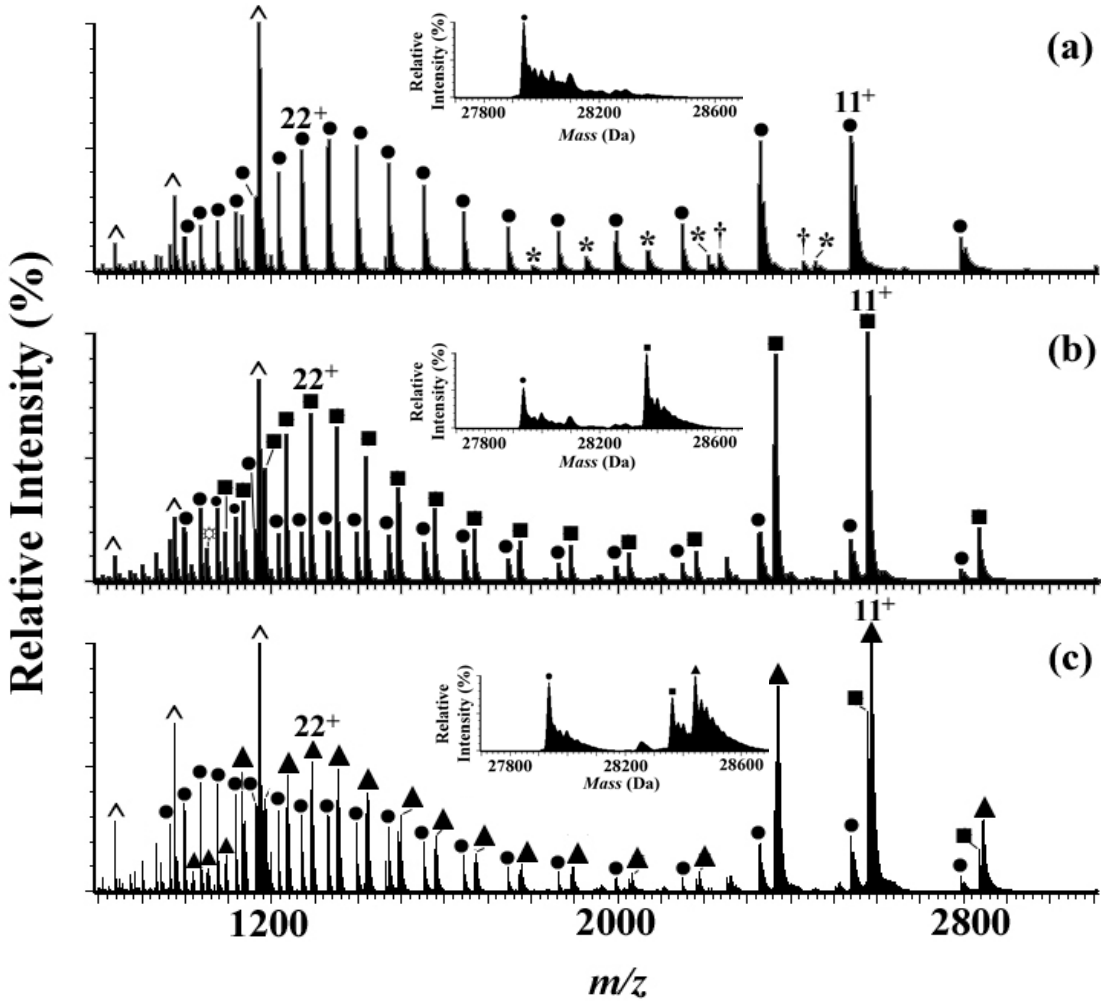


Figure 6.10 ESI mass spectra of DnaC (5 μM) in 10 mM NH_4OAc, pH 7.6. (a) DnaC (b) DnaC + 15 μM ADP (c) DnaC + 15 μM ATP. Solutions containing DnaC and nucleotides were incubated for 60 minutes at room temperature prior to analysis. ● = apoDnaC, * = DnaC dimer, ^ = DnaC fragment (1-51), † = DnaC fragment (52-245), ■ = DnaC+ADP complex and ▲ = DnaC+ATP complex.

The ESI mass spectrum of DnaC in 10 mM NH_4OAc (pH 7.6; figure 6.10 (a)) shows the presence of two ion series. One centred around the 11^+ ion (containing ions from 9^+ to 14^+), and comprising 27% of the intensity of all DnaC ions, and a second around the 21^+ ion (containing ions 15^+ to 28^+), comprising 74% of the intensity of all DnaC ions. Fluorescence studies by Galletto and Bujalowski have also observed a two-state equilibrium for DnaC although 80% of the molecule (a more open conformation) was

believed to be in a nucleotide binding conformation.[391] ESI-MS studies indicate similar populations of molecules in an open and closed conformation. In this ESI-MS spectrum, two DnaC fragments were also observed, one correlating to the mass of residues 1-51 (^) and a second correlating to residues 52-245 (†). These fragments are most likely the result of protease digestion during early stages of protein purification. The intensity of these species did not change significantly when varying instrument parameters suggesting this was not a result of fragmentation during ionization. A small amount of dimer of DnaC (*) was also observed in this spectrum.

With the addition of three time equivalents of ADP, mass shifts corresponding to the binding of one ADP molecule (■; see figure 6.10 (b) insert) were observed. At higher ratios of ADP to DnaC (10:1) a second binding site was observed, however, this was thought to be a non-specific interaction since there have been no other reports of a second binding site. No mass shifts was observed for the 1-51 fragment, however, the 52-245 species did appear to bind ADP. Consistent with this, this latter domain region is believed to contain the AAA+ motif for DnaC.[412] ATP also showed similar binding properties to ADP (see figure 6.10 (c)). The observation of one nucleotide-binding site correlates with UV cross-linking studies and fluorescence studies.[413,414] As seen in the insert of figure 6.10 (c) both DnaC+ATP (▲) and DnaC+ADP (■) complexes could be detected. This indicated that over the 60 minutes where ATP was incubated with DnaC at room temperature, hydrolysis of ATP occurred.

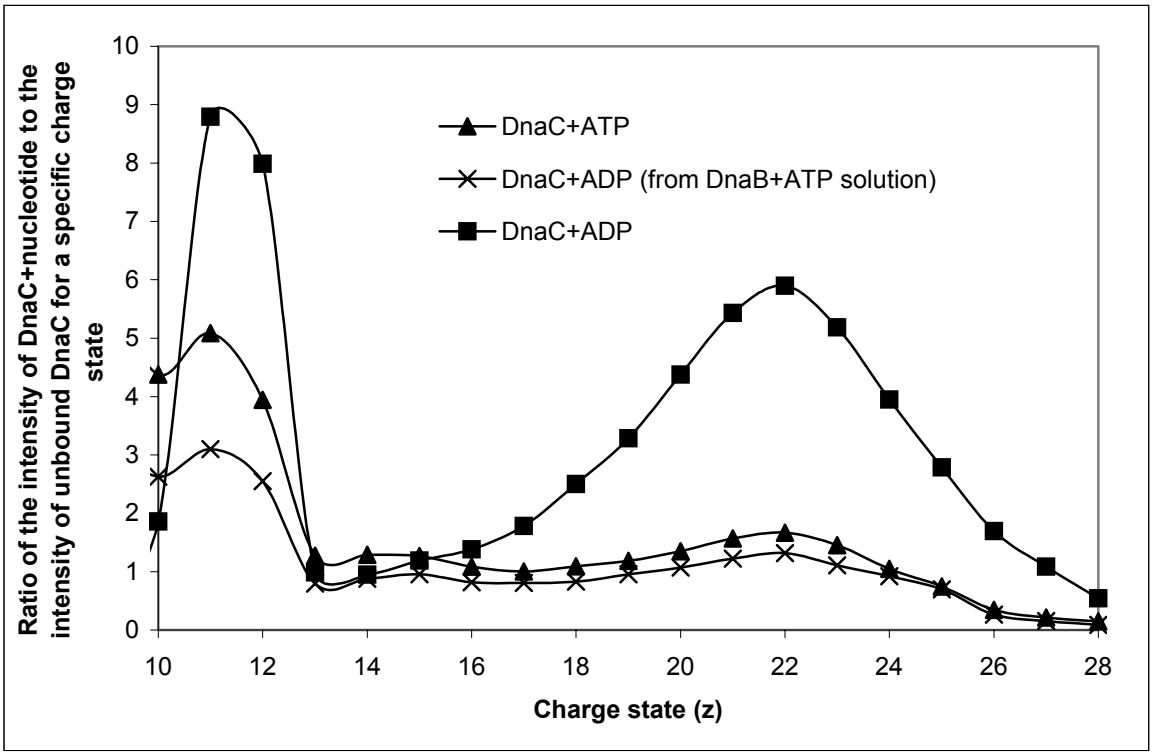


Figure 6.11 Ratios of intensities of DnaC-nucleotide complexes to intensities of unbound DnaC for each individual charge state observed. In each experiment DnaC was equilibrated with nucleotides at a 1:3 ratio for 60 minutes.

To obtain a better indication of the relative affinities of DnaC for nucleotides, for the results shown in figures 6.10 (b) and (c) at each charge state, the intensities of DnaC+nucleotide complexes were divided by the intensity of unbound DnaC. Figure 6.11 shows a plot of these results where the ratio of DnaC+ATP complex to unbound DnaC is represented by ▲, DnaC+ADP complex (from enzyme hydrolysis of ATP) to unbound DnaC from the DnaC+ATP mixture is represented by **x**, and DnaC+ADP complex to unbound DnaC from the ADP mixture is represented by ▲. Both DnaC+ATP and DnaC+ADP mixtures show that for ions with lower charge (11^+ and 12^+) the amount of nucleotide bound was quite high in comparison to unbound DnaC. Reference to the more highly charged ions (15^+-28^+) shows there were significant ions from DnaC-ADP (relative to ions from DnaC alone, peaking at 22^+), but not DnaC-

ATP. The relative abundance of ions from more highly charged DnaC+ATP were comparable with the relative abundance of ions from DnaC alone (ratio ~1). Ions in ESI mass spectra of proteins with a relative high number of charges are thought to arise from molecules present in a conformation which is more unfolded than ions with lower charges.[121,122] This would suggest the structure of the molecules of DnaC which bind ADP are in both a structured and unstructured conformation. Molecules from DnaC complexes with ATP are more likely in the structured conformation. Increasing the ratio of ATP+DnaC to 10:1 (not shown) resulted in the binding of ATP in higher abundances for highly charged ions, as seen for the ADP complex, suggesting that the binding of either nucleotide might cause this conformational change, but that ADP has a higher affinity for DnaC than ATP. The DnaC+ATP spectra are complicated by the hydrolysis of ATP with time to form ADP as the product. The profile from the DnaC+ADP (hydrolysed product) complex (**x**) is similar to the DnaC+ATP complex. Most likely, this is because the product, ADP, is not present at high enough concentrations to cause the conformational change. Ratios of these species are lower than observed for the DnaC+ADP solution as intensities of ions are distributed over three species, DnaC, DnaC+ATP and DnaC+ADP. Similar experiments were attempted in the presence of non-hydrolysable βγ-CH_2-ATP, however, a low binding affinity was observed under these conditions making comparisons between experiments difficult.

Presently there is only limited information regarding the structures of DnaC in complexes with nucleotides. Sedimentation velocity studies by Galletto *et al*. have indicated similar global structures for DnaC+ATP and DnaC+ADP complexes.[415] Fluorescence studies have indicated that in the presence of MANT(3'-O-(N-methylanthraniloyl))-ATP, DnaC has a flexible structure around the nucleotide binding

site. Interactions with MANT-ADP, however, indicated a more rigid structure around the nucleotide-binding site.[415] This would be in contrast to results observed here by ESI-MS, however, due to different solvent conditions and detection methods comparisons are difficult. DnaC has been considered an ATP/ADP switch protein where ATP hydrolysis to ADP allows helicase activities to occur. These different binding properties may suggest mechanisms for the dissociation of DnaC from DnaB, as the presence of small amounts of ADP would allow a change in the conformation of DnaC triggering dissociation from the helicase.

No ATPase activity has previously been reported for DnaC except in the presence of ssDNA, DnaB or both ssDNA and DnaB.[393] This is, therefore, the first report of ATPase activity of DnaC when free from other cofactors. The biological relevance of DnaC being able to hydrolyse ATP without the addition of co-factors is likely to be minimal. The rate of hydrolysis of ATP cannot be determined accurately from the current experiments, however, following determination of dissociation constants of ATP and ADP in the ESI-MS experiment and monitoring hydrolysis over a range of ratios (nucleotide to DnaC) rates may be accurately determined.

6.4 Conclusions

Using an extended mass range ESI mass spectrometer, conditions for the analysis of DnaB oligomers and their interaction with DnaC were determined. Analysis of DnaB resolved at least two oligomeric states of DnaB, the expected hexamer and a previously unreported heptamer. The presence of the heptamer appears to have a dependence on the concentration of the protein and the presence of magnesium but can be favoured in the presence of 30% methanol. These latter conditions may allow characterisation of the heptamer by X-ray crystallography and cryoEM techniques. Another member of the

DnaB like-family has been identified to exist as a hexamer and a heptamer,[395] with the heptameric species possibly having link to DNA recombination activities. Similar properties of the DnaB will be a topic of future experiments.

The interaction between DnaB and is loading partner DnaC was also probed. Titration of DnaC into complexes of DnaB indicated a cooperative binding to DnaB hexamers. No species other than $(DnaB)_6(DnaC)_6$ and $(DnaB)_6(DnaC)_5$ were observed. Tandem mass spectra of $(DnaB)_6(DnaC)_6$ showed the loss of DnaC monomers and the formation of $(DnaB)_6(DnaC)_5$. This observation is consistent with results from similar earlier experiments of other macromolecular complexes. The interactions between DnaC and the nucleotides ATP and ADP were also probed indicating ATP could be hydrolysed by DnaC without the addition of other cofactors (ssDNA, DnaB). ATP and ADP were observed to bind at different affinities to protein molecules carrying few charges in the DnaC ESI mass spectrum.

References

1. Jaenicke, R. (1998) *Biological Chemistry* **379**, 237-243.

2. Radford, S.E. & Dobson, C.M. (1999) *Cell* **97**, 291-298.

3. Stryer, L. (1995) *Biochemistry*; 4th edition, W. H. Freeman and company, New York.

4. Branden, C. & Tooze, J. (1999) *Introduction to protein structure*; 2nd edition, Garland Publishers, New York.

5. Snow, C.D., Nguyen, N., Pande, V.S. & Gruebele, M. (2002) *Nature* **420**, 102-106.

6. Eaton, W.A., Munoz, V., Thompson, P.A., Henry, E.R. & Hofrichter, J. (1998) *Accounts of Chemical Research* **31**, 745-753.

7. Yang, W.Y. & Gruebele, M. (2003) *Nature* **423**, 193-197.

8. Mayor, U., Guydosh, N.R., Johnson, C.M., Grossmann, J.G., Sato, S., Jas, G.S., Freund, S.M.V., Alonso, D.O.V., Daggett, V. & Fersht, A.R. (2003) *Nature* **421**, 863-867.

9. Plaxco, K.W., Simons, K.T. & Baker, D. (1998) *Journal of Molecular Biology* **277**, 985-994.

10. Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. (1963) *Journal of Molecular Biology* **7**, 95–99.

11. Ramachandran, G.N. & Sasisekharan, V. (1968) *Advances in Protein Chemistry* **23**, 283–438.

12. Levinthal, C. (1968) *Journal De Chimie Physique Et De Physico-Chimie Biologique* **65**, 44.

13. Levinthal, C. (1969) in *Mossbauer Spectroscopy in Biological Systems,* 22-24, Allerton House, Monticello.

14. Zwanzig, R., Szabo, A. & Bagchi, B. (1992) *Proceedings of the National Academy of Sciences of the United States of America* **89**, 20-22.

15. Plotkin, S.S. & Onuchic, J.N. (2002) *Quarterly Reviews of Biophysics* **35**, 111-167.

16. Dobson, C.M. (2003) *Nature* **426**, 884-890.

17. Skolnick, J., Kolinski, A. & Yaris, R. (1988) *Proceedings of the National Academy of Sciences of the United States of America* **85**, 5057-5061.

18. Honeycutt, J.D. & Thirumalai, D. (1990) *Proceedings of the National Academy of Sciences of the United States of America* **87**, 3526-3529.

19. Dill, K.A. & Chan, H.S. (1997) *Nature Structural Biology* **4**, 10-19.

20. Brockwell, D.J., Smith, D.A. & Radford, S.E. (2000) *Current Opinion in Structural Biology* **10**, 16-25.

21. Onuchic, J.N. & Wolynes, P.G. (2004) *Current Opinion in Structural Biology* **14**, 70-75.

22. Vendruscolo, M., Paci, E., Dobson, C.M. & Karplus, M. (2001) *Nature* **409**, 641-645.

23. Riddle, D.S., Grantcharova, V.P., Santiago, J.V., Alm, E., Ruczinski, I. & Baker, D. (1999) *Nature Structural Biology* **6**, 1016-1024.

24. Fersht, A.R. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1525-1529.

25. Martinez, J.C. & Serrano, L. (1999) *Nature Structural Biology* **6**, 1010-1016.

26. Chiti, F., Taddei, N., White, P.M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C.M. (1999) *Nature Structural Biology* **6**, 1005-1009.

27. Cheung, M.S., Chavez, L.L. & Onuchic, J.N. (2004) *Polymer* **45**, 547-555.

28. Nelson, D.L. & Cox, M.M. (2005) *Lehninger Principles of Biochemistry*; 4th edition, W.H. Freeman, New York.

29. Garcia, K.C., Degano, M., Pease, L.R., Huang, M.D., Peterson, P.A., Teyton, L. & Wilson, I.A. (1998) *Science* **279**, 1166-1172.

30. Boniface, J.J., Reich, Z., Lyons, D.S. & Davis, M.M. (1999) *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11446-11451.

31. Baldwin, J. & Chothia, C. (1979) *Journal of Molecular Biology* **129**, 175-200.

32. Paoli, M., Liddington, R., Tame, J., Wilkinson, A. & Dodson, G. (1996) *Journal of Molecular Biology* **256**, 775-792.

33. Perutz, M.F., Wilkinson, A.J., Paoli, M. & Dodson, G.G. (1998) *Annual Review of Biophysics and Biomolecular Structure* **27**, 1-34.

34. Shikama, K. & Matsuoka, A. (2003) *European Journal of Biochemistry* **270**, 4041-4051.

35. Schnell, J.R., Dyson, H.J. & Wright, P.E. (2004) *Annual Review of Biophysics and Biomolecular Structure* **33**, 119-140.

36. Huennekens, F.M. (1996) *Protein Science* **5**, 1201-1208.

37. Then, R.L. (2004) *Journal of Chemotherapy* **16**, 3-12.

38. Brunger, A.T. (1997) *Nature Structural Biology* **4**, 862-865.

39. Smyth, M.S. & Martin, J.H.J. (2000) *Journal of Clinical Pathology-Molecular Pathology* **53**, 8-14.

40. Kundrot, C.E. (2004) *Cellular and Molecular Life Sciences* **61**, 525-536.

41. Palmer, A.G., Kroenke, C.D. & Loria, J.P. (2001) in *Nuclear Magnetic Resonance of Biological Macromolecules, Pt B*, **339**, 204-238.

42. Akke, M. (2002) *Current Opinion in Structural Biology* **12**, 642-647.

43. Cole, R. & Loria, J.P. (2002) *Biochemistry* **41**, 6072-6081.

44. Wang, L.C., Pang, Y.X., Holder, T., Brender, J.R., Kurochkin, A.V. & Zuiderweg, E.R.P. (2001) *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7684-7689.

45. Rozovsky, S., Jogl, G., Tong, L. & McDermott, A.E. (2001) *Journal of Molecular Biology* **310**, 271-280.

46. Lu, H.P. (2004) *Current Pharmaceutical Biotechnology* **5**, 261-269.

47. Weiss, S. (2000) *Nature Structural Biology* **7**, 724-729.

48. Witt, S.N. & Slepenkov, S.V. (1999) *Journal of Fluorescence* **9**, 281-293.

49. Weiss, S. (1999) *Science* **283**, 1676-1683.

50. Craig, D.B., Arriaga, E.A., Wong, J.C.Y., Lu, H. & Dovichi, N.J. (1996) *Journal of the American Chemical Society* **118**, 5245-5253.

51. Tan, W.H. & Yeung, E.S. (1997) *Analytical Chemistry* **69**, 4242-4248.

52. Anthonycahill, S.J., Griffith, M.C., Noren, C.J., Suich, D.J. & Schultz, P.G. (1989) *Trends in Biochemical Sciences* **14**, 400-403.

53. Cornish, V.W., Benson, D.R., Altenbach, C.A., Hideg, K., Hubbell, W.L. & Schultz, P.G. (1994) *Proceedings of the National Academy of Sciences of the United States of America* **91**, 2910-2914.

54. Greenfield, N.J. (1999) *Trends in Analytical Chemistry* **18**, 236-244.

55. Barth, A. & Zscherp, C. (2002) *Quarterly Reviews of Biophysics* **35**, 369-430.

56. Mulvaney, S.P. & Keating, C.D. (2000) *Analytical Chemistry* **72**, 145R-157R.

57. Wong, S.F., Meng, C.K. & Fenn, J.B. (1988) *Journal of Physical Chemistry* **92**, 546-550.

58. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y. & Yohida, T. (1988) *Rapid Communications in Mass Spectrometry* **2**, 151-153.

59. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. & Whitehouse, C.M. (1989) *Science* **246**, 64-71.

60. Sundqvist, B. & Macfarlane, R.D. (1985) *Mass Spectrometry Reviews* **4**, 421-460.

61. Cotter, R.J. (1988) *Analytical Chemistry* **60**, 781A-782A.

62. Hansen, G. & Munson, B. (1978) *Analytical Chemistry* **50**, 1130-1134.

63. Barber, M., Bordoli, R.S., Sedgwick, R.D., Tyler, A.N. & Bycroft, B.W. (1981) *Biochemical and Biophysical Research Communications* **101**, 632-638.

64. Karas, M. & Hillenkamp, F. (1988) *Analytical Chemistry* **60**, 2299-2301.

65. Beckey, H.D. (1969) *International Journal of Mass Spectrometry and Ion Physics* **2**, 500-503.

66. Schmelzeisen-Redeker, G., Buetfering, L. & Roellgen, F.W. (1989) *International Journal of Mass Spectrometry and Ion Processes* **90**, 139-150.

67. Dole, M., Mack, L.L. & Hines, R.L. (1968) *Journal of Chemical Physics* **49**, 2240.

68. Iribarne, J.V. & Thomson, B.A. (1976) *Journal of Chemical Physics* **64**, 2287-2294.

69. Thomson, B.A. & Iribarne, J.V. (1979) *Journal of Chemical Physics* **71**, 4451-4463.

70. Cole, R.B. (2000) *Journal of Mass Spectrometry* **35**, 763-772.

71. Fenn, J.B. (1993) *Journal of the American Society for Mass Spectrometry* **4**, 524-535.

72. Felitsyn, N., Peschke, M. & Kebarle, P. (2002) *International Journal of Mass Spectrometry* **219**, 39-62.

73. Schwartz, B.L., Rockwood, A.L., Smith, R.D., Tomalia, D.A. & Spindler, R. (1995) *Rapid Communications in Mass Spectrometry* **9**, 1552-1555.

74. Tolic, L.P., Anderson, G.A., Smith, R.D., Brothers, H.M., Spindler, R. & Tomalia, D.A. (1997) *International Journal of Mass Spectrometry* **165**, 405-418.

75. Fernandez de la Mora, J. (2000) *Analytica Chimica Acta* **406**, 93-104.

76. Gamero-Castano, M. & de la Mora, J.F. (2000) *Analytica Chimica Acta* **406**, 67-91.

77. Kebarle, P. (2000) *Journal of Mass Spectrometry* **35**, 804-817.

78. Nesatyy, V.J. & Suter, M.J.F. (2004) *Journal of Mass Spectrometry* **39**, 93-97.

79. Bai, Y., Milne, J.S., Mayne, L. & Englander, S.W. (1993) *Proteins* **17**, 75-86.

80. Lenormant, H. & Blout, E.R. (1953) *Nature* **172**, 770-771.

81. Hvidt, A. & Linderstrom-Lang, K. (1954) *Biochimica et Biophysica Acta* **14**, 574-575.

82. Leach, S.J. & Hill, J. (1963) *Biochemistry* **128**, 807-813.

83. Saunders, M. & Wishnia, A. (1958) *Annals of the New York Academy of Sciences* **70**, 870-874.

84. Wagner, G., Stassinopoulou, C.I. & Wuthrich, K. (1984) *European Journal of Biochemistry* **145**, 431-436.

85. Wand, A.J., Roder, H. & Englander, S.W. (1986) *Biochemistry* **25**, 1107-1114.

86. Roder, H. & Wuthrich, K. (1986) *Proteins: Structure, Function, and Genetics* **1**, 34-42.

87. Katta, V. & Chait, B.T. (1991) *Rapid Communications in Mass Spectrometry* **5**, 214-217.

88. Thevenon-Emeric, G., Kozlowski, J., Zhang, Z. & Smith, D.L. (1992) *Analytical Chemistry* **64**, 2456-2458.

89. Miranker, A., Robinson, C.V., Radford, S.E., Aplin, R.T. & Dobson, C.M. (1993) *Science* **262**, 896-900.

90. Kaltashov, I.A. & Eyles, S.J. (2002) *Mass Spectrometry Reviews* **21**, 37-71.

91. Zhang, Z.Q. & Smith, D.L. (1993) *Protein Science* **2**, 522-531.

92. Tito, P., Nettleton, E.J. & Robinson, C.V. (2000) *Journal of Molecular Biology* **303**, 267-278.

93. Kim, M.Y., Maier, C.S., Reed, D.J. & Deinzer, M.L. (2001) *Journal of the American Chemical Society* **123**, 9860-9866.

94. Smith, D.L., Deng, Y.Z. & Zhang, Z.Q. (1997) *Journal of Mass Spectrometry* **32**, 135-146.

95. Mandell, J.G., Falick, A.M. & Komives, E.A. (1998) *Analytical Chemistry* **70**, 3987-3995.

96. Ghaemmaghami, S., Fitzgerald, M.C. & Oas, T.G. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8296-8301.

97. Powell, K.D. & Fitzgerald, M.C. (2001) *Analytical Chemistry* **73**, 3300-3304.

98. Powell, K.D., Wales, T.E. & Fitzgerald, M.C. (2002) *Protein Science* **11**, 841-851.

99. Nazabal, A., Dos Reis, S., Bonneu, M., Saupe, S.J. & Schmitter, J.M. (2003) *Biochemistry* **42**, 8852-8861.

100. Akashi, S., Naito, Y. & Takio, K. (1999) *Analytical Chemistry* **71**, 4974-4980.

101. Eyles, S.J., Dresch, T., Gierasch, L.M. & Kaltashov, I.A. (1999) *Journal of Mass Spectrometry* **34**, 1289-1295.

102. Eyles, S.J., Speir, J.P., Kruppa, G.H., Gierasch, L.M. & Kaltashov, I.A. (2000) *Journal of the American Chemical Society* **122**, 495-500.

103. Akashi, S. & Takio, K. (2000) *Protein Science* **9**, 2497-2505.

104. Hoerner, J.K., Xiao, H., Dobo, A. & Kaltashov, I.A. (2004) *Journal of the American Chemical Society* **126**, 7709-7717.

105. Zubarev, R.A., Kelleher, N.L. & McLafferty, F.W. (1998) *Journal of the American Chemical Society* **120**, 3265-3266.

106. Zubarev, R.A., Horn, D.M., Fridriksson, E.K., Kelleher, N.L., Kruger, N.A., Lewis, M.A., Carpenter, B.K. & McLafferty, F.W. (2000) *Analytical Chemistry* **72**, 563-573.

107. McLafferty, F.W., Horn, D.M., Breuker, K., Ge, Y., Lewis, M.A., Cerda, B., Zubarev, R.A. & Carpenter, B.K. (2001) *Journal of the American Society for Mass Spectrometry* **12**, 245-249.

108. Hakansson, K., Cooper, H.J., Emmett, M.R., Costello, C.E., Marshall, A.G. & Nilsson, C.L. (2001) *Analytical Chemistry* **73**, 4530-4536.

109. Roepstorff, P. & Fohlman, J. (1984) *Biomedical Mass Spectrometry* **11**, 601-601.

110. Charlebois, J.P., Patrie, S.M. & Kelleher, N.L. (2003) *Analytical Chemistry* **75**, 3263-3266.

111. Lundblad, R.L. (1995) *Techniques in Protein Modification*; CRC Press, Boca Raton.

112. Lundblad, R.L. (2005) *Chemical Reagents for Protein Modification*; CRC Press, Boca Raton.

113. Sinz, A. (2003) *Journal of Mass Spectrometry* **38**, 1225-1237.

114. Young, M.M., Tang, N., Hempel, J.C., Oshiro, C.M., Taylor, E.W., Kuntz, I.D., Gibson, B.W. & Dollinger, G. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5802-5806.

115. Rappsilber, J., Siniossoglou, S., Hurt, E.C. & Mann, M. (2000) *Analytical Chemistry* **72**, 267-275.

116. Taverner, T., Hall, N.E., O'Hair, R.A.J. & Simpson, R.J. (2002) *Journal of Biological Chemistry* **277**, 46487-46492.

117. Bennett, K.L., Kussmann, M., Bjork, P., Godzwon, M., Mikkelsen, M., Sorensen, P. & Roepstorff, P. (2000) *Protein Science* **9**, 1503-1518.

118. Muller, D.R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S. & Steinmetz, M.O. (2001) *Analytical Chemistry* **73**, 1927-1934.

119. Pearson, K.M., Pannell, L.K. & Fales, H.M. (2002) *Rapid Communications in Mass Spectrometry* **16**, 149-159.

120. Huang, B.X., Kim, H.Y. & Dass, C. (2004) *Journal of the American Society for Mass Spectrometry* **15**, 1237-1247.

121. Chowdhury, S.K., Katta, V. & Chait, B.T. (1990) *Journal of the American Chemical Society* **112**, 9012-9013.

122. Loo, J.A., Loo, R.R.O., Udseth, H.R., Edmonds, C.G. & Smith, R.D. (1991) *Rapid Communications in Mass Spectrometry* **5**, 101-105.

123. Konermann, L. & Douglas, D.J. (1997) *Biochemistry* **36**, 12296-12302.

124. Konermann, L. & Douglas, D.J. (1998) *Journal of the American Society for Mass Spectrometry* **9**, 1248-1254.

125. Grandori, R., Matecko, I. & Muller, N. (2002) *Journal of Mass Spectrometry* **37**, 191-196.

126. Grandori, R., Matecko, I., Mayr, P. & Muller, N. (2001) *Journal of Mass Spectrometry* **36**, 918-922.

127. Konermann, L., Collings, B.A. & Douglas, D.J. (1997) *Biochemistry* **36**, 5554-5559.

128. Dobo, A. & Kaltashov, I.A. (2001) *Analytical Chemistry* **73**, 4763-4773.

129. Gilmanshin, R., Gulotta, M., Dyer, R.B. & Callender, R.H. (2001) *Biochemistry* **40**, 5127-5136.

130. Chi, Z.H. & Asher, S.A. (1999) *Biochemistry* **38**, 8196-8203.

131. Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. & Whitehouse, C.M. (1990) *Mass Spectrometry Reviews* **9**, 37-70.

132. Peschke, M., Blades, A. & Kebarle, P. (2002) *Journal of the American Chemical Society,* **124**, 11519-11530.

133. Blades, A., Peschke, M., Verkerk, U. & Kebarle, P. (2002) *Journal of Physical Chemistry A* **106**, 10037-10042.

134. Grandori, R. (2003) *Journal of Mass Spectrometry* **38**, 11-15.

135. De la Mora, J.F., Van Berkel, G.J., Enke, C.G., Cole, R.B., Martinez-Sanchez, M. & Fenn, J.B. (2000) *Journal of Mass Spectrometry* **35**, 939-952.

136. Lindskog, S. (1997) *Pharmacology & Therapeutics* **74**, 1-20.

137. Mavroidis, C., Dubey, A. & Yarmush, M.L. (2004) *Annual Review of Biomedical Engineering* **6**, 363-395.

138. Muller, V. & Gruber, G. (2003) *Cellular and Molecular Life Sciences* **60**, 474-494.

139. Oster, G. & Wang, H.Y. (2003) *Trends in Cell Biology* **13**, 114-121.

140. Marx, S. (2003) *Journal of Molecular and Cellular Cardiology* **35**, 37-44.

141. Ramakrishnan, V. (2002) *Cell* **108**, 557-572.

142. Doudna, J.A. & Rath, V.L. (2002) *Cell* **109**, 153-156.

143. Culver, G.M. (2003) *Biopolymers* **68**, 234-249.

144. Baker, T.A. & Bell, S.P. (1998) *Cell* **92**, 295-305.

145. Davey, M.J. & O'Donnell, M. (2000) *Current Opinion in Chemical Biology* **4**, 581-586.

146. Benkovic, S.J., Valentine, A.M. & Salinas, F. (2001) *Annual Review of Biochemistry* **70**, 181-208.

147. Nogales, E., Wolf, S.G. & Downing, K.H. (1998) *Nature* **391**, 199-203.

148. Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J.B., Engel, A. & Fujiyoshi, Y. (2000) *Nature* **407**, 599-605.

149. Henderson, R. (2004) *Quarterly Reviews of Biophysics* **37**, 3-13.

150. Chiu, W., Baker, M.L., Jiang, W. & Zhou, Z.H. (2002) *Current Opinion in Structural Biology* **12**, 263-269.

151. Zhou, Z.H., Baker, M.L., Jiang, W., Dougherty, M., Jakana, J., Dong, G., Lu, G.Y. & Chiu, W. (2001) *Nature Structural Biology* **8**, 868-873.

152. Hensley, P. (1996) *Structure* **4**, 367-373.

153. Jelesarov, I. & Bosshard, H.R. (1999) *Journal of Molecular Recognition* **12**, 3-18.

154. Schuler, B., Lipman, E.A. & Eaton, W.A. (2002) *Nature* **419**, 743-747.

155. Szabo, A., Stolz, L. & Granzow, R. (1995) *Current Opinion in Structural Biology* **5**, 699-705.

156. Stafford, W.F., Mabuchi, K., Takahashi, K. & Tao, T. (1995) *Journal of Biological Chemistry* **270**, 10576-10579.

157. McDonnell, J.M. (2001) *Current Opinion in Chemical Biology* **5**, 572-577.

158. Homola, J. (2003) *Analytical and Bioanalytical Chemistry* **377**, 528-539.

159. Morton, T.A., Myszka, D.G. & Chaiken, I.M. (1995) *Analytical Biochemistry* **227**, 176-185.

160. Rich, R.L. & Myszka, D.G. (2003) *Trends in Microbiology* **11**, 124-133.

161. Leavitt, S. & Freire, E. (2001) *Current Opinion in Structural Biology* **11**, 560-566.

162. Bradshaw, J.M., Mitaxov, V. & Waksman, G. (2000) *Journal of Molecular Biology* **299**, 521-535.

163. Lostao, A., El Harrous, M., Daoudi, F., Romero, A., Parody-Morreale, A. & Sancho, J. (2000) *Journal of Biological Chemistry* **275**, 9518-9526.

164. Ye, H. & Wu, H. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8961-8966.

165. Filimonov, V.V. & Rogov, V.V. (1996) *Journal of Molecular Biology* **255**, 767-777.

166. Ganem, B., Li, Y.T. & Henion, J.D. (1991) *Journal of the American Chemical Society* **113**, 6294-6296.

167. Katta, V. & Chait, B.T. (1991) *Journal of the American Chemical Society* **113**, 8534-8535.

168. Loo, J.A., Loo, R.R.O. & Andrews, P.C. (1993) *Organic Mass Spectrometry* **28**, 1640-1649.

169. Cheng, X.H., Harms, A.C., Goudreau, P.N., Terwilliger, T.C. & Smith, R.D. (1996) *Proceedings of the National Academy of Sciences of the United States of America* **93**, 7022-7027.

170. Potier, N., Donald, L.J., Chernushevich, I., Ayed, A., Ens, W., Arrowsmith, C.H., Standing, K.G. & Duckworth, H.W. (1998) *Protein Science* **7**, 1388-1395.

171. Hsieh, Y.L., Li, Y.T., Henion, J.D. & Ganem, B. (1994) *Biological Mass Spectrometry* **23**, 272-276.

172. Drummond, J.T., Loo, R.R.O. & Matthews, R.G. (1993) *Biochemistry* **32**, 9282-9289.

173. Loo, J.A. (1997) *Mass Spectrometry Reviews* **16**, 1-23.

174. Beck, J.L., Colgrave, M.L., Ralph, S.F. & Sheil, M.M. (2001) *Mass Spectrometry Reviews* **20**, 61-87.

175. Veenstra, T.D. (1999) *Biophysical Chemistry* **79**, 63-79.

176. McLafferty, F.W. (1981) *Science* **214**, 280-287.

177. Schwartz, B.L., Bruce, J.E., Anderson, G.A., Hofstadler, S.A., Rockwood, A.L., Smith, R.D., Chilkoti, A. & Stayton, P.S. (1995) *Journal of the American Society for Mass Spectrometry* **6**, 459-465.

178. Fitzgerald, M.C., Chernushevich, I., Standing, K.G., Whitman, C.P. & Kent, S.B.H. (1996) *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6851-6856.

179. Robinson, C.V., Chung, E.W., Kragelund, B.B., Knudsen, J., Aplin, R.T., Poulsen, F.M. & Dobson, C.M. (1996) *Journal of the American Chemical Society* **118**, 8646-8653.

180. Veenstra, T.D., Johnson, K.L., Tomlinson, A.J., Naylor, S. & Kumar, R. (1997) *European Mass Spectrometry* **3**, 453-459.

181. Pramanik, B.N., Bartner, P.L., Mirza, U.A., Liu, Y.H. & Ganguly, A.K. (1998) *Journal of Mass Spectrometry* **33**, 911-920.

182. Wilm, M. & Mann, M. (1996) *Analytical Chemistry* **68**, 1-8.

183. Light-Wahl, K.J., Schwartz, B.L. & Smith, R.D. (1994) *Journal of the American Chemical Society* **116**, 5271-5278.

184. Schwartz, B.L., Light-Wahl, K.J. & Smith, R.D. (1994) *Journal of the American Society for Mass Spectrometry* **5**, 201-204.

185. Green, B.N., Bordoli, R.S., Hanin, L.G., Lallier, F.H., Toulmond, A. & Vinogradov, S.N. (1999) *Journal of Biological Chemistry* **274**, 28206-28212.

186. Wang, Y., Schubert, M., Ingendoh, A. & Franzen, J. (2000) *Rapid Communications in Mass Spectrometry* **14**, 12-17.

187. Guo, M.Q., Zhang, S.Q., Song, F.R., Wang, D.W., Liu, Z.Q. & Liu, S.Y. (2003) *Journal of Mass Spectrometry* **38**, 723-731.

188. Reid, G.E., O'Hair, R.A.J., Styles, M.L., McFadyen, W.D. & Simpson, R.J. (1998) *Rapid Communications in Mass Spectrometry* **12**, 1701-1708.

189. Kitova, E.N., Kitov, P.I., Bundle, D.R. & Klassen, J.S. (2001) *Glycobiology* **11**, 605-611.

190. SannesLowery, K.A., Hu, P.F., Mack, D.P., Mei, H.Y. & Loo, J.A. (1997) *Analytical Chemistry* **69**, 5130-5135.

191. Kuchumov, A.R., Loo, J.A. & Vinogradov, S.N. (2000) *Journal of Protein Chemistry* **19**, 139-149.

192. Guilhaus, M., Selby, D. & Mlynski, V. (2000) *Mass Spectrometry Reviews* **19**, 65-107.

193. Morris, H.R., Paxton, T., Dell, A., Langhorne, J., Berg, M., Bordoli, R.S., Hoyes, J. & Bateman, R.H. (1996) *Rapid Communications in Mass Spectrometry* **10**, 889-896.

194. Krutchinsky, A.N., Chernushevich, I.V., Spicer, V.L., Ens, W. & Standing, K.G. (1998) *Journal of the American Society for Mass Spectrometry* **9**, 569-579.

195. Van Berkel, W.J.H., Van Den Heuvel, R.H.H., Versluis, C. & Heck, A.J.R. (2000) *Protein Science* **9**, 435-439.

196. Tito, M.A., Tars, K., Valegard, K., Hajdu, J. & Robinson, C.V. (2000) *Journal of the American Chemical Society* **122**, 3550-3551.

197. Rostom, A.A. & Robinson, C.V. (1999) *Journal of the American Chemical Society* **121**, 4718-4719.

198. Tahallah, N., Pinkse, M., Maier, C.S. & Heck, A.J.R. (2001) *Rapid Communications in Mass Spectrometry* **15**, 596-601.

199. Chernushevich, I.V. & Thomson, B.A. (2004) *Analytical Chemistry* **76**, 1754-1760.

200. Sobott, F., Hernandez, H., McCammon, M.G., Tito, M.A. & Robinson, C.V. (2002) *Analytical Chemistry* **74**, 1402-1407.

201. Lei, Q.P., Cui, X.Y., Kurtz, D.M., Amster, I.J., Chernushevich, I.V. & Standing, K.G. (1998) *Analytical Chemistry* **70**, 1838-1846.

202. Gale, D.C., Goodlett, D.R., Lightwahl, K.J. & Smith, R.D. (1994) *Journal of the American Chemical Society* **116**, 6027-6028.

203. Gabelica, V., Rosu, F., Houssier, C. & De Pauw, E. (2000) *Rapid Communications in Mass Spectrometry* **14**, 464-467.

204. Loo, R.R.O., Goodlett, D.R., Smith, R.D. & Loo, J.A. (1993) *Journal of the American Chemical Society* **115**, 4391-4392.

205. Loo, J.A., Hu, P.F., McConnell, P., Mueller, W.T., Sawyer, T.K. & Thanabal, V. (1997) *Journal of the American Society for Mass Spectrometry* **8**, 234-243.

206. De Vriendt, K., Sandra, K., Desmet, T., Nerinckx, W., Van Beeumen, J. & Devreese, B. (2004) *Rapid Communications in Mass Spectrometry* **18**, 3061-3067.

207. Zhang, S., Van Pelt, C.K. & Wilson, D.B. (2003) *Analytical Chemistry* **75**, 3010-3018.

208. Ding, J.M. & Anderegg, R.J. (1995) *Journal of the American Society for Mass Spectrometry* **6**, 159-164.

209. Cunniff, J.B. & Vouros, P. (1995) *Journal of the American Society for Mass Spectrometry* **6**, 437-447.

210. Aplin, R.T., Robinson, C.V., Schofield, C.J. & Westwood, N.J. (1994) *Journal of the Chemical Society-Chemical Communications*, 2415-2417.

211. Gabelica, V., De Pauw, E. & Rosu, F. (1999) *Journal of Mass Spectrometry* **34**, 1328-1337.

212. Wu, Q.Y., Gao, J.M., JosephMcCarthy, D., Sigal, G.B., Bruce, J.E., Whitesides, G.M. & Smith, R.D. (1997) *Journal of the American Chemical Society* **119**, 1157-1158.

213. Gupta, R., Hamdan, S.M., Dixon, N.E., Sheil, M.M. & Beck, J.L. (2004) *Protein Science* **13**, 2878-2887.

214. Lafitte, D., Capony, J.P., Grassy, G., Haiech, J. & Calas, B. (1995) *Biochemistry* **34**, 13825-13832.

215. Feng, R. (1995) *Proccedings of the 43rd American Soceity for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics,* Alanta, 1264.

216. Milos, M., Schaer, J.J., Comte, M. & Cox, J.A. (1986) *Biochemistry* **25**, 6279-6287.

217. Milos, M., Comte, M., Schaer, J.J. & Cox, J.A. (1989) *Journal of Inorganic Biochemistry* **36**, 11-25.

218. Lim, H.K., Hsieh, Y.L., Ganem, B. & Henion, J. (1995) *Journal of Mass Spectrometry* **30**, 708-714.

219. Greig, M.J., Gaus, H., Cummins, L.L., Sasmor, H. & Griffey, R.H. (1995) *Journal of the American Chemical Society* **117**, 10765-10766.

220. Sannes-Lowery, K.A., Griffey, R.H. & Hofstadler, S.A. (2000) *Analytical Biochemistry* **280**, 264-271.

221. Daniel, J.M., McCombie, G., Wendt, S. & Zenobi, R. (2003) *Journal of the American Society for Mass Spectrometry* **14**, 442-448.

222. Dotsikas, Y. & Loukas, Y.L. (2003) *Journal of the American Society for Mass Spectrometry* **14**, 1123-1129.

223. Smith, R.D. & Lightwahl, K.J. (1993) *Biological Mass Spectrometry* **22**, 493-501.

224. Fabris, D. & Fenselau, C. (1999) *Analytical Chemistry* **71**, 384-387.

225. Nettleton, E.J., Tito, P., Sunde, M., Bouchard, M., Dobson, C.M. & Robinson, C.V. (2000) *Biophysical Journal* **79**, 1053-1065.

226. Rostom, A.A., Sunde, M., Richardson, S.J., Schreiber, G., Jarvis, S., Bateman, R., Dobson, C.M. & Robinson, C.V. (1998) *Proteins-Structure Function and Genetics*, 3-11.

227. Hernandez, H. & Robinson, C.V. (2001) *Journal of Biological Chemistry* **276**, 46685-46688.

228. Tahallah, N., van den Heuvel, R.H.H., van den Berg, W.A.M., Maier, C.S., van Berkel, W.J.H. & Heck, A.J.R. (2002) *Journal of Biological Chemistry* **277**, 36425-36432.

229. Sirko, A. & Brodzik, R. (2000) *Acta Biochimica Polonica* **47**, 1189-1195.

230. Ha, N.C., Oh, S.T., Sung, J.Y., Cha, K.A., Lee, M.H. & Oh, B.H. (2001) *Nature Structural Biology* **8**, 505-509.

231. Pinkse, M.W.H., Maier, C.S., Kim, J.I., Oh, B.H. & Heck, A.J.R. (2003) *Journal of Mass Spectrometry* **38**, 315-320.

232. Nettleton, E.J., Sunde, M., Lai, Z.H., Kelly, J.W., Dobson, C.M. & Robinson, C.V. (1998) *Journal of Molecular Biology* **281**, 553-564.

233. McCammon, M.G., Scott, D.J., Keetch, C.A., Greene, L.H., Purkey, H.E., Petrassi, H.M., Kelly, J.W. & Robinson, C.V. (2002) *Structure* **10**, 851-863.

234. Sobott, F., Benesch, J.L.P., Vierling, E. & Robinson, C.V. (2002) *Journal of Biological Chemistry* **277**, 38921-38929.

235. Kelly, J.W. (1997) *Structure* **5**, 595-600.

236. Koo, E.H., Lansbury, P.T. & Kelly, J.W. (1999) *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9989-9990.

237. McCutchen, S.L., Colon, W. & Kelly, J.W. (1993) *Biochemistry* **32**, 12119-12127.

238. Peterson, S.A., Klabunde, T., Lashuel, H.A., Purkey, H., Sacchettini, J.C. & Kelly, J.W. (1998) *Proceedings of the National Academy of Sciences of the United States of America* **95**, 12956-12960.

239. Benesch, J.L.P., Sobott, F. & Robinson, C.V. (2003) *Analytical Chemistry* **75**, 2208-2214.

240. Carver, J.A., Lindner, R.A., Lyon, C., Canet, D., Hernandez, H., Dobson, C.M. & Redfield, C. (2002) *Journal of Molecular Biology* **318**, 815-827.

241. Williams, N.K., Prosselkov, P., Liepinsh, E., Line, I., Sharipo, A., Littler, D.R., Curmi, P.M.G., Otting, G. & Dixon, N.E. (2002) *Journal of Biological Chemistry* **277**, 7790-7798.

242. Crouch, T.H. & Klee, C.B. (1980) *Biochemistry* **19**, 3692-3698.

243. Galletto, R., Jezewska, M.J. & Bujalowski, W. (2003) *Journal of Molecular Biology* **329**, 441-465.

244. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000) *Nucleic Acids Research* **28**, 235-242.

245. Martinezbueno, M., Maqueda, M., Galvez, A., Samyn, B., Vanbeeumen, J., Coyette, J. & Valdivia, E. (1994) *Journal of Bacteriology* **176**, 6334-6339.

246. Saether, O., Craik, D.J., Campbell, I.D., Sletten, K., Juul, J. & Norman, D.G. (1995) *Biochemistry* **34**, 4147-4158.

247. Blond, A., Peduzzi, J., Goulard, C., Chiuchiolo, M.J., Barthelemy, M., Prigent, Y., Salomon, R.A., Farias, R.N., Moreno, F. & Rebuffat, S. (1999) *European Journal of Biochemistry* **259**, 747-755.

248. Luckett, S., Garcia, R.S., Barker, J.J., Konarev, A.V., Shewry, P.R., Clarke, A.R. & Brady, R.L. (1999) *Journal of Molecular Biology* **290**, 525-533.

249. Tang, Y.Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C.J., Ouellette, A.J. & Selsted, M.E. (1999) *Science* **286**, 498-502.

250. Craik, D.J., Daly, N.L., Bond, T. & Waine, C. (1999) *Journal of Molecular Biology* **294**, 1327-1336.

251. Hernandez, J.F., Gagnon, J., Chiche, L., Nguyen, T.M., Andrieu, J.P., Heitz, A., Hong, T.T., Pham, T.T.C. & Nguyen, D.L. (2000) *Biochemistry* **39**, 5722-5730.

252. Kessler, H. (1982) *Chimia* **36**, 248-248.

253. Hruby, V.J., Alobeidi, F. & Kazmierski, W. (1990) *Biochemical Journal* **268**, 249-262.

254. Satoh, T., Li, S., Friedman, T.M., Wiaderkiewicz, R., Korngold, R. & Huang, Z.W. (1996) *Biochemical and Biophysical Research Communications* **224**, 438-443.

255. Scott, C.P., Abel-Santos, E., Wall, M., Wahnon, D.C. & Benkovic, S.J. (1999) *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13638-13643.

256. Goldenberg, D.P. & Creighton, T.E. (1983) *Journal of Molecular Biology* **165**, 407-413.

257. Cotton, G.J. & Muir, T.W. (1999) *Chemistry & Biology* **6**, R247-R256.

258. Erlanson, D.A., Chytil, M. & Verdine, G.L. (1996) *Chemistry & Biology* **3**, 981-991.

259. Wilken, J. & Kent, S.B.H. (1998) *Current Opinion in Biotechnology* **9**, 412-426.

260. Paulus, H. (1998) *Chemical Society Reviews* **27**, 375-386.

261. Paulus, H. (2000) *Annual Review of Biochemistry* **69**, 447-496.

262. Trabi, M. & Craik, D.J. (2002) *Trends in Biochemical Sciences* **27**, 132-138.

263. Mills, K.V., Lew, B.M., Jiang, S.Q. & Paulus, H. (1998) *Proceedings of the National Academy of Sciences of the United States of America* **95**, 3543-3548.

264. Shingledecker, K., Jiang, S.Q. & Paulus, H. (1998) *Gene* **207**, 187-195.

265. Wu, H., Xu, M.Q. & Liu, X.Q. (1998) *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1387**, 422-432.

266. Iwai, H., Lingel, A. & Pluckthun, A. (2001) *Journal of Biological Chemistry* **276**, 16548-16554.

267. Iwai, H. & Pluckthun, A. (1999) *FEBS Letters* **459**, 166-172.

268. Zhou, H.X. (2003) *Journal of Molecular Biology* **332**, 257-264.

269. Deechongkit, S. & Kelly, J.W. (2002) *Journal of the American Chemical Society* **124**, 4980-4986.

270. Williams, N.K., Liepinsh, E., Watt, S.J., Prosselkov, P., Matthews, J.M., Attard, P., Beck, J.L., Dixon, N.E. & Otting, G. (2005) *Journal of Molecular Biology* **346**, 1095-1108.

271. Nakayama, N., Arai, N., Kaziro, Y. & Arai, K. (1984) *Journal of Biological Chemistry* **259**, 88-96.

272. Nakayama, N., Arai, N., Bond, M.W., Kaziro, Y. & Arai, K. (1984) *Journal of Biological Chemistry* **259**, 97-101.

273. Fass, D., Bogden, C.E. & Berger, J.M. (1999) *Structure* **7**, 691-698.

274. Weigelt, J., Brown, S.E., Miles, C.S., Dixon, N.E. & Otting, G. (1999) *Structure* **7**, 681-690.

275. Miles, C.S., Weigelt, J., Patrick, N., Stamford, J., Dammerova, N., Otting, G. & Dixon, N.E. (1997) *Biochemical and Biophysical Research Communications* **231**, 126-130.

276. Roder, H., Wagner, G. & Wuthrich, K. (1985) *Biochemistry* **24**, 7396-7407.

277. Ferraro, D.M. & Robertson, A.D. (2004) *Biochemistry* **43**, 587-594.

278. Englander, J.J., Rogero, J.R. & Englander, S.W. (1985) *Analytical Biochemistry* **147**, 234-244.

279. Nemirovskiy, O., Giblin, D.E. & Gross, M.L. (1999) *Journal of the American Society for Mass Spectrometry* **10**, 711-718.

280. Zhu, M.M., Rempel, D.L., Zhao, J., Giblin, D.E. & Gross, M.L. (2003) *Biochemistry* **42**, 15388-15397.

281. Simmons, D.A., Dunn, S.D. & Konermann, L. (2003) *Biochemistry* **42**, 5896-5905.

282. Expasy (2003) *Compute pI/MW*, http://au.expasy.org/tools/pi_tool.html, Vol. 2005

283. Katta, V. & Chait, B.T. (1993) *Journal of the American Chemical Society* **115**, 6317-6321.

284. Simmons, D.A. & Konermann, L. (2002) *Biochemistry* **41**, 1906-1914.

285. Kim, M.Y., Maier, C.S., Reed, D.J. & Deinzer, M.L. (2002) *Protein Science* **11**, 1320-1329.

286. Yi, Q. & Baker, D. (1996) *Protein Science* **5**, 1060-1066.

287. Bai, Y.W., Milne, J.S., Mayne, L. & Englander, S.W. (1994) *Proteins-Structure Function and Genetics* **20**, 4-14.

288. Connelly, G.P., Bai, Y., Jeng, M.F. & Englander, S.W. (1993) *Proteins: Structure, Function, and Genetics* **17**, 87-92.

289. Maier, C.S., Schimerlik, M.I. & Deinzer, M.L. (1999) *Biochemistry* **38**, 1136-1143.

290. Arrington, C.B., Teesch, L.M. & Robertson, A.D. (1999) *Journal of Molecular Biology* **285**, 1265-1275.

291. Myers, J.K. & Oas, T.G. (2001) *Nature Structural Biology* **8**, 552-558.

292. Camarero, J.A., Fushman, D., Sato, S., Giriat, I., Cowburn, D., Raleigh, D.P. & Muir, T.W. (2001) *Journal of Molecular Biology* **308**, 1045-1062.

293. Siebold, C. & Erni, B. (2002) *Biophysical Chemistry* **96**, 163-171.

294. Arrington, C.B. & Robertson, A.D. (2000) *Journal of Molecular Biology* **300**, 221-232.

295. Loo, J.A., Edmonds, C.G., Udseth, H.R. & Smith, R.D. (1990) *Analytical Chemistry* **62**, 693-698.

296. Smith, R.D., Loo, J.A., Loo, R.R.O., Busman, M. & Udseth, H.R. (1991) *Mass Spectrometry Reviews* **10**, 359-451.

297. Grandori, R. (2002) *Protein Science* **11**, 453-458.

298. Mirza, U.A., Cohen, S.L. & Chait, B.T. (1993) *Analytical Chemistry* **65**, 1-6.

299. Sogbein, O.O., Simmons, D.A. & Konermann, L. (2000) *Journal of the American Society for Mass Spectrometry* **11**, 312-319.

300. Iavarone, A.T., Jurchen, J.C. & Williams, E.R. (2000) *Journal of the American Society for Mass Spectrometry* **11**, 976-985.

301. Amad, M.H., Cech, N.B., Jackson, G.S. & Enke, C.G. (2000) *Journal Mass Spectrometry* **35**, 784-789.

302. Halgand, F. & Laprevote, O. (2001) *European Journal of Mass Spectrometry* **7**, 433-439.

303. Verkerk, U.H., Peschke, M. & Kebarle, P. (2003) *Journal of Mass Spectrometry* **38**, 618-631.

304. Mirza, U.A. & Chait, B.T. (1994) *Analytical Chemistry* **66**, 2898-2904.

305. Leblanc, J.C.Y., Wang, J.Y., Guevremont, R. & Siu, K.W.M. (1994) *Organic Mass Spectrometry* **29**, 587-593.

306. Mansoori, B.A., Volmer, D.A. & Boyd, R.K. (1997) *Rapid Communications in Mass Spectrometry* **11**, 1120-1130.

307. Zhou, S.L. & Cook, K.D. (2000) *Journal of the American Society for Mass Spectrometry* **11**, 961-966.

308. Gross, D.S. & Williams, E.R. (1995) *Journal of the American Chemical Society* **117**, 883-890.

309. Schnier, P.D., Gross, D.S. & Williams, E.R. (1995) *Journal of the American Society for Mass Spectrometry* **6**, 1086-1097.

310. Samalikova, M., Matecko, I., Muller, N. & Grandori, R. (2004) *Analytical and Bioanalytical Chemistry* **378**, 1112-1123.

311. Leblanc, J.C.Y., Beuchemin, D., Siu, K.W.M., Guevremont, R. & Berman, S.S. (1991) *Organic Mass Spectrometry* **26**, 831-839.

312. Fligge, T.A., Przybylski, M., Quinn, J.P. & Marshall, A.G. (1998) *European Mass Spectrometry* **4**, 401-404.

313. Xu, Q. & Keiderling, T.A. (2004) *Biopolymers* **73**, 716-726.

314. Hu, P., Ye, Q.-Z. & Loo, J.A. (1994) *Analytical Chemistry* **66**, 4190-4194.

315. Kebarle, P. & Ho, Y. (1997) in *Electrospray Ionisation Mass Spectrometry: Fundamentals, Instrumentation and Applications,* Ed. Cole R.D., John Wiley & Sons, Inc, New York, Chapter 1, 3-65.

316. Hu, P. & Loo, J.A. (1995) *Journal of Mass Spectrometry* **30**, 1076-1082.

317. Winger, B.E., Lightwahl, K.J., Loo, R.R.O., Udseth, H.R. & Smith, R.D. (1993) *Journal of the American Society for Mass Spectromery.* **4**, 536-545.

318. Permyakov, E.A. & Berliner, L.J. (2000) *FEBS Letters* **473**, 269-274.

319. Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S. & Murai, N. (1980) *Biochemical and Biophysical Research Communications* **95**, 1098-1104.

320. Kuwajima, K., Ikeguchi, M., Sugawara, T., Hiraoka, Y. & Sugai, S. (1990) *Biochemistry* **29**, 8240-8249.

321. Ewbank, J.J. & Creighton, T.E. (1993) *Biochemistry* **32**, 3694-3707.

322. Pike, A.C.W., Brew, K. & Acharya, K.R. (1996) *Structure* **4**, 691-703.

323. Kapur, A., Beck, J.L., Brown, S.E., Dixon, N.E. & Sheil, M.M. (2002) *Protein Science* **11**, 147-157.

324. Gross, D.S., Rodriguez-Cruz, S.E., Bock, S. & Williams, E.R. (1995) *Journal of Physical Chemistry* **99**, 4034-4038.

325. Schnier, P.D., Gross, D.S. & Williams, E.R. (1995) *Journal of the American Chemical Society* **117**, 6747-6757.

326. Williams, E.R. (1996) *Journal of Mass Spectrometry* **31**, 831-842.

327. Lavarone, A.T. & Williams, E.R. (2003) *Journal of the American Chemical Society* **125**, 2319-2327.

328. Blades, A.T., Klassen, J.S. & Kebarle, P. (1996) *Journal of the American Chemical Society* **118**, 12437-12442.

329. Rayleigh, L. (1882) *Philosophical Magazine* **14**, 184-186.

330. Kelly, M.A., Vestling, M.M., Fenselau, C.C. & Smith, P.B. (1992) *Organic Mass Spectrometry* **27**, 1143-1147.

331. Torok, K. (2002) *Biochemical Society Transactions* **30**, 55-61.

332. Gao, J., Yin, D.H., Yao, Y.H., Sun, H.Y., Qin, Z.H., Schoneich, C., Williams, T.D. & Squier, T.C. (1998) *Biophysical Journal* **74**, 1115-1134.

333. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B. & Bax, A. (1995) *Nature Structural Biology* **2**, 768-776.

334. Veenstra, T.D., Tomlinson, A.J., Benson, L., Kumar, R. & Naylor, S. (1998) *Journal of the American Society for Mass Spectrometry* **9**, 580-584.

335. Lafitte, D., Heck, A.J.R., Hill, T.J., Jumel, K., Harding, S.E. & Derrick, P.J. (1999) *European Journal of Biochemistry* **261**, 337-344.

336. Hill, T.J., Lafitte, D., Wallace, J.I., Cooper, H.J., Tsvetkov, P.O. & Derrick, P.J. (2000) *Biochemistry* **39**, 7284-7290.

337. Weiss, B., Prozialeck, W.C. & Wallace, T.L. (1982) *Biochemical Pharmacology* **31**, 2217-2226.

338. Babu, Y.S., Bugg, C.E. & Cook, W.J. (1988) *Journal of Molecular Biology* **204**, 191-204.

339. Vetter, S.W. & Leclerc, E. (2003) *European Journal of Biochemistry* **270**, 404-414.

340. Oneil, K.T. & Degrado, W.F. (1990) *Trends in Biochemical Sciences* **15**, 59-64.

341. Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. & Bax, A. (1992) *Science* **256**, 632-638.

342. Meador, W.E., Means, A.R. & Quiocho, F.A. (1992) *Science* **257**, 1251-1255.

343. Vandonselaar, M., Hickie, R.A., Quail, J.W. & Delbaere, L.T. (1994) *Nature Structural Biology* **1**, 795-801.

344. Cook, W.J., Walter, L.J. & Walter, M.R. (1994) *Biochemistry* **33**, 15259-15265.

345. Nemirovskiy, O.V., Ramanathan, R. & Gross, M.L. (1997) *Journal of the American Society for Mass Spectrometry* **8**, 809-812.

346. Nousiainen, M., Vainiotalo, P., Feng, X.D. & Derrick, P.J. (2001) *European Journal of Mass Spectrometry* **7**, 393-398.

347. Nousiainen, M., Derrick, P.J., Lafitte, D. & Vainiotalo, P. (2003) *Biophysical Journal* **85**, 491-500.

348. Chattopadhyaya, R., Meador, W.E., Means, A.R. & Quiocho, F.A. (1992) *Journal of Molecular Biology* **228**, 1177-1192.

349. Seamon, K.B. (1980) *Biochemistry* **19**, 207-215.

350. Levin, R.M. & Weiss, B. (1977) *Molecular Pharmacology* **13**, 690-697.

351. Jackson, A.E. & Puett, D. (1986) *Biochemical Pharmacology* **35**, 4395-4400.

352. Massom, L., Lee, H. & Jarrett, H.W. (1990) *Biochemistry* **29**, 671-681.

353. Comte, M., Maulet, Y. & Cox, J.A. (1983) *Biochemical Journal* **209**, 269-272.

354. Badman, E.R., Hoaglund-Hyzer, C.S. & Clemmer, D.E. (2001) *Analytical Chemistry* **73**, 6000-6007.

355. Myung, S., Badman, E.R., Lee, Y.J. & Clemmer, D.E. (2002) *Journal of Physical Chemistry A* **106**, 9976-9982.

356. Schaeffer, P.M., Headlam, M.J. & Dixon, N.E. (2005) *Iubmb Life* **57**, 5-12.

357. McHenry, C.S. (2003) *Molecular Microbiology* **49**, 1157-1165.

358. McHenry, C.S. (1988) *Annual Review of Biochemistry* **57**, 519-550.

359. Messer, W., Blaesing, F., Jakimowicz, D., Krause, M., Majka, J., Nardmann, J., Schaper, S., Seitz, H., Speck, C., Weigel, C., Wegrzyn, G., Welzeck, M. & Zakrzewska-Czerwinska, J. (2001) *Biochimie* **83**, 5-12.

360. Messer, W. (2002) *FEMS Microbiology Reviews* **26**, 355-374.

361. Frick, D.N. & Richardson, C.C. (2001) *Annual Review of Biochemistry* **70**, 39-80.

362. Johanson, K.O. & McHenry, C.S. (1984) *Journal of Biological Chemistry* **259**, 4589-4595.

363. Dallmann, H.G., Thimmig, R.L. & McHenry, C.S. (1995) *Journal of Biological Chemistry* **270**, 29555-29562.

364. Jeruzalmi, D., O'Donnell, M. & Kuriyan, J. (2002) *Current Opinion in Structural Biology* **12**, 217-224.

365. Ogawa, T. & Okazaki, T. (1980) *Annual Review of Biochemistry* **49**, 421-457.

366. Steitz, T.A., Smerdon, S.J., Jager, J. & Joyce, C.M. (1994) *Science* **266**, 2022-2025.

367. Gao, D.X. & McHenry, C.S. (2001) *Journal of Biological Chemistry* **276**, 4433-4440.

368. De Saro, F.J.L., Georgescu, R.E. & O'Donnell, M. (2003) *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14689-14694.

369. Gao, D.X. & McHenry, C.S. (2001) *Journal of Biological Chemistry* **276**, 4441-4446.

370. Kong, X.P., Onrust, R., Odonnell, M. & Kuriyan, J. (1992) *Cell* **69**, 425-437.

371. Waksman, G., Lanka, E. & Carazo, J.M. (2000) *Nature Structural Biology* **7**, 20-22.

372. Patel, S.S. & Picha, K.M. (2000) *Annual Review of Biochemistry* **69**, 651-697.

373. Caruthers, J.M. & McKay, D.B. (2002) *Current Opinion in Structural Biology* **12**, 123-133.

374. Bujalowski, W. (2003) *Trends in Biochemical Sciences* **28**, 116-118.

375. von Hippel, P.H. & Delagoutte, E. (2003) *Bioessays* **25**, 1168-1177.

376. Ilyina, T.V., Gorbalenya, A.E. & Koonin, E.V. (1992) *Journal of Molecular Evolution* **34**, 351-357.

377. Gorbalenya, A.E. & Koonin, E.V. (1993) *Current Opinion in Structural Biology* **3**, 419-429.

378. Singleton, M.R., Sawaya, M.R., Ellenberger, T. & Wigley, D.B. (2000) *Cell* **101**, 589-600.

379. Niedenzu, T., Roleke, D., Bains, G., Scherzinger, E. & Saenger, W. (2001) *Journal of Molecular Biology* **306**, 479-487.

380. Von Hippel, P.H. & Delagouette, E. (2001) *Cell* **104**, 177-190.

381. Bujalowski, W., Klonowska, M.M. & Jezewska, M.J. (1994) *Journal of Biological Chemistry* **269**, 31350-31358.

382. Biswas, E.E., Biswas, S.B. & Bishop, J.E. (1986) *Biochemistry* **25**, 7368-7374.

383. Bujalowski, W. & Klonowska, M.M. (1993) *Biochemistry* **32**, 5888-5900.

384. Sanmartin, M.C., Stamford, N.P.J., Dammerova, N., Dixon, N.E. & Carazo, J.M. (1995) *Journal of Structural Biology* **114**, 167-176.

385. Yu, X., Jezewska, M.J., Bujalowski, W. & Egelman, E.H. (1996) *Journal of Molecular Biology* **259**, 7-14.

386. San Martin, C., Radermacher, M., Wolpensinger, B., Engel, A., Miles, C.S., Dixon, N.E. & Carazo, J.M. (1998) *Structure* **6**, 501-509.

387. Donate, L.E., Llorca, O., Barcena, M., Brown, S.E., Dixon, N.E. & Carazo, J.M. (2000) *Journal of Molecular Biology* **303**, 383-393.

388. Barcena, M., San Martin, C., Weise, F., Ayora, S., Alonso, J.C. & Carazo, J.M. (1998) *Journal of Molecular Biology* **283**, 809-819.

389. Scherzinger, E., Ziegelin, G., Barcena, M., Carazo, J.M., Lurz, R. & Lanka, E. (1997) *Journal of Biological Chemistry* **272**, 30228-30236.

390. Barcena, M., Ruiz, T., Donate, L.E., Brown, S.E., Dixon, N.E., Radermacher, W. & Carazo, J.M. (2001) *EMBO Journal* **20**, 1462-1468.

391. Galletto, R. & Bujalowski, W. (2002) *Biochemistry* **41**, 8907-8920.

392. Galletto, R. & Bujalowski, W. (2002) *Biochemistry* **41**, 8921-8934.

393. Davey, M.J., Fang, L.H., McInerney, P., Georgescu, R.E. & O'Donnell, M. (2002) *EMBO Journal* **21**, 3148-3159.

394. Arai, K. & Kornberg, A. (1981) *Journal of Biological Chemistry* **256**, 5253-5259.

395. Toth, E.A., Li, Y., Sawaya, M.R., Cheng, Y.F. & Ellenberger, T. (2003) *Molecular Cell* **12**, 1113-1123.

396. Stasiak, A.Z., Larquet, E., Stasiak, A., Muller, S., Engel, A., Van Dyck, E., West, S.C. & Egelman, E.H. (2000) *Current Biology* **10**, 337-340.

397. Miyata, T., Yamada, K., Iwasaki, H., Shinagawa, H., Morikawa, K. & Mayanagi, K. (2000) *Journal of Structural Biology* **131**, 83-89.

398. Yu, X., VanLoock, M.S., Poplawski, A., Kelman, Z., Xiang, T., Tye, B.K. & Egelman, E.H. (2002) *EMBO Reports* **3**, 792-797.

399. Kaplan, D.L. & O'Donnell, M. (2002) *Molecular Cell* **10**, 647-657.

400. Versluis, C. & Heck, A.J.R. (2001) *International Journal of Mass Spectrometry* **210**, 637-649.

401. Nesatyy, V.J. (2001) *Journal of Mass Spectrometry* **36**, 950-959.

402. Aquilina, A.J., Benesch, J.L.P., Bateman, O.A., Slingsby, C. & Robinson, C.V. (2003) *Proceedings of the National Academy of Sciences of the United States of America* **100**, 10611-10616.

403. Jurchen, J.C. & Williams, E.R. (2003) *Journal of the American Chemical Society* **125**, 2817-2826.

404. Felitsyn, N., Kitova, E.N. & Klassen, J.S. (2001) *Analytical Chemistry* **73**, 4647-4661.

405. Yin, S., Xie, Y. & Loo, J.A. (2005) *Proceedings of the 53rd American Soceity for Mass Spectrometry conference,* San Antonio, 23.

406. Wickner, S. & Hurwitz, J. (1975) *Proceedings of the National Academy of Sciences of the United States of America* **72**, 921-925.

407. Kobori, J.A. & Kornberg, A. (1982) *Journal of Biological Chemistry* **257**, 3770-3775.

408. Ludlam, A.V., McNatt, M.W., Carr, K.M. & Kaguni, J.M. (2001) *Journal of Biological Chemistry* **276**, 27345-27353.

409. Funnell, B.E., Baker, T.A. & Kornberg, A. (1987) *Journal of Biological Chemistry* **262**, 10327-10334.

410. Wahle, E., Lasken, R.S. & Kornberg, A. (1989) *Journal of Biological Chemistry* **264**, 2463-2468.

411. Felitsyn, N., Kitova, E.N. & Klassen, J.S. (2002) *Journal of the American Society for Mass Spectrometry* **13**, 1432-1442.

412. Koonin, E.V. (1992) *Nucleic Acids Research* **20**, 1997.

413. Biswas, S.B. & Biswas, E.E. (1987) *Journal of Biological Chemistry* **262**, 7831-7838.

414. Galletto, R., Rajendran, S. & Bujalowski, W. (2000) *Biochemistry* **39**, 12959-12969.

415. Galletto, R., Maillard, R., Jezewska, M.J. & Bujalowski, W. (2004) *Biochemistry* **43**, 10988-11001.

416. Clemmer, D.E. & Jarrold, M.F. (1997) *Journal of Mass Spectrometry* **32**, 577-592.

417. Suckau, D., Shi, Y., Beu, S.C., Senko, M.W., Quinn, J.P., Wampler, F.M. & McLafferty, F.W. (1993) *Proceedings of the National Academy of Sciences of the United States of America* **90**, 790-793.

Appendix

Table A.1 Extent of reduction of disulfide bonds of ribonuclease A following treatment with 10 mM NH_4OAc pH 7.1, 1 mM DTT and 30% CH_3CN.[Φ]

	%				
Time (min)	**0 cysteines reduced**	**2 cysteines reduced**	**4 cysteines reduced**	**6 cysteines reduced**	**8 cysteines reduced**
0	100	0	0	0	0
5	7.7	26.1	38.4	22.5	5.3
15	0	2.4	10.7	34.4	52.4
30	0	0	0	3.4	96.6
45	0	0	0	5.7	94.3
60	0	0	0	7.3	92.7

[Φ] The percentage of each reduced form was calculated by comparing the abundance of ions from a reduced state with the total abundace of ions from ribonuclease A.

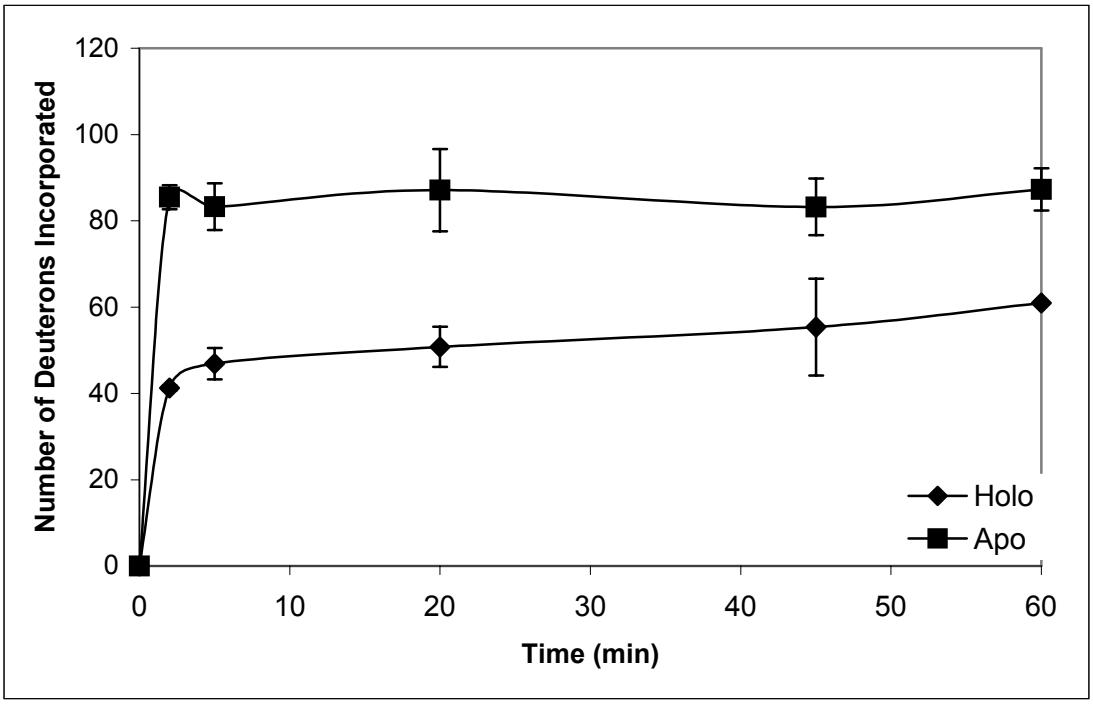


Figure A.1 The hydrogen deuterium exchange results for α-lactalbumin when in 10 mM NH_4OAc pH 7.1 (holo) and following incubation in 10 mM NH_4OAc pH 7.1 and 5.6 mM DTT for 1 hour (apo). For these experiments, α-lactalbumin and DTT were 200 μM and 0.112 M, respectively in 10 mM NH_4OAc pH 7.1. Exchange of amide protons of holo-α-lactalbumin was carried out by mixing 10 μL of α-lactalbumin at 200 μM with 10 μL of 10 mM NH_4OAc (pH 7.1) then adding 180 μL of the D_2O solution (prepared as described in 2.4.1) adjusted to pH 7.1 with dilute NH_4OH in water. For apo samples 10 μL aliquots of α-lactalbumin at 200 μM and DTT (0.112 M) were mixed and incubated at room temperature for 1 hour before exchange was commenced as described above. Exchange was quenched using the quenching method described in section 2.4.2.

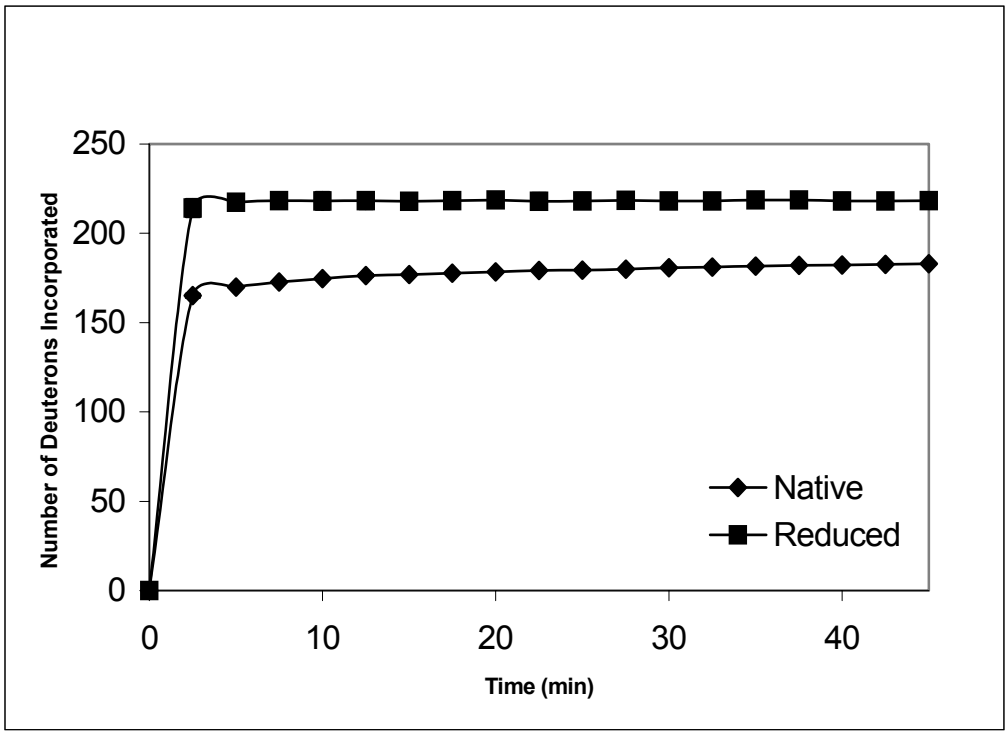


Figure A.2 The hydrogen deuterium exchange results of ribonuclease A in 10 mM NH_4OAc pH 7.1 (native) and following incubation in 10 mM NH_4OAc pH 7.1, 1 mM DTT and 30% CH_3CN for 45 minutes (reduced). Spectra were acquired by the direct injection method where stocks of 200 µM ribonuclease A, 0.1 M DTT in 10 mM NH_4OAc, pH 7.1 and 100% CH_3CN were prepared. Two deuterium oxide solutions were also prepared, one containing only 10 mM NH_4OAc at pH 7.1 in D_2O (native D_2O solution) and a second solution containing 10 mM NH_4OAc at pH 7.1 in D_2O and 30% CD_3CN (reducing D_2O Solution). For the native sample, 25 µL of ribonuclease A at 200 µM was diluted to 100 µM with 25 µM of 10 mM NH_4OAc prior to addition of 450 µL of the native D_2O solution. This solution was then injected into the mass spectrometer at room temperature. The reduced sample was first reduced by mixing 25 µL of the ribonuclease stock with 15 µL of CH_3CN, 5 µL of the DTT (0.1 M) and 5 µL of 10 mM NH_4OAc at pH 7.1 to produce a solution of 100 µM ribonuclease A, 30% CH_3CN and 0.01 M DTT solution in 10 mM NH_4OAc. This solution was than heated at 45 ^oC for 45 minutes. Once cooled to room temperature, 450 µL of the reducing D_2O solution was then added and this solution was injected into the mass spectrometer at room temperature.