Mass spectrometric studies of non-covalent biomolecular complexes

Thitima Urathamakul
University of Wollongong

Mass Spectrometric Studies of Non-Covalent Biomolecular Complexes

A thesis submitted in (partial) fulfilment of the requirements
for the award of the degree

Doctor of Philosophy

from

University of Wollongong

by

Thitima Urathamakul
Bachelor of Science (Honours)

Department of Chemistry

October 2006

DECLARATION

I, Thitima Urathamakul, 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 work has not been submitted for qualification at any other academic institution.

Thitima Urathamakul

24th October 2006

ACKNOWLEDGEMENTS

While this thesis is a culmination of three years' worth of work and study, my contribution in the form of its writing is but a small part of the overall process. The following is a list of people who have played an integral part in my life over the past several years – people who have provided me with guidance and support both immeasurable and invaluable. In short, people without whom this thesis would not have been possible.

Firstly, my supervisors Dr Jennifer Beck, Dr Stephen Ralph and Professor Margaret Sheil. Margaret, thank you for giving me the opportunity to complete my postgraduate research here at the University of Wollongong. You have never failed to help and encourage me through difficult times.

Steve, your enthusiasm has been a real driving force that has kept the ruthenium work (and my focus) on track. I have found your energy infectious and your ideas a constant source of inspiration.

Jenny, you have been an amazing supervisor throughout my time working with you. You have always found time to help me both academically and personally, your commitment to your students and your work is tremendous. You are my mentor and my confidant.

The three people mentioned above have been the best supervisors that any student could hope for.

Mr Larry Hick for his knowledgeable advice and assistance with the mass spectrometer. You have always been approachable and helpful to everyone. The lab is a warmer place with your presence. Larry, you are a legend.

Raj Gupta and Stephen Watt for their help in teaching me invaluable skills for the various instruments in the early stages of my degree. David Harman, Karin Maxwell, Stephen Blanksby, Roger Kanitz, Todd Mitchell, Jihan Talib, Karina Gornall, Linda Jessop, Michael Thomas, Jane Deeley, and other past and present members of the Mass Spectrometry group for making this a fun and enjoyable place to work.

Dr Nicholas Dixon and his group (Research School of Chemistry Australian National University) for their kindness in providing the proteins used in this study.

Dr Janice Aldrich-Wright (School of Science, Food and Horticulture, University of Western Sydney, Australia University of Western Sydney) for the ruthenium drugs used for the DNA work.

The Department of Chemistry, around which so much of my life has revolved over the past 8 years. The friendliness, support and guidance I have experienced during my time here have been truly memorable.

My family – my brothers, sister, and most especially my mother for all her support over the years and for giving me the opportunity to study overseas in the first place.

Last but not least, my dearest husband Min for all his endless support and patience, particularly during the tough time of writing up. Min, thank you for believing in me and for always being there for me.

PUBLICATIONS

Beck, J.L., Gupta, R., **Urathamakul, T**., Williamson, N.L.; Sheil, M.M., Aldrich-Wright, J. R. and Ralph, S.F. (2003) Probing DNA Selectivity of Ruthenium Metallointercalators Using ESI Mass Spectrometry. *Chem. Commun.*, **5**, 626-7.

Urathamakul, T., Beck, J.L., Sheil, M.M., Aldrich-Wright, J.R. and Ralph, S.F. (2004) A Mass Spectrometric Investigation of Non-Covalent Interactions Between Ruthenium Complexes and DNA. *Dalton Trans.*, **17**, 2683-2690.

Beck, J.L., **Urathamakul, T**., Watt, S.J., Sheil, M.M., Schaeffer, P.M. and Dixon, N.E. (2006) Proteomic Dissection of DNA Polymerisation. *Expert Rev. Proteomics*, **3**, 197-211.

Watt, S.J., **Urathamakul, T**., Schaeffer, P.M., Sheil, M.M., Dixon, N.E. and Beck, J.L. (2006) Electrospray Ionisation Mass Spectrometry of Oligomers of *E. coli* DnaB Helicase and Mutants. *Rapid Commun. Mass Spectrom.*, **21**, 132-140.

ABSTRACT

Electrospray ionisation mass spectrometry (ESI-MS) was employed to investigate non-covalent associations of macromolecules with ligands, metal ions and other macromolecules. Firstly, ESI-MS was used to examine the interactions of six ruthenium compounds with three different DNA sequences (D1, D2 and D3). The relative binding affinities of these ruthenium compounds towards dsDNA was determined to be: $[Ru(phen)_2(dppz)]^{2+} \geq [Ru(phen)_2(dpqMe_2)]^{2+} > [Ru(phen)_2(dpqC)]^{2+} > [Ru(phen)_2(dpq)]^{2+} > [Ru(phen)_2(pda)]^{2+} > [Ru(phen)_3]^{2+}$. This order was in good agreement with that obtained from DNA melting temperature experiments. Competition experiments involving ruthenium compounds and organic drugs were also conducted to obtain information about the DNA binding modes of the ruthenium compounds. These studies provide strong support for the routine application of ESI-MS as a tool for analysis of non-covalent complexes between metallointercalators and dsDNA.

ESI-MS also proved to be a rapid and efficient tool for investigation of interactions between the N-terminal domain of ε (ε186, the exonuclease proofreading subunit of *E. coli* DNA) and three different metal ions (Mn^{2+}, Zn^{2+} and Dy^{3+}). The dissociation constants (K_d) for binding of Mn^{2+}, Zn^{2+} and Dy^{3+} to ε186 were determined from ESI-MS data to be 38.5 x 10^{-6}, 3.7 x 10^{-6} and 2.0 x 10^{-6} M, respectively. Despite binding the least tightly to the protein, incorporation of Mn^{2+} into the enzyme resulted in the highest enzymatic activity as measured by spectrophotometric studies. This suggested that Mn^{2+} is possibly the native metal ion present in ε186. The ability of the metal ions to enhance ε186 enzymatic activity was found to follow the order:

Mn^{2+} >> Zn^{2+} > Dy^{3+}. The results of these experiments also provided evidence that the presence of two divalent metal ions was essential for efficient enzyme-catalysed hydrolysis.

The distribution of different oligomeric forms of wild-type *E. coli* DnaB helicase and DnaB helicase mutants (F102E, F102H, F102W and D82N) was examined using a factory-modified Q-ToF mass spectrometer equipped with a 32,000 *m/z* quadrupole. Previous experiments showed that the heptameric form of the wild-type protein was favoured in the presence of methanol (30% *v/v*). In the current work, mixtures of hexamer, heptamer, decamer and dodecamer were observed in solutions containing 1000 mM NH_4OAc, 1 mM Mg^{2+} and 0.1 mM ATP, pH 7.6. When the proteins were prepared in solutions containing a lower concentration of Mg^{2+} (0.1 mM), only the hexameric form was observed for all proteins except D82N, which showed a mixture of hexamer and heptamer. These observations suggest that the higher order structures were stabilised at high concentrations of Mg^{2+}. In addition, the hexamers of DnaB and mutants ($(DnaB)_6$, $(F102W)_6$ and $(D82N)_6$) formed complexes with four to six molecules of the helicase loading partner, DnaC.

ESI-MS was used in conjunction with hydrogen/deuterium exchange studies to probe the unfolding mechanisms of linear and cyclised DnaB-N (the N-terminal domain of DnaB helicase) containing linkers comprised of different numbers of amino acid residues (3, 4, 5 and 9). The unfolding rates for all the cyclised proteins were about ten-fold slower than for the corresponding linear proteins. These observations suggest that enhancement of protein stability against unfolding could be achieved

through cyclisation. Furthermore, the HDX data showed that all the proteins examined exhibited a rare EX1 mechanism at near neutral pH.

ABBREVIATIONS

ε186	N-terminal domain of ε
A_{420}	Absorbance at 420 nm wavelength
ADP	Adenosine-5′-diphosphate
AMP-PNP	β, γ-imidoadenosine-5′-triphosphate
ATP	Adenosine 5′-triphosphate
BIRD	Blackbody infrared radiative dissociation
bp	Base pair
bpy	2,2′-Bipyridine
BSA	Bovine serum albumin
CD	Circular dichroism
CI	Chemical ionisation
CID	Collision-induced dissociation
Da	Dalton
DAPI	4′,6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
dppz	Dipyrido[3,2-*a*:2′,3′-c]phenazine
dpq	Dipyrido[3,2-*d*:2′,3′-*f*]quinoxaline
dpqC	Dipyrido[3,2-*a*:2′,3′-*c*](6,7,8,9-tetrahydro)phenazine
$dpqMe_2$	Dipyrido[6,7-d:2′,3′-f]2,3-dimethylquinoxaline
DSC	Differential scanning calorimetry
dsDNA	Double-stranded DNA
DTT	D, L-Dithiothreitol
$Dy(OAc)_3$	Dysprosium(III) acetate

ECD	Electron-capture dissociation
EDTA	Ethylenediaminetetraacetic acid
EI	Electron ionisation
EM	Electron microscopy
EPR	Electron paramagnetic resonance
ESI	Electrospray ionisation
FAB	Fast atom bombardment
FD	Field desorption
FTICR	Fourier transform ion cyclotron resonance
HDX	Hydrogen/deuterium exchange
HSQC	Heteronuclear single quantum correlation
HMQC	Heteronuclear multiple quantum correlation
HX	Hydrogen exchange
ICP	Inductively coupled plasma
IR	Infrared
ITC	Isothermal titration calorimetry
k_{cat}	Turnover number (Michaelis-Menten kinetics)
K_d	Dissociation constant
kDa	Kilo Dalton
KF	Klenow fragment of Pol I (contains exonuclease domain)
kV	Kilovolts
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionisation

$Mg(OAc)_2$	Magnesium(II) acetate
MLCT	Metal-to-ligand charge transfer
$Mn(OAc)_2$	Manganese(II) acetate
M_r	Molecular mass
MS	Mass spectrometry
MWCO	Molecular weight cut off
NH	Amide hydrogen
NH_4OAc	Ammonium acetate
NMR	Nuclear magnetic resonance
NTP	Nucleoside triphosphate
PAGE	Polyacrylamide gel electrophoresis
PAP	Purple acid phosphatase
PD	Plasma desorption
Pda	9,10-diaminophenanthrene
PEG	Polyethylene glycol
phen	1,10-Phenanthroline
pm	Picometres
*p*NP-TMP	5′-*p*-nitrophenyl ester of thymidine-5′-monophosphate
Pol I	DNA polymerase I
Pol III	DNA polymerase III
Q-ToF	Quadrupole-time-of-flight
RNA	Ribonucleic acid
SPR	Surface plasmon resonance
SUPREX	Stability of unpurified proteins from rates of H/D exchange
ssDNA	Single-stranded DNA

TMP	Thymidine-5′-monophosphate
Tris-HCl	Tris (hydroxymethyl) amino methane hydrochloride
UV	Ultraviolet
$Zn(OAc)_2$	Zinc(II) acetate

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction to Biological Mass Spectrometry

1.1 Development of Biological Mass Spectrometry

Mass spectrometry (MS) is a technique that is used for structural analysis of molecules ranging in size from small organic compounds to large biological polymers such as proteins and nucleic acids. Mass spectrometry involves measurement of the mass-to-charge (*m/z*) ratio of gas phase ions, thus a sample must first be vaporised and ionised before being analysed by the mass spectrometer. Early ionisation methods, such as electron ionisation (EI)[1] and chemical ionisation (CI)[2] were limited to volatile compounds. In EI, volatile molecules are ionised directly by using an electron beam (generated by a heated filament in the ion source) and in CI ionisation occurs via reaction with gaseous ions generated by passing electrons through a reagent gas such as methane in the source of the mass spectrometer.[1,2] Since these two ionisation methods are only useful for volatile compounds, non-volatile samples had to be transformed into volatile derivatives prior to mass analysis, which generally limits the range of analysable compounds to those with molecular masses less than 1,000 Da. The development of field desorption (FD) in 1969 enabled the ionisation of thermally labile, non-volatile compounds, such as small peptides, with molecular masses less than 2,000 Da without transformation into volatile derivatives.[3] However, FD has not been widely used owing to tedious sample preparation procedures and technical difficulties.[3,4]

Plasma desorption (PD) was introduced as a new ionisation method capable of analysing large non-volatile biomolecules in 1974.[5] Torgerson *et al*. used the interactions of hundreds of heavy, high-energy ions (from the spontaneous fission of californium-252, ^{252}Cf) in a solid matrix to induce desorption and ionisation.[5] The essential feature of ^{252}Cf PD-MS was that the energy was highly concentrated, and the excitation lasted for only a short period of time. These conditions allowed large thermally labile molecules, such as insulin (Mr ~6,000 Da), and proteins as large as 30 kDa to be ionised without any fragmentation.[5]

In 1981, Barber and co-workers introduced fast atom bombardment (FAB), which is a soft ionisation method that results in minimal fragmentation of analyte molecules.[6] In a typical FAB analysis, a solution of sample is dissolved in the FAB matrix (usually glycerol), and introduced into the mass spectrometer vacuum system. The matrix-sample solution is then bombarded by fast atoms (8 keV, Xe) or ions (20 keV, Cs^+) and energy is transferred to the matrix-sample solution, and the resulting $[M + H]^+$ ions are vaporised along with the protonated matrix clusters. However, FAB was limited to intact proteins with molecular mass ≤ 10,000 Da and a single analysis requires a fairly large amount of sample (~1 nanomole),[7] in comparison to newer ionisation methods.

The development of two further soft ionisation techniques: matrix-assisted laser desorption ionisation (MALDI)[8,9] and electrospray ionisation (ESI),[10,11] allowed the observation of molecular ions of intact proteins using significantly smaller amounts of sample (≤10 picomoles). Through the advancement of ESI and MALDI techniques and other recent developments, mass spectrometry has become an

increasingly important technology for the analysis of biomolecules and in new field such as proteomics.[12,13] Identification of proteins can be accomplished by combining mass spectrometry, which provides information on the molecular masses of proteins, with partial amino acid sequences of peptides derived from tandem mass spectrometry (MS/MS) experiments, and database searching.[14,15]

1.2 Current Ionisation Techniques Used in Biological Mass Spectrometry

Since the development of MALDI[8,9] and ESI,[16] analysis of intact proteins using only small quantities of sample has become possible. These two ionisation methods play important complementary roles in structural studies in biomolecular research. In this section, the characteristics of MALDI and ESI that make these techniques so suitable for biological applications will be discussed.

1.2.1 Matrix-assisted laser desorption ionisation (MALDI) mass spectrometry

Laser desorption ionisation of large biomolecules was first reported in 1988 by Tanaka *et al*..[8] It was achieved by using a 337 nm low energy (nitrogen) laser and a matrix of glycerol containing colloidal particles (see Figure 1.1). Singly- and doubly-charged molecular ions were detected using a time-of-flight (ToF) mass spectrometer.[8] At about the same time, Karas and Hillenkamp reported the use of an ND-YAG laser for detection of large protein molecules (M_r ~14,000 - ~67,000 Da).[9] Singly-charged monomeric and dimeric protein ions as well as doubly-charged

protein ions were observed.[9] Since then, the technique has been improved further, mainly by Hillenkamp and co-workers.[9]

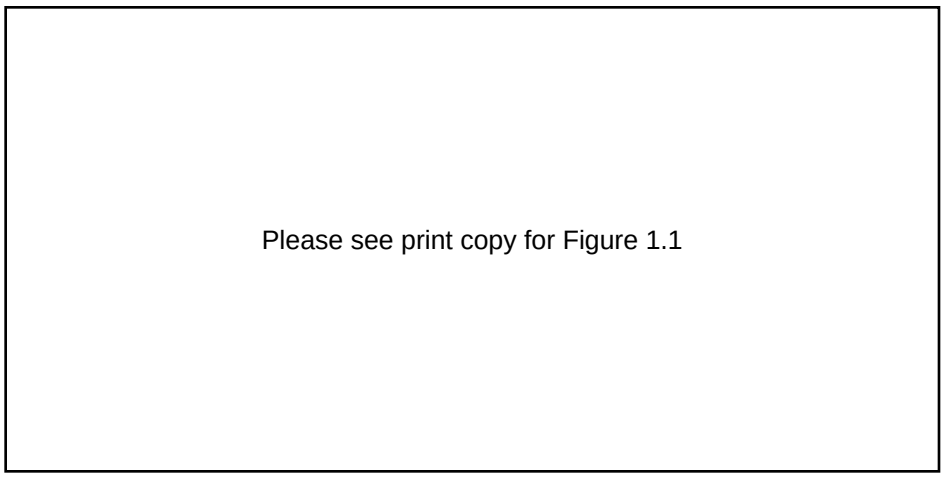


Figure 1.1 A schematic representation of the matrix-assisted laser desorption ionisation (MALDI) process. Laser desorption occurs when the sample (M), is irradiated with a focused laser beam. Most of the applied energy resulting from the laser irradiation is first absorbed by the matrix, resulting in vaporisation and ionisation of the matrix and sample molecules. Adapted from Akashi.[7]

MALDI is most often coupled to time-of-flight (ToF) mass analysers.[17,18] This is an ideal arrangement because both MALDI and ToF mass analysis are pulsed events. MALDI shares some similarities with FAB, however, instead of the sample being dissolved in a glycerol matrix, laser energy is directed at the co-crystallisation product derived from the sample and a light absorbing matrix. Light energy is absorbed by the crystals and dissipated, with the result that protonated (or deprotonated) sample molecules and matrix are vaporised. Owing to its robustness and high tolerance for impurities in comparison to ESI, MALDI is applicable to a large variety of compounds. Very large proteins (M_r ~1 MDa) have been ionised by MALDI.[19] Since a MALDI target plate can be loaded with as many as several hundred sample spots, and a single spectrum can be obtained in less than one minute, the technique is amenable to high sample throughput.[20-22] It is not, however, the

method of choice for very large proteins as signals arising from the resulting singly or doubly charged ions are often broad (partially as a result of adduct formation), typically in the order of several hundred Daltons for a protein with a molecular mass of 100 kDa.[23]

MALDI has found widespread application in the field of proteomics, especially in characterising proteins by their so-called peptide mass finger prints.[24-26] Once its primary structure has been determined, complete characterisation of a protein involves investigating its interaction with other biological molecules. There are only a limited number of reports about the detection of intact non-covalent complexes involving proteins using MALDI-MS.[27-34] One of the reasons for this is because the combination of a dried sample and an organic matrix (often acidic) is not an ideal environment for the maintenance of non-covalent biological complexes. In addition, it is difficult to distinguish the molecular ions of specific non-covalent complexes from those of non-specific interactions. Consequently, application of MALDI-MS for detection of biomolecular non-covalent complexes is not routine. In contrast, the other soft ionisation technique, ESI, is now rapidly expanding as a tool for analysis of non-covalent biological complexes.

1.2.2 Electrospray ionisation (ESI) mass spectrometry

The concept of ESI-MS was first introduced by Dole and co-workers in the late 1960s.[35] However, the first successful analysis of a large intact molecule, polyethylene glycol (PEG), by ESI-MS was reported by Yamashita and Fenn in 1984.[10,11] Later, in 1989, Fenn and his co-workers use ESI-MS to measure the molecular mass of proteins.[16] Since then, the application of ESI-MS for studying

non-covalent complexes has become increasingly important in biomolecular research.[18,36-38] In the electrospray process, small charged droplets containing the analyte are initially formed at the tip of a capillary, which is typically subjected to a potential of 1-4 kV. As a result, molecular ions with multiple charges, such as $[M + nH]^{n+}$ (when a positive voltage is applied to the capillary) or $[M - nH]^{n-}$ (when a negative voltage is applied to the capillary) are generated. The droplets diminish in size by solvent evaporation, assisted by a warm flow of nitrogen gas which passes across the front of the ionisation source (Figure 1.2). While the droplets decrease in radius, the charge is conserved, therefore at some critical radius Coulombic forces overcome the surface tension of the liquid leading to fission of the droplets (at the Rayleigh limit) into even smaller droplets.

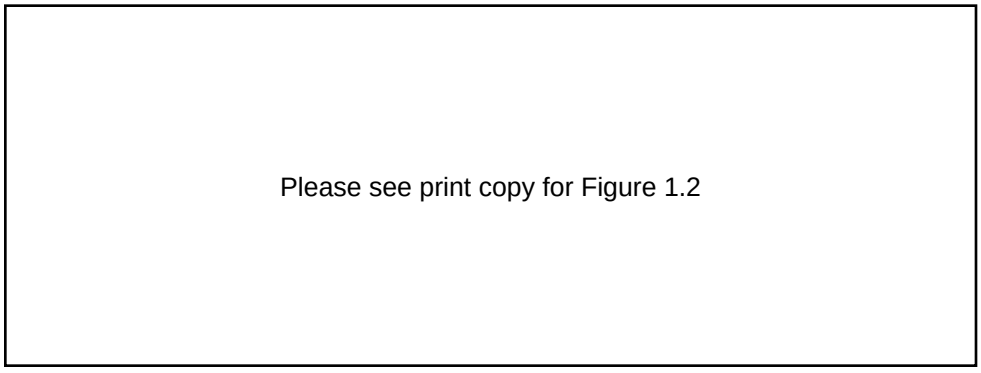


Figure 1.2 A Schematic representation of droplet formation at atmospheric pressure inside an ESI mass spectrometer source. Adapted from Gaskell.[39]

The evaporation and fission processes continue until the point is reached that either an ion desorbs from a droplet[40,41] or solvent is completely removed,[42] resulting in very small charged droplets, which are the precursors of the gas phase ions. However, the exact mechanism of ion formation, whether it is by ion evaporation (ion-evaporation model)[40] or by complete solvent removal (charge-residue model),[42] from the charged droplet is still under debate. There is some evidence suggesting that

different mechanisms occur in different situations, and both mechanisms may operate to some extent during the evaporation and ionisation processes.[43-46] The highly charged droplets are then passed down a potential and pressure gradient towards the analyser in a high vacuum system. Electrospray ionisation sources are often coupled with ion trap,[47-50] quadrupole (or triple quadrupole),[11,51-53] time-of-flight[54] or quadrupole time-of-flight mass analysers[55-57]

Since electrospray ionisation takes place at atmospheric pressure, it is very gentle and hence there is no significant fragmentation of the analyte ions (maintaining not only covalent bonds but often also weaker non-covalent associations). As a consequence of its ability to maintain non-covalent interactions in the gas phase, ESI-MS has been successfully employed to study a wide variety of non-covalent complexes (discussed in section 1.3.2), including multimeric proteins,[36,58-60] protein-ligand complexes,[61-66] protein-protein complexes,[61,67] protein-metal complexes,[37,68,69] protein-DNAcomplexes,[70-72] double-stranded oligonucleotides and higher order nucleotide complexes,[73-75] and non-covalent complexes involving small organic molecules such as drugs bound to oligonucleotides.[76-81]

ESI-MS has several important advantages over more traditional methods of analysis that can be exploited for studies of non-covalent complexes. ESI-MS allows exact molecular masses (within 0.01%) of large biomolecules such as proteins and their complexes to be determined.[16] Furthermore, ESI-MS can determine the identities as well as the stoichiometries of components of a multimeric complex. ESI-MS has been referred to as having the "S" advantages in a review by Loo in 1997.[82] The most obvious "S" advantages of ESI-MS are its speed and sensitivity. In comparison to

techniques such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, mass spectrometry can acquire data in a very short period of time (less than one minute). ESI-MS is capable of detecting analytes in the picomole to femtomole range,[58,83] while for many NMR experiments, micromoles of sample are required. MS also has the ability to distinguish between, for example, complexes containing different numbers of metal ions (Ca^{2+}) bound to the protein calmodulin.[68,84] In other words MS can directly provide the stoichiometry of protein-metal complexes.[68] The final "S" advantage is specificity, which refers to the ability of ESI mass spectra to reflect differences in binding affinity between different binding partners. For example, ESI-MS was used to study the binding of the $pp60^{v\text{-}Src}$ SH2 (Src homology 2) domain protein with non-phosphorylated and phosphorylated peptides. This study showed there was a greater binding affinity displayed by the phosphorylated peptides to Src SH2 than non-phosphorylated peptides with the same sequence.[85]

1.3 Non-Covalent Complexes

The application of ESI-MS to protein sequence determination and identification ("proteomics") is well established and most protein researchers now have access to laboratories dedicated to these tasks. A full understanding of biological processes, however, depends on understanding specific, non-covalent interactions between molecules. For example, proteins may interact with other proteins, peptides, small molecules, nucleic acids and oligonucleotides, lipids and polysaccharides. These interactions regulate many cellular processes such as cell division, cell signalling, ion transport, gene transcription, translation and homeostasis. It is important to determine factors including: (i) the identity of the binding partners, (ii) the stoichiometry of

binding, (iii) the strength of binding, (iv) the contact points between binding partners, and (v) whether there are conformational changes on binding.

In parallel to the use of mass spectrometry in proteomics for the elucidation of the primary structure of proteins, the technique has also been used as a complementary tool in structural biology for the investigation of higher order structure of protein complexes[86] and their interactions with DNA,[87-89] RNA,[90] ligands[91,92] and cofactors.[93-95] In addition, mass spectrometry has also been used to study the interactions of drugs with DNA and proteins and ternary drug-DNA-protein complexes.[61,96] These complexes occur through non-covalent associations involving electrostatic interactions, hydrogen bonds, and hydrophobic interactions.

There are many established techniques that have been used historically to study non-covalent interactions, including circular dichroism (CD) spectroscopy, light scattering, fluorescence spectroscopy, infrared (IR) spectroscopy, ultraviolet (UV) spectroscopy, surface plasmon resonance (SPR), differential scanning calorimetry (DSC), isothermal titration calorimetry (ITC), nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography.[97] Each of these methods has its own strengths and weaknesses. The techniques vary in many ways including the amount of sample that is required, the required levels of sample purity, the extent to which the data obtained are qualitative or quantitative, and the level of detail in the structural information that is obtained.

1.3.1 Brief overview of techniques for studying non-covalent complexes

Methods such as UV, IR, CD and fluorescence spectroscopy can be used to investigate changes in the three dimensional structures of biomolecules. [97-100] In addition, they can also be useful for characterisation of biomolecular complexes. For example, Otto-Bruc *et al.* measured the affinity of the GTP-bound α-subunit of the G-protein transduction for the γ-subunit of retinal cyclic GMP phosphodiesterase by monitoring changes in intrinsic fluorescence.[98] Dissociation constants in the range of 10^{-6}-10^{-11} M can be obtained using spectroscopic methods.[97]

Analytical ultracentrifugation is used for determination of molecular mass, conformation and shape, dissociation constants, and also for characterisation of macromolecular complexes (e.g. through determination of binding stoichiometries).[97] The analytical ultracentrifugation technique allows dissociation constants between 10^{-3}-10^{-8} M to be measured.[97] In this technique, a sample being centrifuged can be monitored in real time through an optical detection system (typically using UV light). This allows the observation of changes in sample concentration caused by variations in the axis of rotation of the centrifugal field. An analytical ultracentrifuge can be used to perform two types of experiments: sedimentation velocity and sedimentation equilibrium. Sedimentation velocity is a hydrodynamic technique and is sensitive to the mass and shape of the macromolecular species, whereas sedimentation equilibrium is a thermodynamic technique that is sensitive to the mass, but not the shape, of the macromolecular species.[101-103]

Another method that can be used to measure equilibrium constants and detect intermolecular interactions directly is surface plasmon resonance (SPR). It is a label-free, real-time binding technique which allows analysis of the kinetics (on- and off-rates) of macromolecular interactions. One component of a macromolecular complex is attached (immobilised) to a gold surface and the other is allowed to flow past. Interactions between the two species can be monitored by the change in refractive index at the sensor surface.[104] SPR allows measurement of dissociation constants in the range of 10^{-4}-10^{-11} M,[97] indicating that it has the ability to study high affinity interactions. SPR is a very useful technique for rapid screening of conditions for visualising protein-protein interactions, and only small amounts of protein are required.[97] Despite several great advantages of SPR, the immobilisation process may modify the protein and prevent its normal interactions with other molecules from being observed.

Isothermal titration calorimetry (ITC) is one of the most robust methods for characterising protein-protein interactions,[105] protein dynamics[106] and drug-DNA interactions.[107] In ITC experiments the heat change that occurs when a complex is formed at a constant temperature is monitored, allowing the enthalpy change associated with complex formation to be measured.[108] This is achieved by titrating one of the binding partners of the complex into a reaction cell containing the other partner, and comparing the energy absorbed or released to that of a reference cell. Since the heat change upon complex formation is directly proportional to the amount of binding occurring, dissociation constants in the range of 10^{-6}-10^{-11} M can be measured.[97]

Another calorimetric approach to studying interactions with or between macromolecules is differential scanning calorimetry (DSC). This technique can be used to measure protein stability and very high affinity protein-protein interactions. DSC can also provide information on conformational changes that result from macromolecular binding interactions. It is a thermoanalytical technique in which the amount of heat required to increase the temperature of a sample (compared with a reference) is measured as a function of temperature. The basic principle underlying the DSC technique is that there is a measurable heat change when a molecule undergoes a physical transformation.[109] This allows, for example, measurement of the thermodynamic parameters associated with a change in protein conformation between a folded and an unfolded state. DSC can be used to measure dissociation constants in the range of 10^{-9}-10^{-20} M.[97]

Traditional chromatographic- and electrophoretic-based assays, such as size exclusion chromatography and gel electrophoresis, have long been used to detect and determine the molecular masses of biomolecules as well as the stoichiometry of biomolecular complexes.[110-117] However, the mass accuracy of these methods is rather low (± 15%). This is because these techniques rely on comparisons of elution times (for chromatography) or migration distances (for electrophoresis), which are sensitive to the shape and physical properties of the protein. When a complex forms, the binding partners often undergo conformational changes. As a result, errors in mass determination may occur.

X-ray crystallography is an important and powerful technique for determining protein structures to near atomic resolution. When an X-ray beam bombards a

crystalline lattice, the pattern produced by diffraction of the X-ray beam through the closely spaced lattice of atoms in the crystal is recorded and analysed using Bragg's Law.[118,119] This then affords the molecular structure of the molecule.[119] To determine a structure, the molecule of interest must first be crystallised and frozen (to reduce radiation damage incurred during data collection and decrease thermal motion within the crystal). Crystals are used because the diffraction pattern from one single molecule can be insignificant, whereas many identical individual molecules in a crystal amplify the diffraction pattern. However, there are difficulties in growing high quality crystals of some proteins, especially those containing unstructured domains. Acquisition of high quality X-ray images can also be hampered by discontinuities in the crystal structure, caused by temperature variations within the growing crystal.[118] Moreover, X-ray crystal structures trap the protein in a single conformation, and therefore may sometimes not reveal other conformations which may be crucial for its function.[120,121]

NMR spectroscopy is a powerful non-destructive technique that can be used to provide information on the three dimensional structures of proteins. It can also be used to study protein dynamics, protein-protein complexes and protein-ligand interactions.[122] NMR spectroscopy takes advantage of the magnetic properties of atomic nuclei. When placed in an external magnetic field, NMR- active nuclei such as ^{1}H or ^{13}C resonate at a specific frequency, dependent in part on the strength of the applied field, which is converted into a field-independent value known as the chemical shift. In addition, different nuclei of the same type (e.g. ^{1}H) in a molecule resonate at slightly different frequencies, depending on the local chemical environment. By understanding how different chemical environments influence

chemical shifts, signals can be assigned to specific atoms or groups of atoms. This information, in combination with signal integrations (peak areas) and spin-spin coupling values, provides information about the identity of molecules. Two dimensional NMR methods such as NOESY (nuclear Overhauser effect spectroscopy), HSQC (heteronuclear single quantum correlation) and HMQC (heteronuclear multiple quantum correlation) provide more information about a molecule, including its three dimensional structure in solution, than can be obtained by one dimensional NMR techniques. NMR spectroscopy and X-ray crystallography are complementary methods in that they provide high resolution three dimensional structures of biomolecules and complexes in the solution phase and solid phase, respectively.[123] However, they both share some disadvantages. For example, both methods require a large amount of sample, and analysis of data can be rather complicated and time consuming. Mass spectrometry offers an alternative and complementary technique for studying non-covalent macromolecular interactions in a highly sensitive and more timely fashion, and with high mass accuracy.

1.3.2 ESI-MS studies of non-covalent complexes

The properties of non-covalent complexes in the gas phase need to be comparable to those observed in solution in order for ESI-MS to be used as an alternative and/or complementary tool for biological research. ESI-MS was first used to observe specific non-covalent complexes in 1991 by Ganem *et al.*.[124,125] They examined non-covalent receptor-ligand complexes formed between the naturally occurring cytoplasmic receptor FKBP and the immunosuppressive agents FK506 protein and rapamycin.[125] Enzyme-substrate interactions between lysozyme and N-acetylglycosamine have also been studied using the same method,[124] while Katta and

Chait observed the specific non-covalent interaction between heme and globin in its native state at pH 4.4.[126] These early reports established not only the practicality of the ESI-MS method, but also the importance of experimental design to ensure valid observations. Subsequently, numerous studies on the application of ESI-MS to the study of non-covalent associations have been reported. In addition, several reviews on this area have been published.[7,23,37,82,127-131] The section below describes some of the ESI-MS investigations that have been carried out on protein-DNA, protein-metal/ligand, dsDNA, dsDNA-drug and multimeric protein non-covalent complexes which are of relevance to this thesis.

1.3.2.1 ESI-MS of protein-DNA complexes

Protein-DNA interactions are involved in many cellular processes. In one of the earlier ESI-MS studies of protein-DNA interactions, Cheng *et al*. examined the complexes formed between the gene V protein from bacteriophage f1 and a series of different oligonucleotides (13-18 bases).[88] This protein stabilises ssDNA during phage replication. The ESI-MS data showed that complexes of the protein with oligonucleotides shorter than 15 bases have 2:1 protein:oligonucleotide stoichiometries, whereas complexes with oligonucleotides containing 16 bases or more had a 4:1 stoichiometry.[88] These binding stoichiometries were in agreement with those obtained from solution phase studies.[132,133]

Electrostatic interactions often play an important role in stabilising protein-DNA complexes. These interactions are thought to be strengthened *in vacuo*,[82] whereas the use of high capillary voltages has been shown to cause their dissociation.[70] The choice of ion mode for ESI-MS observation of protein-DNA complexes must also be

considered carefully. Proteins are usually detected in positive ion mode, whereas DNA is normally detected in negative ion mode. A review by Beck *et al*. provides some examples of ESI-MS studies of protein-DNA complexes up to 2001.[128] The first ESI-MS study in which a complex of dsDNA with its protein binding partner was investigated involved the DNA-binding domain of the eukaryotic transcription factor PU. 1 (PU.1 DBD).[70] ESI-MS showed that a 1:1 protein-dsDNA complex was formed only when DNA of a specific sequence was used. These results were consistent with those obtained by gel electrophoresis mobility shift assay.[70]

Kapur *et al*. used ESI-MS to examine the binding interactions between a replication terminator protein (Tus) and double-stranded DNA with a variety of sequences including the specific DNA sequence (*Ter*) recognised by Tus *in vivo*.[71] This specific interaction halts the chromosomal replication of *E. coli*. Only a 1:1 complex of Tus protein and *Ter* DNA was observed using ESI-MS, suggesting that the Tus-*Ter* complex was the result of a specific interaction.[71] Kapur *et al*. also measured dissociation constants for the binding of Tus mutant proteins with *Ter* DNA. These results were also in agreement with those obtained from solution studies.[71] In addition, the extent of dissociation of the complex was found to be increased with an increase in NH_4OAc concentration, indicating the overall stability of the complex was maintained by electrostatic interactions.[71]

ESI-MS has also been used to detect conformational changes of a protein upon DNA binding.[72] The protein studied was the catalytic domain of bacteriophage λ integrase that catalyses site-specific DNA recombination.[72] When free protein was examined, three distinct charge distributions of ions in the ESI mass spectrum attributed to

unfolded, folded and dimeric protein were observed. Addition of a cognate DNA sequence to the solution containing the protein changed the charge distribution, giving ions that were attributed to the protein-DNA complex and the free folded protein. The absence of the unfolded conformation suggested that DNA binding stabilised the global fold of the protein, and that DNA was bound specifically to monomeric λ-integrase.[72] In this study, ESI-MS allowed observation of populations of different protein conformations in a single spectrum rather than showing an average of different folded states, such as is observed using most spectroscopic techniques.[72]

1.3.2.2 ESI-MS of protein-metal and protein-ligand complexes

Metal ions are important for many proteins for the formation of stable conformations, and/or for enzyme catalysis. The potential of ESI-MS to determine protein-metal ion stoichiometries has been demonstrated by many reports.[68,84,134-137] For example, the Ca^{2+}-binding proteins bovine calmodulin, rabbit parvalbumin, and bovine α-lactalbumin were found to bind specifically to four, two and one Ca^{2+} ions, respectively.[68,84] ESI-MS was also used to accurately determine the metal-binding stoichiometry of the Ca^{2+}-binding protein calbindin D_{28K}. Results from NMR and CD spectroscopy, fluorescence and gel electrophoresis indicated that between 3-6 moles of Ca^{2+} bind to every mole of protein,[138-140] while Veenstra *et al*. used ESI-MS to show that the protein binds 4 moles of Ca^{2+} per 1 mole of protein.[141]

ESI-MS has also been shown to be invaluable for studies of metalloproteins that contain more than one type of metal ion. For example, the ESI-MS study of a metalloproteinase, matrilysin, by Feng *et al*. revealed that the protein contains two

Zn^{2+} and two Ca^{2+} binding sites.[142] Since the atomic masses of Zn^{2+} and Ca^{2+} are significantly different, the individual stoichiometry of each metal relative to the protein was easily determined.

Changes in the overall appearance of ESI mass spectra of multiply-charged proteins have been used to detect protein conformational changes upon binding to metal ions.[135,141,142] One example involved the DNA-binding domain of the vitamin D receptor, which binds Zn^{2+}. At a 1:1 ratio of Zn^{2+} to protein, where only one metal is bound, there was very little change observed in the ESI mass spectrum compared to that of the free protein. However, when the ratio was increased to 2:1, a significant change in the spectrum was observed as a result of a change in protein conformation caused by the binding of a second Zn^{2+} ion. These results were consistent with changes observed in CD spectra.[143]

1.3.2.3 ESI-MS of dsDNA

In contrast to proteins, oligo- and polynucleotides have been mainly studied using negative ion mode ESI-MS.[37,144-146] Double-stranded (ds) DNA is stabilised by a number of factors, with the major contributions being from hydrogen bonding between bases on the two strands, base stacking within each strand,[147] the aqueous environment on the outside of the duplex,[148] and the presence of counterions.[37] Within cells electrostatic interactions between the phosphodiester groups and Na^+ or K^+ ions in the surrounding medium are an important stabilising factor for dsDNA. In ESI-MS experiments, however, the presence of involatile alkali metal ions leads to formation of complex adducts with DNA, which decrease sensitivity in ESI-MS

analysis.[73,149] Therefore, relatively high concentrations of ammonium acetate have been used in ESI-MS experiments to stabilise the duplex.

Reports by Ganem *et al*.[74] and Light-Wahl *et al*.[73] in 1993 demonstrated that dsDNA could be successfully transferred from solution to the gas phase as an intact molecule using ESI-MS. Another study by Doktycz *et al*. using an ESI ion trap mass spectrometer illustrated that DNA duplexes could survive in the gas phase for hundreds of milliseconds. These results indicated that the ions were kinetically stable during ion injection, storage, and mass analysis at a relatively high pressure (1 mTorr) of bath gas.[150] Bayer *et al*. examined the stability of double-stranded oligonucleotides of different lengths (8-30-mer) with natural and chemically modified oligonucleotides.[151] They observed that as the length of the DNA strands increased, the stability of the duplex also increased, along with the signal intensity of the duplexes observed in the mass spectra.[151]

Ding and Anderegg examined specific and non-specific dimer formations of oligonucleotides (between 6- to 15-mer) in the presence of complementary and non-complementary DNA strands using negative ion ESI-MS.[152] They found that formation of dimers was concentration-dependent, which could only be observed when the concentration of each oligonucleotide in the mixtures was ≥ 100 μM.[152] In addition, they also studied the effect of the energy of the incoming ions (orifice potential) on duplex formation. Unlike other studies on oligonucleotide duplexes by Ganem *et al*.[74] and Gabelica *et al*.,[153] as the energy of the orifice potential was increased, the abundance of ions assigned to dsDNA increased with respect to those assigned to ssDNA.[152]

DNA duplexes have been shown to be readily observed using ESI-MS as intact molecules under gentle ionisation conditions.[73,74,144,151,152] However, under these conditions the overall gas phase structures and the interactions between the two strands cannot be determined. Several research groups have attempted to address these issues by studying dissociation of the duplexes. For example, Schnier *et al*. examined the dissociation kinetics of a series of complementary and non-complementary DNA duplexes using blackbody infrared radiative dissociation (BIRD) in a Fourier transform mass spectrometer.[154] Their results provided strong evidence supporting formation of Watson-Crick base pairs in complementary DNA duplexes, which could exist in the complete absence of solvent in the gas phase. The activation energy for dissociation of the complementary duplex, $(A_7 \cdot T_7)^{3-}$, was notably higher than that of the two related non-complementary duplexes, $(A_7 \cdot A_7)^{3-}$ and $(T_7 \cdot T_7)^{3-}$, indicating a stronger interaction between strands with a specific base pairing sequence. An extensive loss of neutral adenine base was observed for $(A_7 \cdot A_7)^{3-}$, and $(A_7 \cdot C_7)^{3-}$ but not for $(A_7 \cdot T_7)^{3-}$, suggesting the loss of bases was prevented by Watson-Crick hydrogen bonding. A correlation observed between the gas phase dissociation activation energy and the dimerization enthalpy ($-\Delta H_d$) in solution, and molecular dynamics studies, indicated that Watson-Crick base pairing in $(A_7 \cdot T_7)^{3-}$ was preserved in the gas phase.[154]

Gabelica and De Pauw examined a series of 16-mer DNA duplexes using collision-induced dissociation (CID) by varying the capillary-skimmer voltage (0-135 V) in a quadrupole-ToF hybrid mass spectrometer.[155,156] The relative kinetic stabilities in the gas phase were found to correlate well with those obtained in solution measured by thermal denaturation (monitored by UV spectrophotometry).[155] Their studies showed

that dsDNA of the same length that contained a higher GC content had higher thermal stability (in both solution and gas phase) than those containing a lower GC content. They attributed these results to the hydrogen bonding and base stacking interactions in solution, which were maintained in the gas phase. Furthermore, their results indicated that the size of the duplexes were also crucial in CID experiments, as the larger duplex (with the same GC content) had the greater V_m value (voltage at which half the fragmentation in CID experiments was observed).[155] A study of the dissociation mechanism of duplex DNA has also been conducted using CID experiments by Gabelica and De Pauw.[157] The results of this study showed that the dissociation of duplex DNA into single strands, which involves cleavage of non-covalent bonds, occurred under fast activation conditions, which favoured entropy driven dissociations in the collision cell of a hybrid quadrupole-ToF instrument. In contrast loss of neutral bases from duplex DNA by cleavage of covalent bonds was favoured by slow activation conditions using a quadrupole ion trap.[157] Furthermore, evidence from this study also suggested that the dissociation of dsDNA into ssDNA is a multi-step process involving a progressive unzipping of the molecule, which occurs preferentially at terminal positions.[157]

1.3.2.4 ESI-MS of dsDNA-drug complexes

The specific interactions that occur when small organic molecules bind non-covalently to dsDNA provides the basis for many antiviral, antitumour and antibiotic drugs.[158-161] This is because compounds that have a high affinity towards DNA can greatly affect cell replication, translation and transcription. DNA duplexes can bind to drugs via several different modes of binding including electrostatic interactions,

groove binding and intercalation (see section 3.2.2). In many cases drugs bind to DNA using a combination of these binding modes.[162]

An ESI-MS study of a non-covalent dsDNA-drug complex was first reported by Gale *et al*. in 1994.[163] They examined a dsDNA molecule formed by a self-complementary 12-mer sequence, 5′-dCGCAAATTTGCG-3′, and a non-covalent complex of the duplex with distamycin A, a minor groove binder. Non-covalent complexes with a 1:1 or 2:1 distamycin:duplex DNA ratio were observed, depending on the relative concentrations of the binding partners. These observations were consistent with other data from solution studies.[163] In addition to the gentle conditions used in ESI-MS, careful selection of salt and buffer concentrations also played an important part in maintaining complex stability.[163] Following on from this work a more detailed investigation of the non-covalent complexes formed between dsDNA formed from the same 12-mer-self-complementry DNA sequence and three minor groove binders (distamycin A, pentamidine and Hoechst 33258) was conducted.[164] In this study various electrospray ionisation interface parameters including the capillary-skimmer potential and the capillary temperature, as well as collision-induced dissociation were utilised to characterise the nature and stability of the non-covalent complexes. There was no evidence for the presence of non-specific dimers, trimers, or other aggregations in thenESI mass spectra. The stability of a non-covalent complex with a 2:1 distamycin:duplex DNA ratio was found to be greater than that for the corresponding 1:1 complex, which was in agreement with results from solution studies.[164] These early investigations demonstrated the potential of ESI-MS for characterisation of non-covalent complexes formed between small molecules and dsDNA.[164]

Wan *et al*. studied complexes formed between ten different compounds, including minor groove binders, intercalators, porphyrins and metallophophyrins, and dsDNA derived from 6- to 12-mer self-complementary oligonucleotides, using an ESI-ion trap mass spectrometer.[165] Their results showed that minor groove binders (distamycin, Hoechst 33258, Hoechst 33342, and berenil) preferred to bind to AT-rich dsDNA, whereas the intercalator actinomycin D preferred to bind to GC-rich dsDNA. The order of binding affinities of the minor groove binders obtained from competition experiments in the gas phase was found to be: Hoechst 33342 > Hoechst 33258 > distamycin > berenil.[165]

1.3.2.5 ESI-MS of multimeric protein subunits

Most cellular events, such as replication, transcription and translation, are regulated by large multimeric protein complexes rather than by individual proteins. These processes are essential in metabolic control and differentiation. In addition to coordinating multiple enzyme activities, it is believed that protein oligomerisation improves the stability of the proteins against proteolysis and thermal denaturation.[23] MS is a sensitive method to determine the stoichiometries of protein subunits since multimeric interactions have been preserved during the transfer of complexes from solution to the gas phase. The measured molecular mass of the oligomer directly reveals the number of subunits in the quaternary protein structure. The first ESI-MS observation of a protein subunit complex was reported by Baca and Kent in 1992.[166] Molecular ions of a ternary complex between a homodimeric enzyme HIV-1 protease and a substrate-based inhibitor were observed. Since an accurate mass of the complex was obtained by ESI-MS, the binding stoichiometry of the dimeric protein

and its binding partner was determined without ambiguity.[166] Soon after this successful demonstration, many more studies on multimeric protein subunit interactions were reported (e.g. tetramers of concanavalin A,[60] avidin,[36] streptavidin,[83] and haemoglobin;[36] a dimer and tetramer of alcohol dehydrogenase[59], a hexamer of insulin[167] and dimer of a leucine zipper peptide[168]).

Fitzgerald *et al*. investigated enzyme oligomers of 4-oxalocrotonate tautomerase (4OT) and four mutant enzyme monomers.[58] Under non-denaturing conditions, an intact hexameric form of native 4OT was detected by ESI-MS, whereas for the mutant proteins only monomeric forms were observed. These results were consistent with structural data obtained from CD spectroscopy.[58] In addition to providing information about the oligomerisation behaviour of the wild type protein the study provided evidence of specific residues that might be important in stabilising the hexameric structure. ESI-MS has also been used to determine solution association constants for oligomeric forms of the enzyme citrate synthase from *E. coli*.[169] The association constant for the formation of the hexameric form of the protein from the dimeric form was obtained, and found to be within an order of magnitude of the value measured using analytical ultracentrifugation.[169]

Lei *et al*. examined the state of oligomerisation and the metal atom stoichiometry of the non-heme iron-containing multimeric proteins hemerythin and rubrerythrin using ESI-MS.[170] Under non-denaturing conditions, they found that both proteins existed in octameric forms, with molecular masses of ~110 kDa.[170] Furthermore in-source CID experiments, involving increasing the capillary-skimmer voltages, allowed the exact number of metal atoms present in each subunit to be determined. These experiments

also provided information on the oxidation state of the metal from the number of hydrogen atoms displaced during metal binding.[170]

1.4 Scope of the Thesis

This thesis presents the results of three investigations that highlight selected applications of ESI-MS for studying non-covalent interactions of biological macromolecules with ligands, metal ions or other macromolecules. Further, the effect of tethering the N- and C-termini of proteins on unfolding was investigated by using ESI-MS to measure the exchange rates of amide protons. Proteins of the *E. coli* replisome, the multiprotein-nucleic acid complex that contains the enzymes and protein scaffolds that replicate the bacterial chromosome, were used in these studies. The binding of small molecules (potential nucleic acid probes/drugs) to double-stranded DNA was also investigated, as these interactions are expected to interfere with processes such as DNA replication.

In chapter three, the non-covalent binding of various ruthenium-based metallointercalators to 16-mer dsDNA molecules is outlined. A series of titration experiments involving six different ruthenium molecules and three different DNA sequences are described, as well as competition experiments in which pairs of ruthenium compounds were allowed to complete for a single DNA molecule. Results obtained from these experiments using ESI-MS provided information about the order of relative binding affinities and DNA sequence selectivities of the ruthenium compounds. In addition, competition experiments involving ruthenium compounds and organic DNA-binding drugs were conducted in order to obtain more information about their modes of DNA binding and possibly their preferred DNA-binding sites.

The focus in chapter four shifts to the protein subunit of *Escherichia coli* DNA polymerase III (ε), that is involved in proof-reading newly synthesized DNA. The exonucleolytic activity of this enzyme requires a metal ion, most likely Mn^{2+}. ESI-MS was used to determine and compare the dissociation constants for the binding of the metal ions Mn^{2+}, Zn^{2+} and Dy^{3+} to the enzyme. The magnitudes of the dissociation constants were compared with the enzymatic activities of the protein measured in the presence of the metal ions.

While the experiments described in the first chapters of this thesis were performed using a commercial ESI Q-ToF mass spectrometer (quadrupole to *m/z* 4,000), the work reported in chapter 5 was carried out using a factory-modified ESI Q-ToF mass spectrometer (quadrupole to *m/z* 32,000) with the facility to increase the argon pressure in the ion optics region. This enables "collisonal cooling" which enhances sensitivity, especially for very large ions. This mass spectrometer was used to compare the oligomeric forms of wild-type and mutant forms of the hexameric helicase from *E. coli*, DnaB. This helicase unwinds DNA in advance of DNA polymerase III, allowing the template DNA to be copied by the polymerase. Heptamers of DnaB were observed in the absence of organic solvent for the first time, as were higher order oligomers such as decamers and dodecamers. The propensity to form these higher order structures was found to vary among the mutants.

Finally, chapter six highlights the application of ESI-MS to understanding fundamental aspects of protein unfolding. Aspects of the mechanism of protein unfolding were probed by cyclising the N-terminal domain of DnaB protein by

insertion of amino acid linkers of varying length between the N- and C-terminus. The rates of unfolding of linear and cyclised proteins with varying linker lengths were obtained from hydrogen/deuterium exchange experiments. The length of the amino acid linkers did not have any influence on the rate of protein unfolding, since similar rates were obtained using proteins with different linker lengths. However, the rates obtained from all the cyclised proteins were approximately ten-fold slower than those from the linear version. These results suggest that cyclisation leads to an enhancement of protein folding stability. In addition, the amide proton exchange process was found to follow a rare EX1 kinetics mechanism at the near neutral pH levels used in this study.

Chapter 2

Materials & Methods

2.1 Materials

All solvents and chemical reagents used were the highest grade commercially available. MilliQ™ water from Millipore (Molsheim, France) was used in all experiments.

Ruthenium compounds (PF_6 salts): $[Ru(phen)_2(dpq)]^{2+}$, $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$ were kindly provided by Dr Janice Aldrich-Wright (School of Science, Food and Horticulture, University of Western Sydney, Australia).[a] Proteins theta (θ), epsilon 186 (ε186, residues 1-185), DnaB, DnaB-N (N-terminal domain of DnaB), DnaB mutants (F102E, F102H, F102W, D82N) and DnaC were a kind gift from Dr Nicholas Dixon (Research School of Chemistry, Australian National University, Australia).

Ethylenediaminetetraacetic acid (EDTA, free acid), daunomycin (daunorubicin), distamycin A, the 5′-*p*-nitrophenyl ester of thymidine-5′-monophosphate (*p*NP-TMP), magnesium acetate ($Mg(OAc)_2$), manganese(II) acetate ($Mn(OAc)_2$) and

[a] phen = 1,10-phenanthroline; dpq = dipyrido[3,2-*d*:2′,3′-*f*]quinoxaline; dpqC = dipyrido[3,2-*a*:2′,3′-*c*](6,7,8,9-tetrahydro)phenazine; dppz = dipyrido[3,2-*a*:2′,3′-c]phenazine; pda = 9,10-diaminophenanthrene; $dpqMe_2$ = dipyrido[6,7-d:2′,3′-f]2,3-dimethylquinoxaline. Structures of ruthenium compounds containing these ligands are shown in section 3.3.

dysprosium(III) acetate ($Dy(OAc)_3$) were purchased from Sigma-Aldrich (St Louis, USA). $[Ru(phen)_3]^{2+}$ was obtained as the chloride salt from Aldrich (Milwaukee, USA). Ammonia, ammonium acetate (NH_4OAc), acetic acid, acetonitrile (HPLC-grade), formic acid, methanol (HPLC-grade), sodium chloride, manganese(II) chloride and zinc(II) acetate ($Zn(OAc)_2$) were obtained from Ajax Finechem (Seven Hills, Australia). Tris hydrochloride (Tris-HCl), adenosine-5′-diphosphate (ADP, free acid), adenosine-5′-triphosphate (ATP, free acid) and dithiothreitol (DTT) were purchased from ICN Biomedicals (now MP Biomedicals; Aurora, USA). Deuterium oxide (D_2O) was purchased from Cambridge Isotope Laboratories (Andover, USA). Dialysis tubing (3,500 and 10,000 molecular weight cut-off (MWCO)) were purchased from Crown Scientific (Moorebank, Australia). Millipore Biomax centrifugal filters (5,000 molecular weight cut-off) were obtained from Millipore (Bedford, USA). Reagents for the Bio-Rad DC protein assay were obtained from Bio-Rad (Hercules, USA). Nanospray capillaries (Au/Pd coated, medium size) were obtained from Proxeon (Odense, Denmark). Custom-made oligonucleotides were obtained from Geneworks (Adelaide, Australia).

2.2 Methods

2.2.1 Reactions of oligonucleotides with ruthenium compounds

Preparation of oligonucleotides

Oligonucleotides were obtained as "trityl-off" derivatives. Dried single-stranded (ss) DNA was dissolved in 1 mL of 10 mM NH_4OAc prior to purification using a Beckman high performance liquid chromatography (HPLC) system as described

previously.[171] A C18 octadecylsilyl column (8 x 100 mm Waters Delta Pak Radial Pak Cartridge) was used in all purifications. A linear gradient of 0-60% aqueous acetonitrile in 10 mM NH_4OAc (35 minutes; 1 mL/min flow rate) was used to elute the ssDNA. The peak corresponding to the ssDNA was collected and freeze-dried using a Savant Speed Vac (Selby-Biolab, Australia). The dried ssDNA was then redissolved in MilliQ water and stored at -20 °C. DNA concentrations were determined from the UV absorbance at 260 nm, using the Beer-Lambert law. The molar extinction coefficient for each oligonucleotide was calculated using values of ε_{260} for adenine, guanine, cytosine and thymine of 15,200, 12,010, 7,050 and 8,400 $M^{-1}cm^{-1}$, respectively obtained from the website "Oligonucleotide Properties Calculator".[172]

Preparation of 16-mer double-stranded DNA (dsDNA)

Equimolar amounts of complementary single-stranded (ss) DNA strands were mixed together, freeze-dried and then redissolved using 50 μL of 100 mM NH_4OAc, pH 8.5, giving a final concentration of dsDNA of 1 mM. The mixture was then heated for 15 min in the water bath at 20 °C higher than the calculated melting temperature (Oligonucleotide Properties Calculator)[172] for dsDNA. The solution was allowed to cool slowly overnight to room temperature, giving 1 mM double-stranded (ds) DNA in 100 mM NH_4OAc, pH 8.5.

Titration of dsDNA with ruthenium complexes

These experiments were performed to determine the number of ruthenium complexes bound to dsDNA. Stock solutions of ruthenium complexes in 100 mM NH_4OAc, pH

8.5, were 200 μM. Reaction mixtures containing different ratios of ruthenium complex:dsDNA were prepared, giving a final concentration of dsDNA of 25 μM in 100 mM NH_4OAc, pH 8.5, and a final volume of 80 μL. The relative amounts of different ruthenium complexes added to the reaction mixtures are given in the relevant sections of the text. Prior to ESI-MS, the mixtures were diluted with 100 mM NH_4OAc, pH 8.5, giving a final concentration of dsDNA of 10 μM. ESI-mass spectra were obtained using a Micromass Q-ToF-2™ mass spectrometer (Wyntheshawe, UK) equipped with a Z-spray electrospray ionisation source and a quadrupole/ToF mass analyser. The instrument conditions are described in section 2.2.5.1. All samples were directly injected into the source of the mass spectrometer using a Harvard model 22 syringe pump (Natick, MA, USA) at a flow rate of 20 μL/min.

Competition for dsDNA among ruthenium compounds

Competition experiments were performed to determine the relative binding affinities of the six ruthenium complexes ($[Ru(phen)_3]^{2+}$, $[Ru(phen)_2(dpq)]^{2+}$, $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+)}$, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$) for different DNA duplexes. Each reaction mixture contained two of the ruthenium complexes, and the ratio of ruthenium 1: ruthenium 2: dsDNA was 3:3:1. Reaction mixtures were prepared by adding 2 μL of 1 mM dsDNA to 30 μL of two solutions each containing 200 μM ruthenium complex and 18 μL of 100 mM NH_4OAc, pH 8.5. The final concentration of dsDNA in reaction mixtures was 25 μM in a final volume of 80 μL. After allowing to stand overnight in the dark (23 °C), the reaction mixtures were diluted by adding 120 μL of 100 mM

NH_4OAc, pH 8.5, prior to analysis by ESI-MS. All mass spectra were obtained using a Waters extended mass range Q-ToF Ultima™ (Wyntheshawe, UK; see section 2.2.3 for a description of the instrument). All samples were directly injected into the source of the mass spectrometer using a Harvard model 22 syringe pump (Natick, MA, USA) at a flow rate of 20 μL/min; see section 2.2.3 for a description of the instrument).

Competition between ruthenium compounds and organic drugs

Competition experiments were performed in which the abilities of ruthenium compounds to compete for sites on dsDNA with either distamycin or daunomycin were tested. In most cases reaction mixtures contained a 10:6:1 ratio of organic drug: ruthenium compound: dsDNA, and were prepared by combining 2 μL of 1 mM dsDNA with 4 μL of 5 mM organic drug and the correct amount of 100 mM NH_4OAc, pH 8.5 and allowing the solution to stand at room temperature for 1 hr. After this period of time an appropriate amount of ruthenium compound was added and left to stand at room temperature overnight. Reaction mixtures involving $[Ru(phen)_3]^{2+}$ contained a 10:30:1 ratio of organic: ruthenium compound: dsDNA. The final dsDNA concentration was 25 μM in a final reaction volume of 80 μL. Table 2.1 shows the volume of reagents added to the reaction mixtures. ESI mass spectra were obtained 16 hours after addition of the ruthenium compound to the mixtures. Prior to ESI-MS, the reaction mixtures were diluted with 120 μL of 100 mM NH_4OAc, pH 8.5, giving a final concentration of dsDNA of 10 μM.

In a second set of experiments, reaction mixtures were prepared as described above, except that the order of addition of the ruthenium compound and organic drug was reversed. All mass spectra were obtained using a Waters extended mass range Q-ToF Ultima™ (Wyntheshawe, UK; see section 2.2.3 for a description of the instrument).

Table 2.1 Compositions of reaction mixtures used for competition experiments among ruthenium compounds and organic drugs

Ratio of organic drug:Ru:dsDNA	Volume 1 mM dsDNA (μL)	Volume organic drug (μL)	Volume ruthenium compound (μL)	Volume 100 mM NH_4OAc, pH 8.5 (μL)
10:6:1[a]	2	4 μL of 5 mM stock	60 μL of 200 μM stock	14
30:10:1[b]	2	5 μL of 5 mM stock	12 μL of 5 mM stock	61

[a] Reaction mixtures involving $[Ru(phen)_2(dpq)]^{2+}$, $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$.
[b] Reaction mixture involving $[Ru(phen)_3]^{2+}$.

Melting temperatures of drug-DNA complexes determined by UV spectroscopy

Reaction mixtures for melting temperature analysis contained ruthenium compounds and dsDNA at a ratio of 3:1. The final dsDNA concentration was 1 μM. The reaction mixtures were prepared by combining 10 μL of 1 mM dsDNA and 100 μL of 100 μM ruthenium complex in 100 mM NH_4OAc, pH 8.5, and making the volume up to 1 mL using 100 mM NH_4OAc, pH 8.5. The mixtures were left to stand at room temperature for 30 minutes before being analysed with the Thermal-UV software supplied with a Cary 500 UV-Vis NIR spectrophotometer (Varian, Mulgrave, Australia). Measurements of solution absorbance were performed using a wavelength

of 260 nm. The start and end temperatures were 25 °C and 80 °C, respectively, the ramping rate was 1 °C/min, data interval 0.3 °C, and filter size 101.

2.2.2 Preparation of proteins, protein-metal and protein-protein complexes

Determination of protein concentrations

Concentrations of proteins were determined spectrophotometrically using a Shimadzu PharmaSpec UV-1700 UV-Visible spectrophotometer (Shimadzu, Japan). The molar extinction coefficients, $\varepsilon_{280,}$ are shown in Table 2.2.

Table 2.2 Extinction coefficients (ε_{280}) used to determine protein concentrations.

Please see print copy for Table 2.2

Concentrations of DnaB, DnaB mutants and DnaC for studies of oligomerisation and complex formation with DnaC (chapter 5) were determined using a Bio-Rad DC protein assay. Standard solutions of BSA (Sigma, Australia), 0-20 mg/mL, were prepared from 100 mg/mL stock BSA solution in the same buffer used for DnaB (50 mM NH_4OAc, pH 7.6, 1 mM $Mg(OAc)_2$ and 0.1 mM ATP). Aliquots (3 μL) were taken from each standard and the protein stock solution. Bio-Rad reagent A (37.5 μL) was added to each standard or protein sample, vortexed and pulse centrifuged

before 300 μL of Bio-Rad reagent B was added. The mixtures were again vortexed and pulse centrifuged.

After 20 minutes at room temperature, mixtures were analysed using a Shimadzu 1700 PharmaSpec UV spectrophotometer set to 750 nm. The BSA standards were prepared and measured in duplicate. A standard curve was obtained from the average absorbance of the readings for the BSA standards, and concentrations of proteins were interpolated from the curve.

Metal ion binding to ε186

θ and ε subunits of *E. coli* DNA polymerase III were over-expressed in *E. coli* and purified in the laboratory of Dr Nicholas Dixon (Australian National University) as described previously by Keniry *et al*.[176] and Hamdan *et al*..[177] Proteins were stored at -80 °C in 25 mM Tris-HCl, pH 7.6, 2 mM DTT, 1 mM EDTA, 150 mM NaCl and 10% (*v/v*) glycerol. Prior to MS analysis, aliquots of proteins (50-120 μL) were taken from the frozen stock (-80 °C) and thawed on ice, then diluted to 300 μL with 10 mM NH_4OAc. The diluted proteins were then dialysed against 2 L of 10 mM NH_4OAc at 4 ºC overnight. The solutions were changed three more times over a period of 24 hours.

For ESI-MS analysis, three different metal ions: manganese, zinc and dysprosium were used to study the interaction of metal ions with ε186. An appropriate volume of either $Mn(OAc)_2$, $Zn(OAc)_2$ or $Dy(OAc)_3$ in 100 mM NH_4OAc, pH 8.0, was added to give different ratios of ε186: metal ion (ratios were in the range of 1:1 to 1:300).

The final protein concentration in solution mixtures was 2 μM in 100 mM NH_4OAc, pH 8.0. The final concentrations of metal ions were from 2 to 600 μM. The final volume of the mixtures was 100 μL. The mixtures were kept on ice for 15 minutes before injection directly into the mass spectrometer.

Spectrophotometric assay of ε186 activity

In a typical assay, stock solutions of 200 mM 5′-*p*-nitrophenyl ester of thymidine-5′-monophosphate (*p*NP-TMP, in assay buffer), 100 mM $Mn(OAc)_2$, $Zn(OAc)_2$ or $Dy(OAc)_3$ (in H_2O) were diluted with assay buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl) giving final concentrations of 0.05-10 mM, and 50 and 500 mM, respectively, in a 1 mL quartz cuvette. To start the reaction (*p*NP-TMP hydrolysis), an aliquot of ε186 was added, giving a final enzyme concentration of 0.1 μM. The solution was quickly and thoroughly mixed with a bent glass rod, and the production of *p*-nitrophenolate ion was followed at 420 nm for several minutes using a Cary 500 UV-Vis-NIR spectrophotometer (Varian, Mulgrave, Australia) at 25 ºC.[173] Initial rates (v_0) were measured from slopes of A_{420} curves versus time. Rates of *p*NP-TMP hydrolysis were calculated using a value of 12,950 $M^{-1}cm^{-1}$ for ε_{420} of *p*-nitrophenol at pH 8.0.[173]

Oligomerisation of DnaB and DnaB mutants

DnaB and DnaB mutants (F102E, F102H, F102W and D82N) were stored at -80°C in 50 mM Tris-HCl, 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. Prior to mass spectrometry analysis, an

aliquot of each protein (20-60 μL) was taken from the frozen stock (-80 °C), thawed on ice, then diluted with 400 μL of ice-cold 50 mM NH_4OAc, pH 7.6, 1 mM $Mg(OAc)_2$ and 0.1 mM ATP. Buffer exchange was then carried out using Millipore Biomax centrifugal filters (5,000 MWCO). The solution was centrifuged at 10,000 g using an Eppendorf 5415C bench top centrifuge (Crown Scientific, Moorebank, Australia) in a 4 °C cold cabinet. Three additional buffer exchanges were performed by adding 400 μL aliquots of the cold solution of 50 mM NH_4OAc, pH 7.6, 1 mM $Mg(OAc)_2$ and 0.1 mM ATP before protein samples were concentrated to ~20 μL. In experiments aimed at examining the formation of complexes of DnaB or DnaB mutants with DnaC, the DnaB and mutants were exchanged into a solution containing the same reagents except that the concentration of $Mg(OAc)_2$ was 0.1 mM.

Oligomerisation mixtures were prepared by adding appropriate volumes of 5 M NH_4OAc, pH 7.6, 10 mM $Mg(OAc)_2$ (in 50 mM NH_4OAc, pH 7.6) and 1 mM ATP (in 50 mM NH_4OAc, pH 7.6), followed by an addition of the appropriate amount of DnaB. The final protein concentration for MS analysis was 10 μM in 10 μL in different solvents. The final concentrations of NH_4OAc, Mg^{2+}, ATP and protein are given in the relevant sections of the text. The protein mixtures were then analysed using nanospray ESI-MS under the conditions listed in Table 2.4.

Formation of $(DnaB)_6(DnaC)_x$ complexes

DnaC is a protein that loads the helicase DnaB onto primed template DNA.[178] DnaC was stored at -80°C in 50 mM Tris-HCl, pH 7.6, 25 mM NaCl, 5 mM $MgCl_2$, 2 mM

DTT, 0.1 mM ATP, 1 mM EDTA and 20% (*v/v*) glycerol. To obtain nucleotide-free DnaC for ESI-MS analysis, 20 μL of thawed protein was diluted to 300 μL with 10 mM NH_4OAc, pH 5.5, and then dialysed overnight against 2 L of 10 mM NH_4OAc, pH 5.5 at 4 °C. The solutions were changed three more times with 2 L of 10 mM NH_4OAc, pH 7.6, over the next 24 hours.

For analysis of DnaB/DnaC complexes, DnaC was prepared as described above, then concentrated to ~20 μL using a Millipore Biomax centrifugal filter (5,000 MWCO), which had been washed twice with 10 mM NH_4OAc, pH 7.6. The protein concentration was then determined using a Bio-Rad DC protein assay as described above.

An appropriate volume of the concentrated DnaB was prepared to give a final concentration of 10 μM hexameric DnaB ($(DnaB)_6$) in 10 μL solution. Appropriate volumes of a solution containing 300 mM NH_4OAc, pH 7.6, 0.1 mM $Mg(OAc)_2$ and 0.1 mM ATP were added, before an appropriate volume of DnaC was added to give $(DnaB)_6$ to DnaC ratios of 1:1, 1:3, 1:6 and 1:8. Table 2.3 shows examples of the compositions of each solution mixture. NanoESI-MS analysis of (DnaB mutant)$_6$(DnaC)$_x$ complexes were carried out using solutions prepared in a similar fashion.

Hydrogen/deuterium (H/D) exchange of linear and cyclised DnaB-N

Deuterium oxide (D_2O) solution

D_2O solution was prepared by adding 50 μL of 3.0 M NH_4OAc in H_2O to deuterium oxide (99.9% D), giving a final volume of 15.0 mL. This produced 10 mM NH_4OAc in deuterium oxide. The pH was then adjusted to 7.2 by adding either solutions of 3% (*v/v*) ammonium hydroxide or 3% (*v/v*) acetic acid in deuterium oxide. During the preparation, nitrogen gas was constantly flushed over the solution.

Table 2.3 Examples of compositions of $(DnaB)_6(DnaC)_x$, $(F102W)_6(DnaC)_x$ or $(D82N)_6(DnaC)_x$ oligomerisation mixtures. The final concentration of the hexameric protein was 10 μM in 10 μL.

	Volumes (μL)		
$(DnaB)_6$:DnaC ratio	**Concentrated DnaB[a] (or F102W or D82N)**	**Solution containing 300 mM NH_4OAc, pH 7.6, 0.1 mM $Mg(OAc)_2$ and 0.1 mM ATP**	**Concentrated DnaC[b]**
1:0	2.5	7.5	-
1:1	2.5	7.05	0.45
1:3	2.5	6.14	1.36
1:6	2.5	4.77	2.73
1:8	2.5	3.86	3.64

[a] DnaB concentration = 40 μM.
[b] DnaC concentration = 220 μM.

Quenching method for analysis of H/D exchange by ESI-MS

DnaB is the replicative helicase of *E. coli*.[179,180] (section 5.1.1.1). The N-terminus (residues 24-136) is involved in the helicase activity. Linear and cyclised DnaB-N with linkers of varying lengths (3, 4, 5 and 9 amino acids) were expressed in *E. coli*

and purified in the laboratory of Dr Nicholas Dixon (Australian National University) as described by Williams *et al.*.[174] Stock DnaB-N with different linker lengths (linear or cyclised) were stored at -80 °C in 50 mM Tris-HCl, pH 7.6, 100 mM NaCl, 5 mM $MgCl_2$ and 15% (*v/v*) glycerol. An aliquot (20-40 μL) of each frozen protein stock was thawed on ice before dialysis against 2 L of 10 mM NH_4OAc, pH 7.2, at 4 °C overnight. The solution was changed three more times over a period of 24 hours. The dialysed proteins were then concentrated to ~20 μL using a Millipore Biomax centrifugal filter (5,000 MWCO), which had been pre-washed twice with 10 mM NH_4OAc, pH 7.2. The protein concentration was then determined (Table 2.2).

Stock solutions containing 1 mM protein (10 μL) were prepared in 10 mM NH_4OAc, pH 7.2. An aliquot of 1 mM protein solution and the D_2O solution above were equilibrated separately at 10 °C in a water bath for 10 minutes. A t = 0, an aliquot of the D_2O solution (99x the volume of the protein, 10 °C) was added to the protein solution giving a final protein concentration of 10 μM and a deuterium percentage of ~99%. The deuterium-protein mixtures were incubated in a 10 °C water bath. At specific time points, 4 μL of the deuterated protein solution were taken and diluted with 36 μL of ice-cold quenching solution (water:methanol:formic acid (90:9:1), pH 2.1). An aliquot (10 μL) of the quenched solution was then injected into the mass spectrometer through a Rheodyne injector with a 10 μL sample loop. The ice-cold quenching buffer was used as the mobile phase throughout the experiment and injected using a Harvard model 22 syringe pump (Natick, MA, USA) at a flow rate of 50 μL/min.

Another set of experiments was also performed using 100 mM NH_4OAc, pH 7.2, with the same percentage of deuterium.

2.2.3 Mass spectrometry

Conditions for mass spectrometry

Mass spectra were acquired using either a Micromass Q-ToF2™ (Wyntheshawe, UK) or a Waters extended mass range Q-ToF Ultima™ (Wyntheshawe, UK) ESI mass spectrometer equipped with a Z-spray probe. The mass spectrometers had a *m/z* range in the quadrupole of 4,000 and 32,000, respectively. Prior to analyses of samples, the mass spectrometers were externally calibrated using a solution of 1 mg/mL or 10 mg/mL (for high molecular mass using the Q-ToF Ultima) caesium iodide. Table 2.4 shows experimental conditions used for analyses of ruthenium-DNA and protein samples on the Micromass Q-ToF-2™ and/or the Waters extended mass range Q-ToF Ultima™ mass spectrometer.

Processing data

Typically 20-50 acquisitions (see Table 2.4) were combined to obtain a representative spectrum. Each spectrum was then background subtracted using a polynomial order of 11 with 40 % below the curve, and smoothed using a Savitzky Golay smoothing method, where 20 channels were smoothed twice.

For the analysis of the binding of metal ions to the protein ε186, and H/D exchange of DnaB-N, data processing was carried out as mentioned above and spectra were

centred and transformed to a mass scale using the transform function in the MassLynx software, in order to reduce complexity for data analysis.

Table 2.4 ESI-MS conditions used for the analysis of ruthenium-DNA and protein samples on the Micromass Q-ToF-2™ and the Waters extended mass range Q-ToF-Ultima™ mass spectrometer.

	Sample					
MS Parameters	**Ruthenium**[a]	**Ruthenium**[b1]	**Epsilon**[b1]	**DnaB-N**[b1]	**DnaC**[b1]	**DnaB ± DnaC**[b2]
Ion mode	- ve	- ve	+ ve	+ ve	+ ve	+ ve
Capillary (V)	2500	2500	2500	2500	2500	1500
Cone (V)	50	100	100	50	100	170
RF lens 1 energy (V)	N/A	70	55	120	100	190
Source block temperature (°C)	60	N/A	N/A	N/A	N/A	N/A
Desolvation temperature (°C)	80	100	100	120	120	N/A
Collision energy (V)	10	4	2	4	4	4
Transport /Aperture (V)	2/12	5/5	5/5	5/5	5/5	5/13
Acquisitions over mass range (*m/z*)	500-3000	500-3000	500-4500	500-3000	500-4500	500-20000
Number of acquisitions	50	40	40	3	20	20-30
Collision cell gas gauge (bar)	N/A	N/A	N/A	N/A	N/A	0.75
Ion optic region pressure (mbar)	N/A	3.6×10^{-3}	3.6×10^{-3}	3.6×10^{-3}	3.6×10^{-3}	1×10^{-1}

[a] ESI mass spectra were obtained using Q-ToF-2™.
[b1] ESI mass spectra were obtained using a Waters extended mass range Q-ToF Ultima™ (conventional electrospray).
[b2] ESI mass spectra were obtained using a Waters extended mass range Q-ToF Ultima™ (nanospray).

Chapter 3
Non-Covalent Interactions between DNA and Metallointercalators

3.1 Structure of DNA

Nucleic acids carry the genetic code which contains the information used to synthesise proteins. There are two types of nucleic acids, ribonucleic acid (RNA) and deoxyribonucleic acid (DNA).[181] Both RNA and DNA are polymeric molecules, consisting of thousands of repeating units known as nucleotides, connected by covalent bonds. The primary structure of nucleic acids is determined by the sequence of nucleotides along the polynucleotide chain (Figure 3.1). Each nucleotide comprises a five-carbon sugar (ribose in RNA, 2′-deoxyribose in DNA), a phosphate group and a purine or pyrimidine base. In both types of nucleic acids, successive nucleotides are connected by phosphodiester bonds involving the 5′-carbon of one sugar unit and the 3′-carbon of the next. The phosphate groups in both RNA and DNA are strong acids, and are therefore deprotonated at physiological pH.

Before the secondary structure of DNA was known, there were questions as to how it could carry the enormous amount of information required to recreate life, how this information could be processed, and most important of all how it could accurately be replicated during cell division.

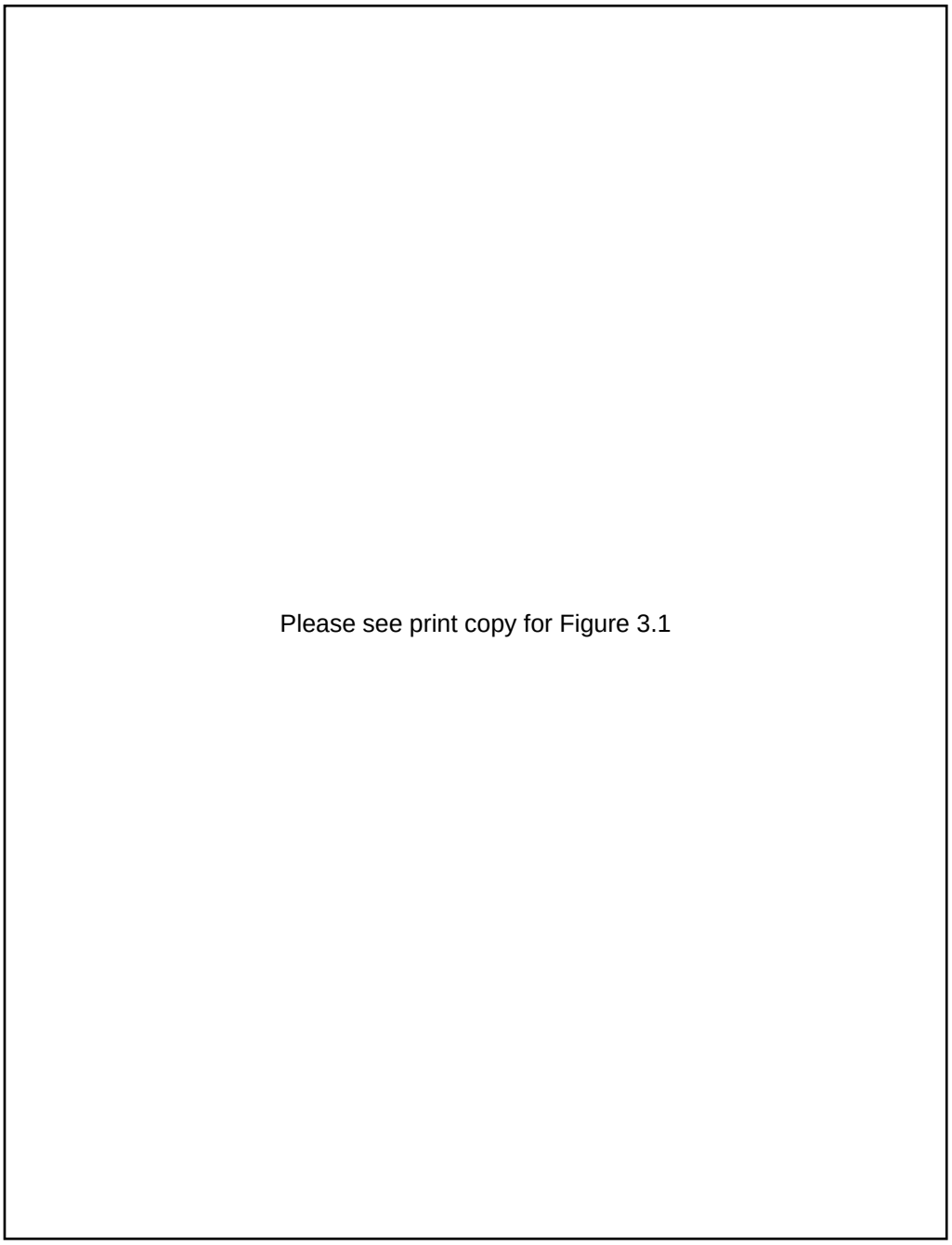


Figure 3.1 Essential features of the structure of double-stranded (ds) DNA. Each successive nucleotide is connected through a phophodiester linkage, involving the 5′-carbon of one sugar unit and the 3′-carbon of the next. The two antiparallel strands, each running in the 5′ to 3′ direction, are held together by hydrogen bonds between A-T and G-C base pairs.[182] The structures of the purine and pyrimidine bases are also shown.

These questions were finally answered by the elucidation of the three dimensional structure of DNA in 1953 by Watson and Crick from analysis of X-ray diffraction data.[183] Their discovery not only provided a physical picture of DNA, it also explained other data that had not been understood until then. DNA is a double-stranded polymer consisting of two antiparallel polynucleotide chains that are joined together by hydrogen bonds between complementary purine and pyrimidine base pairs. In DNA adenine will only pair with thymine (T), and cytosine (C) will only pair with guanine (G) as shown in Figure 3.1.

There are three principal conformations of double-stranded DNA: the A, B and Z forms (see Figure 3.2). All three conformations have the same composition, and only differ in their secondary structures.[184] Under physiological conditions, the most common form of DNA is B-DNA, which is the form that was first described by Watson and Crick.[183] B-DNA is a right-handed helix, ~20 Å in diameter. It contains 10 base pairs (bp) per full helical turn, and has a helical twist of 36° per bp. The partial stacking of aromatic bases results in a pitch or rise per helical turn of 34°, with a 6° tilt from the axis of the double helix. Each base pair in B-DNA has approximately the same width, resulting in almost perfect symmetry for the DNA molecule. This leads to the possibility of base pairs being interchangeable at any position in the DNA molecule. For example an A-T base pair can be changed to T-A base pair, or G-C to C-G without disrupting the sugar phosphate backbone.[184] B-DNA has a wide and deep major groove, and a narrow and deep minor groove (Figure 3.2).

DNA has been shown to adopt different conformations depending on both its sequence and environmental conditions. Under dehydrating conditions, B-DNA undergoes a reversible conformational change to A-DNA,[185] which is also a right-handed helix that is ~26 Å in diameter. It contains 11 bp per helical turn and has a pitch of 25 Å. An interesting feature of A-DNA is that the planes of its base pairs are tilted by ~20° with respect to the helix axis. Furthermore, since the helix axis does not pass through its base pairs, A-DNA has a deep major groove and a rather shallow minor groove, resulting in a wider and flatter helix than B-DNA. The A form of nucleic acid structure is also present when DNA is base paired with RNA and in dsRNA.[186]

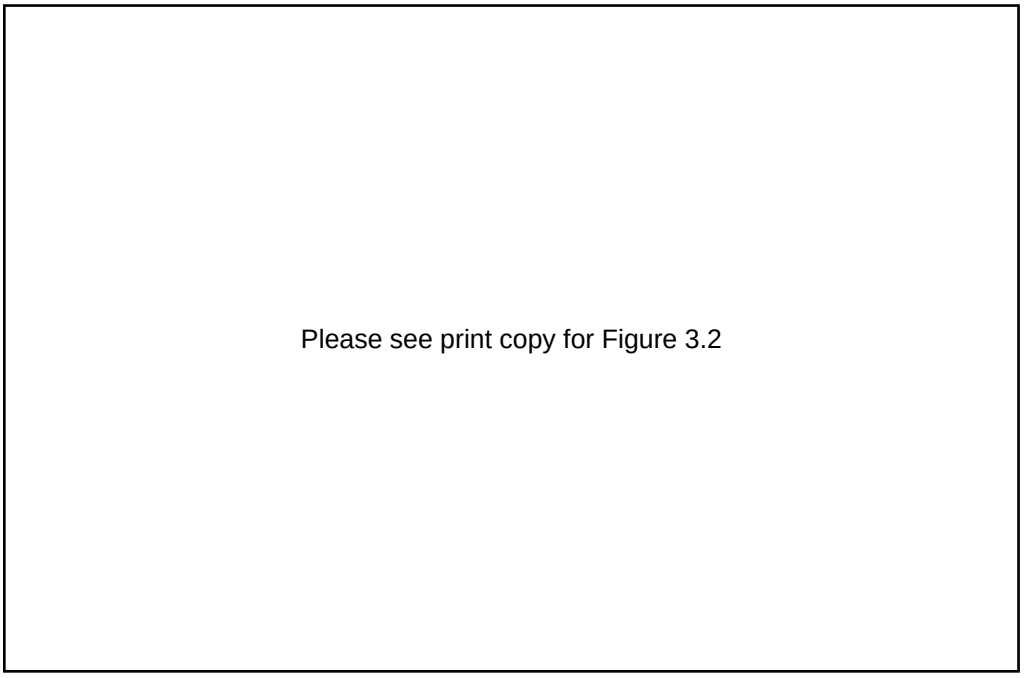


Figure 3.2 The A-, B- and Z-conformations of DNA.[187]

Z-DNA has an unusual structure even though it also primarily consists of two antiparallel strands connected by hydrogen-bonding between base pairs. Z-DNA has a left-handed helical structure and is formed in solutions with high salt

concentrations.[188] Z-DNA received its name from the zig-zag conformation that characterises the sugar-phosphate backbone. It contains 12 bp per helical turn, and has a diameter of ~18 Å. Z-DNA has a pitch of ~45 Å, a 60° helical twist per bp, a flat major groove and deep minor groove. Z-DNA structure also occurs with DNA sequences containing alternating sequences of purines and pyrimidines such as in the promoter regions of some genes.[188-190] The presence of these three different conformations of DNA in solution under different conditions has been confirmed by various NMR studies.[191-193]

Both RNA and DNA contain coded information, in the sequence of their purine and pyrimidine bases, for the synthesis of proteins. They can interact reversibly or irreversibly with a wide range of chemical species including water, metal ions and small organic molecules. The recognition and binding of ligands or drugs by DNA usually occurs in the floor of either the major or minor groove, where unique environments are generated.[194] The floor of these grooves differs depending on the base-pair sequence present. This provides variations in hydrogen bond donor and acceptor sites, and gives rise to the ability of DNA to bind selectively in some instances to different ligands or drugs. The sugar-phosphate backbone lies along the exterior of the DNA helix, and is polyanionic. This provides additional atoms for interactions with polar atoms in ligands. Base stacking causes the core of DNA to be hydrophobic, providing an environment for non-polar interactions. Section 3.2 describes small molecules that bind to nucleic acids, in particular DNA, and their modes of binding.

3.2 DNA-Drug Interactions

DNA is a potential target for drugs designed to alter a variety of biological functions. Consequently there is considerable interest in studying the interactions of different drugs with DNA. A difficulty in using DNA as a target for drug interactions is that there is a risk of disturbing normal cellular functions. This generally limits the use of such drugs to very serious, life threatening diseases such as cancer. However, greater understanding of the mechanisms by which small molecules, including metal complexes, peptides, and organic compounds interact with DNA may facilitate the development of new therapeutic agents and approaches. For example, transcription therapy is the treatment of cancer by blocking or re-activating transcription factors displaying aberrant behaviour owing to inherent or acquired damage to DNA.[195,196] The attractiveness of this approach lies in recognising that while the origins of cancer are numerous, its onset is a result of incorrect processing of information through a relatively small number of signalling pathways controlled by a limited number of transcription factors.[196,197] Recently it was shown that a rhodium metallointercalator can competitively inhibit the binding of a transcription factor to its DNA recognition element,[198] providing strong support for exploring the potential of transcription therapy using this general class of compounds. In order to design drugs for transcription therapy or other applications, it is therefore important to understand the different types of non-covalent interactions that can take place between the drugs and DNA.

3.2.1 Covalent (irreversible) binding

Covalent binding to DNA generally involves the nitrogen and oxygen atoms in purine bases, particularly the N7 of guanine, which is the most nucleophilic site.[199] There are many molecules that can bind covalently to DNA. These include cisplatin (a coordination compound), mechlorethamine (an alkylating agent) and hedamycin (an intercalative alkylator), see Figure 3.3.

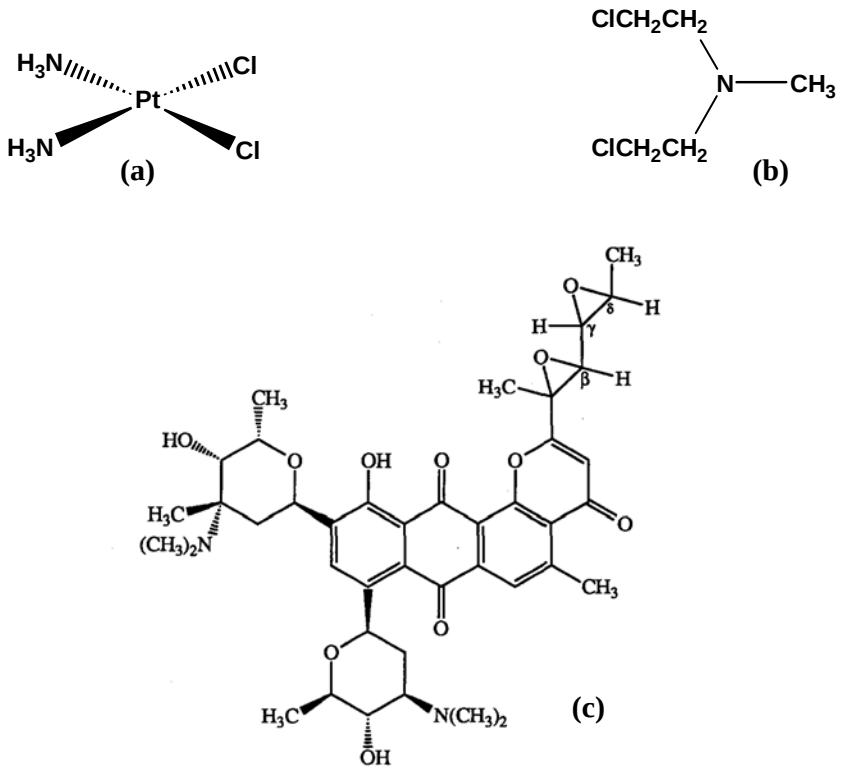


Figure 3.3 Examples of small molecules that covalently bind to DNA. (a) Cisplatin, a coordination compound, (b) mechlorethamine, an alkylating agent, and (c) hedamycin, an intercalating alkylator.

The potential anticancer properties of platinum complexes was first discovered serendipitously in 1965 by Rosenberg and co-workers during a study of the effect of electric fields on *E. coli* cell growth.[200] They found that cell division was inhibited by platinum complexes which formed by reaction of the platinum electrode with the

bacterial growth medium. Later it was shown that *cis*-diamminedichloroplatinum(II) (cisplatin, Figure 3.3 (a)), which had been known for over 100 years, also inhibited division of *E. coli* cells. In 1971 cisplatin entered clinical trials (phase I), and later on became a major drug for treatment of several human malignancies, including testicular,[201,202] ovarian,[203,204] bladder[205] and head and neck cancers.[206-208] Cisplatin binds to DNA preferentially at the N7 atoms of two adjacent guanine (GG) bases on the same DNA strand, forming an intrastrand crosslink. It also forms a significant proportion of intrastrand crosslinks in which platinum is bonded to the N7 atoms of both bases in a 5′-AG-3′ sequence, and much smaller amounts of crosslinks involving two guanine residues separated by a third base on the same DNA strand.[209] Interstrand crosslinks involving guanine N7 atoms on different DNA strands are also formed in small quantities. Coordination of cisplatin to DNA results in significant destabilisation of the double helix. NMR spectroscopy, gel electrophoresis and X-ray crystallography revealed cisplatin binding results in DNA being unwound by ~13° and bent by ~34-60° towards the major groove.[210] Furthermore, spectroscopic and calorimetric studies showed that there was a DNA conformational change from the B-form to the A-form and a reduction in thermal stability (melting temperature lowered by ~9 °C) and thermodynamic stability (lowered by 6.3 kcal/mol).[211]

Nitrogen mustards, such as mechlorethamine (Figure 3.3 (b)), chlorambucil and melphalan, are widely used in clinical treatment of lymphoma, leukaemia, multiple myeloma and ovarian carcinoma.[212] The cytotoxic and anticancer activity of nitrogen mustards correlate closely with the formation of DNA-DNA cross-links. These bifunctional lesions are capable of blocking DNA replication and transcription, which leads to inhibition of tumour growth and cell death.[213,214] Mechlorethamine is

a bifunctional alkylating agent which is known to undergo intramolecular nucleophilic substitution reactions, producing an aziridium ion intermediate. This reactive intermediate is capable of alkylating multiple sites in DNA such as N7 of guanine, and the N1, N3 and N7 of adenine.[215-218] Of these sites it has been shown that it is the N7 position of guanine that mechlorethamine prefers to bind to.[219] Mechlorethamine has also been shown to link distal guanine bases in the opposite strands of the 5′-GGC-3′ sequences (interstrand cross-linking).[217,219] This type of guanine-guanine cross-link is thought to be responsible for the cytotoxicity of mechlorethamine and other nitrogen mustards.[214]

Hedamycin (Figure 3.3 (c)) is a naturally occurring antitumour antibiotic which can both intercalate into, and alkylate dsDNA.[220-222] Intercalation of hedamycin into DNA involves positioning of carbohydrate substituents into both the major and minor grooves, while alkylation occurs via epoxide-mediated nucleophilic attack at the N7 atom of a guanine residue, located on the 3′ side of the drug molecule.[222] Many studies including gel electrophoresis, NMR and ESI-MS studies have shown that alkylation is sequence specific,[171,222-225] with hedamycin showing a preference for binding to the guanine in 5′-CGT sequences, and to a lesser extent in 5′-CGG sequences.[221,224]

3.2.2 Non-covalent (reversible) binding

DNA offers a variety of binding sites and binding modes for non-covalent interactions with small molecules. The three most important modes are: (i) electrostatic interactions, (ii) groove binding and (iii) intercalation. Some drugs use

more than one of these binding modes to maximise their DNA-binding affinity and selectivity.

Electrostatic interactions occur between cationic metal ions or ligands and the polyanionic phosphate backbone of DNA. Groove binding involves direct interactions between a metal complex or organic molecule and functional groups present on the edge of the base pairs in either the major or minor groove of DNA.[226] The major and minor grooves differ significantly in many aspects, including size, electrostatic potential, hydrogen bonding capabilities, steric properties and level of hydration.[226] These differences enable groove binding molecules to be separated into different classes which have different characteristics. For instance, many oligonucleotides and protein molecules bind specifically to the major groove, and are therefore classed as major groove binders, whilst smaller molecules prefer to bind to the minor groove, and are classed as minor groove binding ligands. The latter ligands often prefer to bind in the minor groove region because it allows them to maximise their H-bonding and van der Waals interactions. A characteristic of minor groove binders is that they contain aromatic rings with torsional freedom allowing the molecules to twist and adjust to the shape of the minor groove. The crescent shape of these molecules (e.g. distamycin, netropsin, Hoechst 33258; Figure 3.4) also matches closely the shape of the floor of the minor groove.

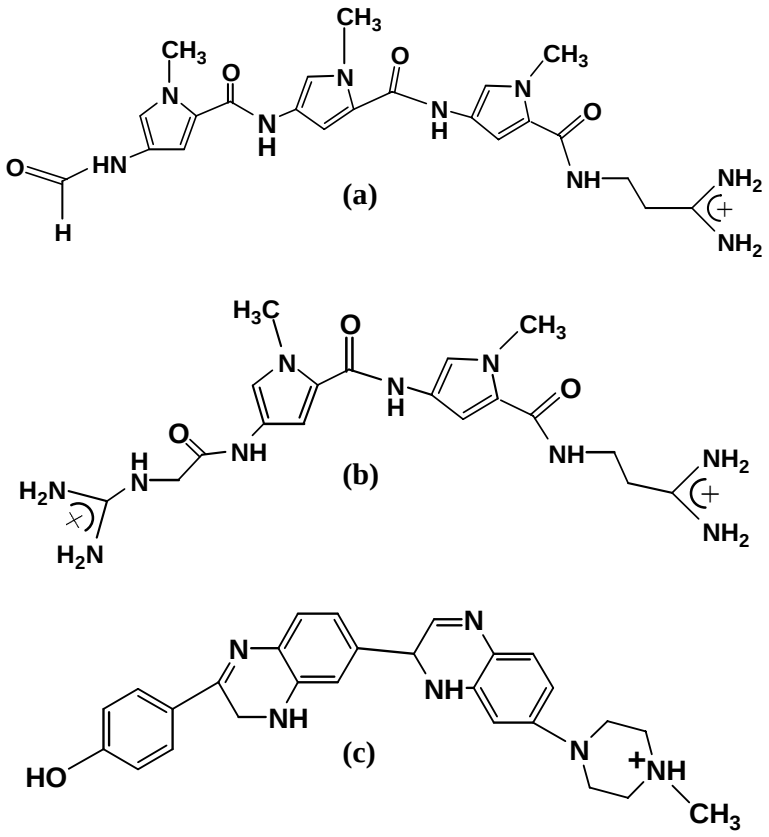


Figure 3.4 Structures of well known minor groove binders. (a) Distamycin, (b) netropsin, and (c) Hoechst 33258.

Figure 3.5 (a) shows an image of distamycin bound to the minor groove of DNA based on an X-ray crystal structure.[227] Distamycin is a tripeptide containing three *N*-methylpyrrole units joined together giving a curved shape that facilitates binding to the minor groove of DNA. Many studies have shown that distamycin prefers to bind to AT-rich DNA sequences containing at least four base pairs.[228-231] In addition, it has been shown that the minor groove can expand to accommodate two distamycin molecules lying side by side.[232] Other minor groove binders also show a preference for binding to DNA sequences containing at least four consecutive AT base pairs.[233] The positively charged end of distamycin usually lies in the bottom of the minor groove near the N3 of an adenine base, which has the greatest negative electrostatic

potential. In contrast, the exocyclic amino groups of guanine residues provide steric hindrance that prevents drug binding to GC-rich DNA sequences.[234,235]

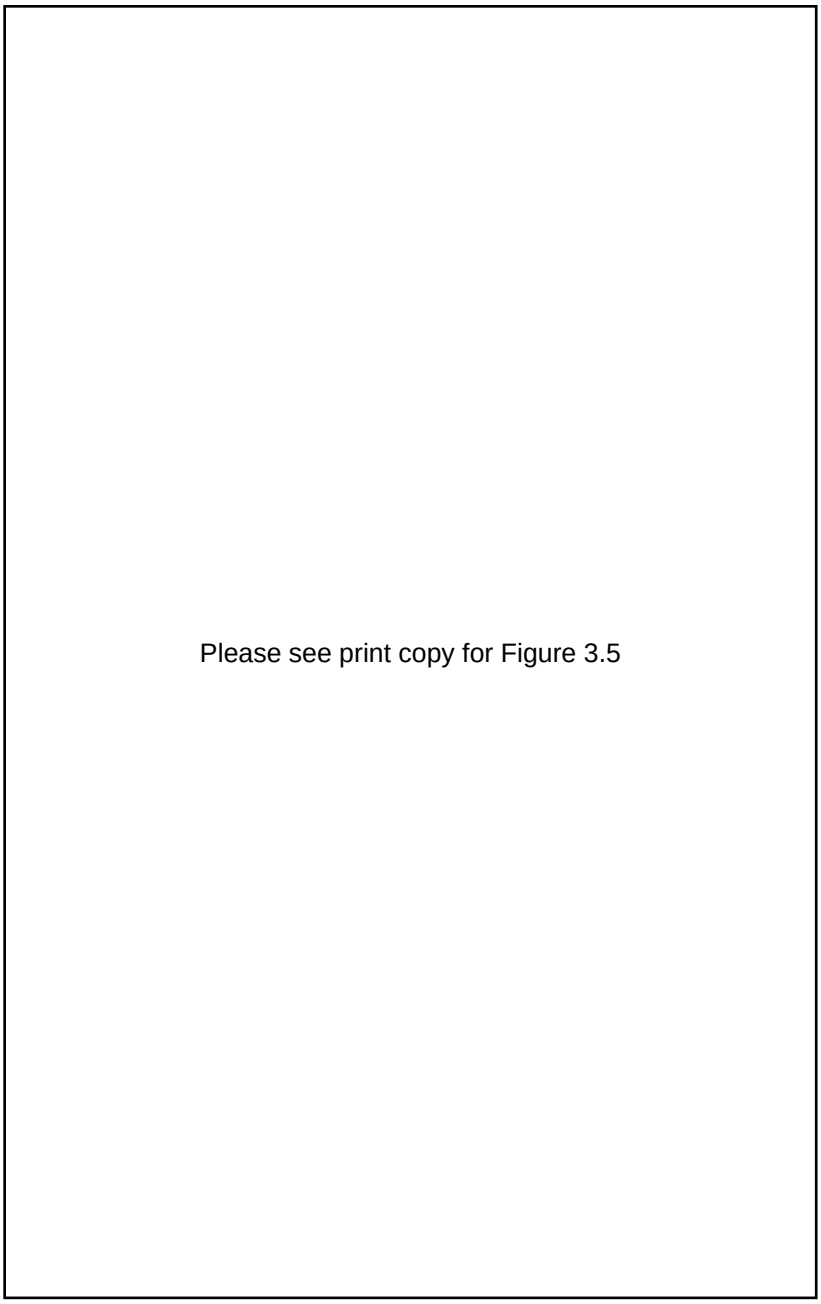


Figure 3.5 X-ray crystallographic structures of: (a) Distamycin bound to the minor groove of d(GGCCAATTGG)$_2$ (Protein data bank 1jtl),[227]and (b) a complex of two anthracycline molecules with d(CGATCG)$_2$, (Protein data bank 1nab).[236] The dsDNA is shown in blue, adenine and thymine bases are in orange and red, respectively, and drug molecules are represented by green balls and sticks.

Intercalation is a common mode for non-covalent binding of small molecules to DNA. Intercalators generally have polycyclic planar aromatic or heterocyclic ring systems and are positively charged. These properties allow intercalators to insert and stack in between base pairs in the hydrophobic interior of helical dsDNA. In addition, electrostatic interactions and hydrogen bonding may also play an important role in stabilising the overall binding interaction.[237-239] In general, intercalation occurs from the major groove of DNA and increases the vertical separation of adjacent base pairs. As a result, the DNA helix is distorted. In order to compensate for such disruption, the sugar-phosphate backbone is lengthened, resulting in partial unwinding of the supercoiled helix.[240-242]

Figure 3.6 shows the structures of some intercalators. Ethidium (Figure 3.6 (a)) is a small, simple organic intercalator that is often included as a reference compound in studies involving novel intercalators. It is widely used in molecular biology laboratories as a fluorescent stain to visualise DNA on agarose gels. Daunomycin (daunorubicin) and adriamycin (doxorubicin), shown in Figure 3.6 (b) and (c), respectively, are examples of more complex intercalators that exhibit high selectivity when binding to DNA, and as a result are used clinically as chemotherapeutic agents.[243,244] Daunomycin contains an anthracycline ring system that can intercalate into DNA, and an amino sugar that resides in and interacts with the minor groove. Daunomycin has been shown to bind preferentially to right-handed B form DNA and displays an increase in binding affinity towards DNA sequences with greater GC content.[245-247] Footprinting titration experiments have shown that triplet sequences containing GC or CG base pairs with either an A or T at the 5′-end are preferred binding sites.[248] X-ray diffraction analysis confirms direct intercalation of

daunomycin between GC base pairs.[160] Figure 3.5 (b) shows the X-ray crystal structure of a complex in which two anthracycline molecules are intercalated into a dsDNA base stack.[236]

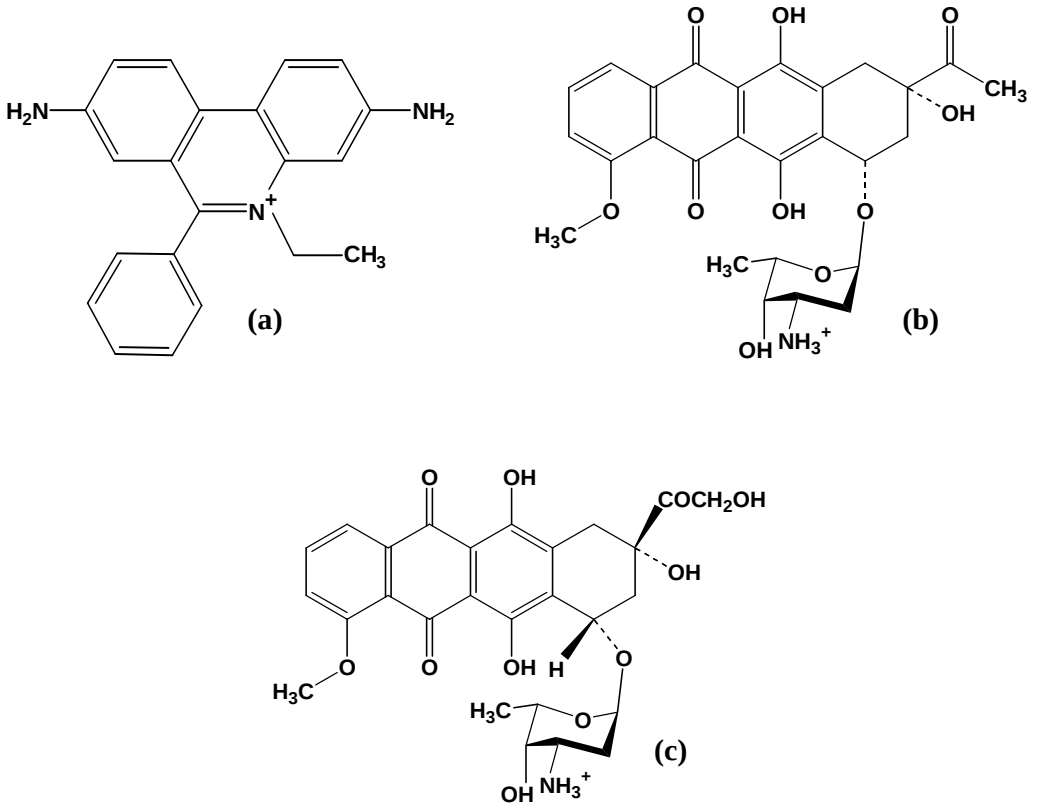


Figure 3.6 Structures of some intercalators. (a) Ethidium, (b) daunomycin, and (c) adriamycin.

3.3 Transition Metal Complexes

Designing small molecules that can bind and interact with specific sequences of DNA has increasingly become important. This approach may not only lead to novel chemotherapeutics, but also to reagents that detect certain DNA structures and could potentially be used as highly sensitive diagnostic agents.[245] The discovery of the anticancer activity of cisplatin over two decades ago[200] triggered renewed interest in

using transition metal complexes as anticancer drugs, or as other types of therapeutic or diagnostic agents.[249]

Many transition metal complexes display properties that can be exploited so they can be used as DNA probes to gain a greater understanding of how molecules interact with DNA.[250] These properties include stability, inertness and water solubility. Furthermore, many transition metals have favourable spectroscopic properties that enable them to act as "reporter" molecules after binding to DNA. Ruthenium complexes are amongst the most widely used transition metal complexes in studies of DNA recognition. This is because in addition to forming complexes that are kinetically and thermodynamically stable, ruthenium ions can bind to many different types of ligands.

There are several different ways that ruthenium complexes can interact with DNA. Since many ruthenium complexes are positively charged, they are often able to non-specifically bind to the negatively charged phosphate backbone of DNA through electrostatic interactions.[251] Ruthenium complexes containing ligands such as chloride, water or dimethylsulfoxide are able to bind covalently to DNA, while those containing planar, aromatic ligands can intercalate between DNA base pairs. Ruthenium complexes such as $[Ru(phen)_3]^{2+}$, which contains three bidentate ligands, are also chiral, and it has been shown that they can interact in an enantioselective manner with chiral B-form DNA (see Figure 3.7).[252] In the section below, the DNA-binding properties of ruthenium-based metallointercalators are further discussed.

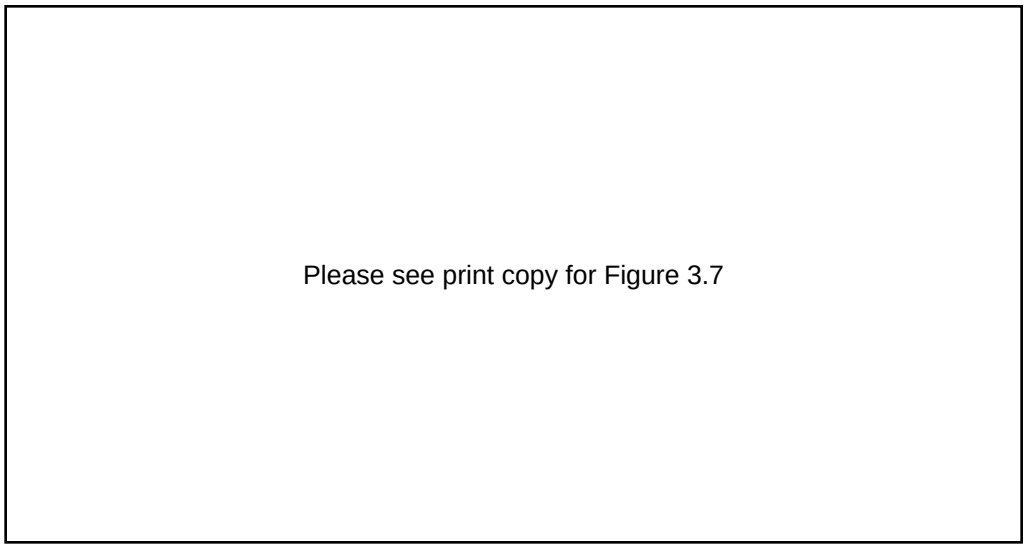


Figure 3.7 Enantioselective interactions of a ruthenium compound with B-DNA. (a) Λ-$[Ru(phen)_3]^{2+}$ and (b) Δ-$[Ru(phen)_3]^{2+}$. The Δ-isomer easily fits into the right-handed DNA groove, while the Λ-isomer does not properly fit owing to steric interactions with the DNA phosphate backbone. Adapted from Barton *et al*..[253]

3.4 Interactions of Ruthenium-Based Intercalators with dsDNA

The first studies on metallointercalators by Lippard and co-workers investigated square-planar platinum(II) complexes containing aromatic terpyridyl or phenanthroline ligands.[254] Later studies of metallointercalators focused on the binding of tris(phenanthroline) complexes of zinc, cobalt and ruthenium to DNA.[255-258] However, in recent years, ruthenium(II) complexes have been of more interest because of: (i) their kinetic inertness (owing to the low spin d^6 electronic state),[259,260] and (ii) the sensitivity of their photophysical properties, most notably an intense metal-to-ligand charge transfer (MLCT) band in the visible spectrum, to DNA binding.[261] Another reason for the interest in ruthenium complexes compared to, for example, platinum(II) complexes is that the greater size of these octahedral

molecules can result in a greater number of intermolecular interactions with DNA, therefore providing more opportunities for selective molecular recognition. Furthermore, the discovery that $[Ru(phen)_3]^{2+}$ and related complexes can bind non-covalently in an enantioselective manner to B-DNA (Figure 3.7),[252] raised other possibilities for selective binding to various DNA structures. For example, one enantiomer might favour binding to the left-handed Z-DNA helix that can occur in some DNA sequences,[188-190] over the right-handed B-DNA helix.

Initially the two enantiomers of $[Ru(phen)_3]^{2+}$ were suggested by Barton and co-workers to bind to DNA in the major groove via two different modes: intercalation of a single phenanthroline ligand into the DNA base stack, and surface binding along the major groove.[261] Early studies showed that Δ-$[Ru(phen)_3]^{2+}$ binds more tightly to B-DNA than its isomeric counterpart, Λ-$[Ru(phen)_3]^{2+}$. Later studies using viscosity measurements suggested that neither enantiomer binds to DNA like a classical intercalator.[262,263] Debate over the mechanism of DNA binding of the isomers of $[Ru(phen)_3]^{2+}$ continues, with some studies (using two dimensional proton NMR spectroscopy and CD spectroscopy) suggesting that they interact with DNA in the minor groove in a fashion independent of the sequence of DNA bases present.[258,264-267] In one of these studies the Δ-$[Ru(phen)_3]^{2+}$ enantiomer was found to bind to DNA by insertion of two phenanthroline ligands into the minor groove, whereas Λ-$[Ru(phen)_3]^{2+}$ was bound to the minor groove by a different mode that left the DNA structure unaffected.[264]

Another study (using normal absorption, linear and circular dichroism spectroscopy and computer modelling) suggested that the mechanism of binding is dependent on

the DNA sequence and, for the Δ-isomer, the relative concentrations of metal complex and DNA.[268] For the Λ-isomer, a single phenanthroline ligand was proposed to be inserted parallel to the base pair planes in the major groove in a binding mode referred to as partial insertion. For the Δ-isomer, when the metal:DNA ratio was below 1:4-6, two phenanthroline ligands of $[Ru(phen)_3]^{2+}$ were found to bind in the minor groove. In contrast, when the ratio was higher, the most favourable complex was that in which one phenanthroline ligand was inserted along the minor groove, while with some Δ-complexes there was partial insertion of a ligand into the major groove.[268]

Many more studies have been conducted to investigate the DNA binding properties of ruthenium metallointercalators. Some of these studies have centred on complexes with the general formula $[Ru(phen)_2L]^{2+}$, where L is an extended planar aromatic ligand capable of intercalating deeply into the DNA base stack and increasing the overall strength of binding interactions with DNA.[269-273] As the surface area of the unique ligand L is increased, so does the strength of intercalative binding to DNA. As a result, metallointercalators that contain extended aromatic heterocyclic ligands may be powerful tools for probing nucleic acid structure. The structures of some of the most widely studied ruthenium complexes of this type, including those used in this work, are illustrated in Figure 3.8 ((a)-(f)), together with the structure of $[Ru(phen)_3]^{2+}$.

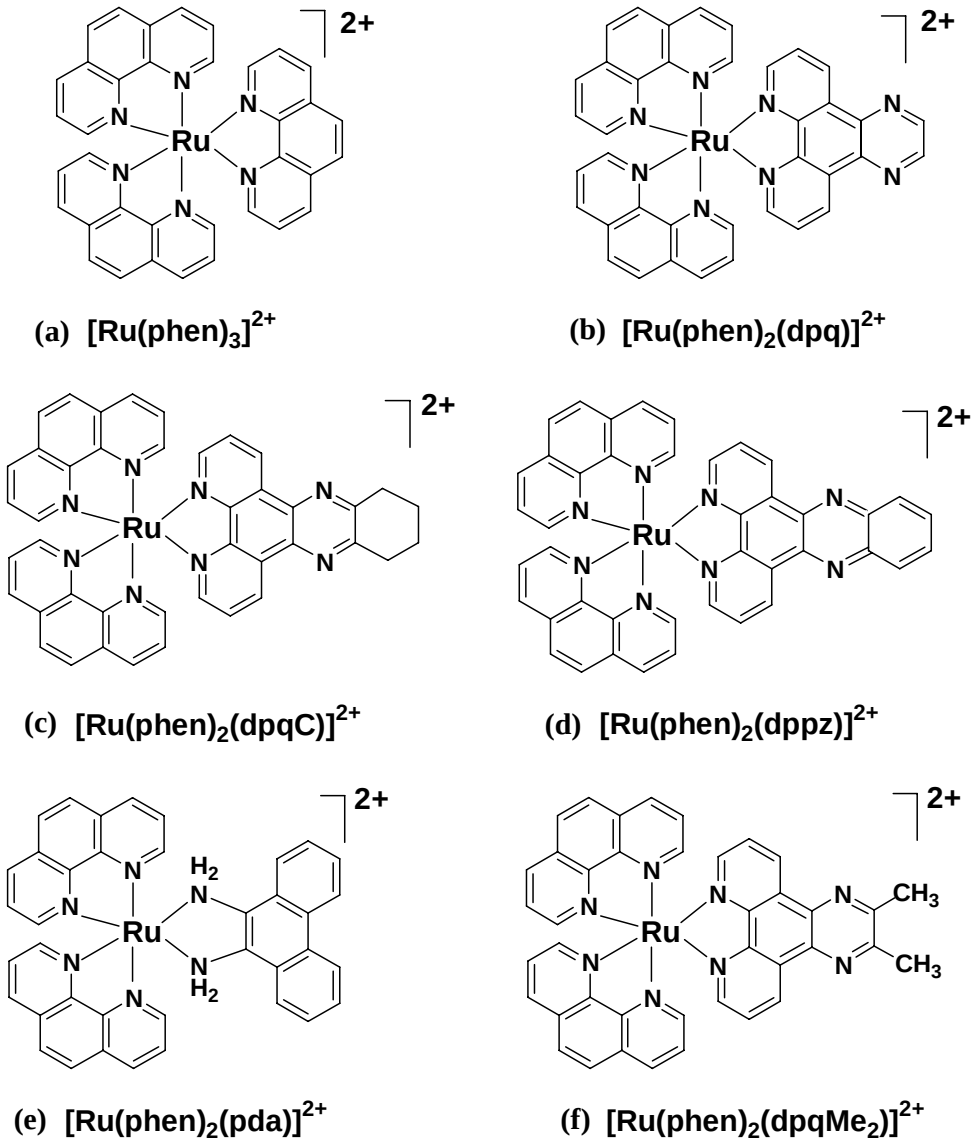


Figure 3.8 Structures of ruthenium metallointercalators used in this study. (a) $[Ru(phen)_3]^{2+}$; (b) $[Ru(phen)_2(dpq)]^{2+}$; (c) $[Ru(phen)_2(dpqC)]^{2+}$; (d) $[Ru(phen)_2(dppz)]^{2+}$; (e) $[Ru(phen)_2(pda)]^{2+}$ and (f) $[Ru(phen)_2(dpqMe_2)]^{2+}$.

The dppz ligand (dppz = dipyrido[3,2-*a*:2′,3′-c]phenazine) has an extremely large surface area. As a result ruthenium complexes containing this ligand generally have greater affinity for DNA ($>10^6$ M^{-1})[274-277] compared to complexes containing the other ligands shown in Figure 3.8. Binding affinities to calf thymus DNA have been obtained using a variety of techniques, including fluorescence spectroscopy.

For example, both $[Ru(bpy)_2(dppz)]^{2+}$ (bpy = 2,2′-bipyridine) and $[Ru(phen)_2(dppz)]^{2+}$ have been shown to display luminescence in non-aqueous solvents.[277,278] However, in aqueous solution, these complexes do not luminesce as a result of the ability of water molecules to deactivate the excited state through hydrogen bonding with the intercalating ligands.[277-279] Upon binding of these metal complexes to DNA, photoluminescence is observed. This is a consequence of intercalation of the dppz ligand into the DNA base stack, which prevents the nitrogen atoms on the intercalating ligand from being protonated by the surrounding aqueous solvent, and consequently increases the excited-state lifetime.[277] This effect has been extensively characterised and described as the "molecular light switch" effect, see Figure 3.9. It has been suggested that this property might serve as the basis for some ruthenium complexes acting as photophysical probes of nucleic acid structure or conformation.[250]

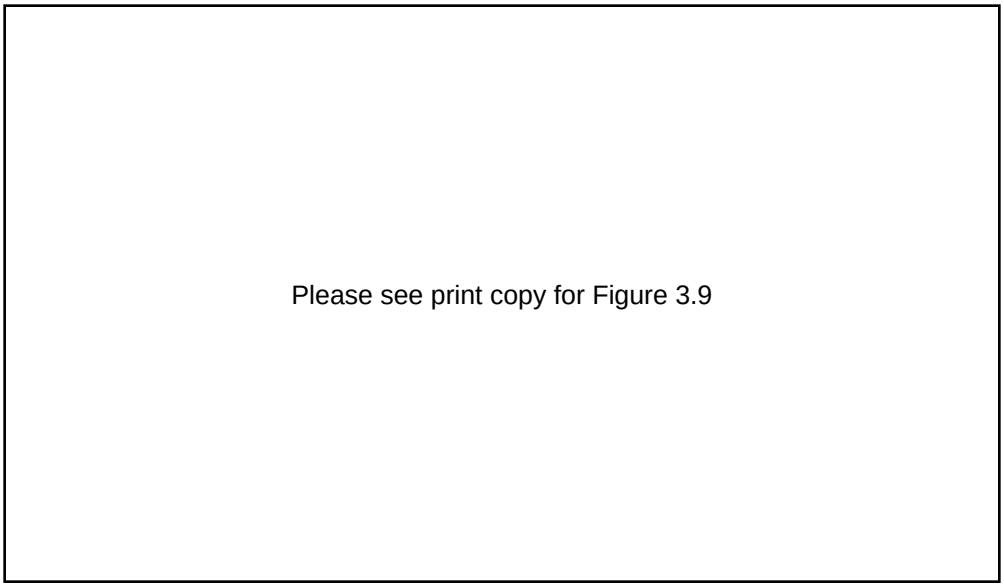


Figure 3.9 The "molecular light switch" effect as displayed by $[Ru(bpy)_2(dppz)]^{2+}$. In aqueous solution luminescence is quenched (lower spectrum). Upon binding to calf thymus DNA, the intercalating ligand is protected from solvent quenching, resulting in the observation of luminescence. Adapted from Friedman *et al*..[277]

Photophysical experiments have also provided evidence that the Δ-enantiomer of $[Ru(phen)_2(dppz)]^{2+}$ intercalates more deeply into right-handed B-DNA than the Λ-isomer, and consequently binds more tightly.[280] Fluorescence studies by Holmlin *et al.* showed that Δ- $[Ru(phen)_2(dppz)]^{2+}$ displays a slight, but significant, preference for AT-rich DNA sequences.[273] The complexes $[Ru(phen)_2(dpq)]^{2+}$ and $[Ru(phen)_2(dpqC)]^{2+}$ both contain chelating ligands consisting of an extended aromatic ring system similar to that in dppz, which enables then to intercalate into the DNA base stack. NMR spectroscopic studies showed that $[Ru(phen)_2(dpq)]^{2+}$ binds by intercalation to DNA via the minor groove, with a preference for purine-purine/pyrimidine-pyrimidine sequences.[269,270] Even though there has been no detailed binding study for $[Ru(phen)_2(dpqC)]^{2+}$, preliminary NMR experiments suggested that it also binds to DNA by intercalation from the minor groove, in a similar fashion to $[Ru(phen)_2(dpq)]^{2+}$.[269]

3.5 Applications of Ruthenium and Other Metal-Based Metallointercalators

Ruthenium metallointercalators have been suggested to be useful for a number of applications, including nucleic acid probes,[281,282] synthetic restriction enzymes,[250,283] DNA repair agents,[284-286] and potential therapeutic agents in gene regulation.[198] The potential of metallointercalators to be employed as synthetic restriction enzymes was demonstrated by construction of a metallointercalator-peptide chimera (Figure 3.10) by Barton and co-workers.[283] The chimera consisted of a Zn^{2+}-coordinated peptide tethered covalently to the rhodium metallointercalator $[Rh(phi)_2bpy]^{3+}$. When the

chimera became bound to DNA, the Zn^{2+} performed hydrolytic cleavage with some sequence specificity.

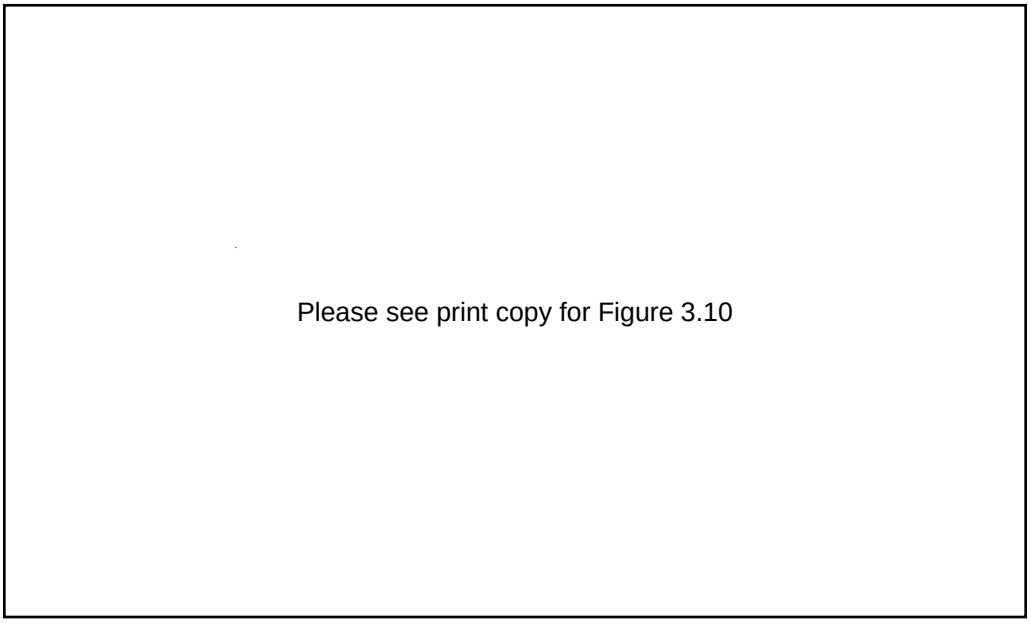


Figure 3.10 Structure of the Zn^{2+}-coordinated metallointercalator-peptide chimera, used as a synthetic restriction enzyme by Barton and co-workers.[250]

In another study a rhodium metallointercalator was shown to have the ability to repair DNA damaged by external agents such as ultraviolet radiation.[284,286] Exposure to UV radiation can damage genetic material and lead to mutations and cancer. The thymine dimer is the most common photochemical lesion in DNA. Barton and co-workers showed that when a rhodium metallointercalator became non-covalently bound to a DNA duplex containing a thymine dimer lesion, and was subsequently irradiated with 400 nm light, oxidative repair occurred, most likely by charge transfer from the tethered rhodium intercalator (Figure 3.11).[285,286] If a metallointercalator such as this could be selectively targeted to damaged sites in DNA, this type of repair could potentially reduce the risk of passing on mutations to successive generations.

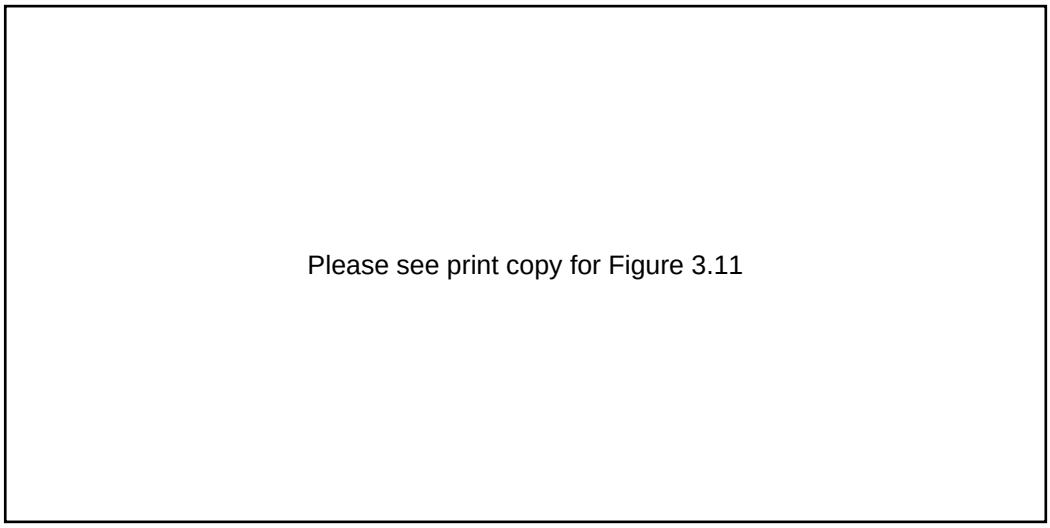


Figure 3.11 Oxidative repair of UV-damaged DNA by a rhodium metallointercalator. Irradiation using light of 400 nm wavelength resulted in repair of the thymine dimer lesion. Adapted from Erkkila *et al.*.[250]

3.6 Scope of This Chapter

Many techniques have been used to investigate binding of ruthenium metallointercalators such as those shown in Figure 3.8 to DNA. For example, circular dichroism spectroscopy provides information on the enantioselectivity of binding interactions as well as binding affinities, but does not provide significant detail about the specific atoms involved in the intermolecular interactions. On the contrary, NMR spectroscopy provides detailed information about the contacts between binding partners. However, owing to the complexity of NMR spectra, only short oligonucleotides (6 and 10 bp),[264,267,270,287] which may not have structures representative of cellular DNA, have been used in many early studies. A further limitation of NMR spectroscopy is that it may not be able to provide the same level of information about binding interactions for paramagnetic metal complexes as it can for diamagnetic complexes. One technique that can, however, be applied to all metal

complexes, but so far has received little attention for studying non-covalent interactions of metal complexes to DNA, is electrospray ionisation mass spectrometry (ESI-MS). In recent years, ESI-MS has become a routine technique for characterisation of biopolymers such as proteins and nucleic acids.[14,288,289] In addition, ESI-MS has also been widely used to study non-covalent interactions between biopolymers,[82,128] and between biopolymers and small organic molecules.[128,163-165,290,291]

In this chapter, the results of an ESI-MS investigation into the binding interactions of six ruthenium compounds (shown in Figure 3.8) with three different non-self complementary 16 base pair oligonucleotides (D1, D2 and D3) are presented. In each of these 16-mer duplexes, the first four and last four base pairs are identical, however the middle eight base pairs are different. In D1, this variable region is GC-rich, while in D2 it contains an equal mix of GC and AT base pairs, and in D3 the variable region is AT-rich and is therefore expected to favour drugs that bind as classical minor groove binders (e.g. distamycin).

D1 d(CCTCGGCCGGCCGACC/GGTCGGCCGGCCGAGG)

D2 d(CCTCATGGCCATGACC/GGTCATGGCCATGAGG)

D3 d(CCTCAAAATTTTGACC/GGTCAAAATTTTGAGG)

One aim of the study was to determine the relative binding affinities of the ruthenium compounds for these duplex DNA molecules. A second aim was to obtain information about their exact binding modes by performing competition binding experiments involving distamycin (minor groove binder) or daunomycin (intercalator).

3.7 Results and Discussion

3.7.1 Reactions of ruthenium compounds with individual 16-mer duplexes

3.7.1.1 Titration experiments

The three DNA duplexes D1, D2 and D3 were each titrated with increasing amounts of $[Ru(phen)_3]^{2+}$, $[Ru(phen)_2(pda)]^{2+}$, $[Ru(phen)_2(dpq)]^{2+}$, $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$ and $[Ru(phen)_2(dppzMe_2)]^{2+}$, and ESI mass spectra of the resulting solutions obtained, in order to determine the maximum number of drug molecules that could be bound to these dsDNA molecules. This series of ruthenium compounds allowed the effects of changing the surface area of the unique ligand (phen, pda and dpq, dpqC, dppz) and introducing methyl substituents on the outer surface of these ligands (dpq and $dpqMe_2$) on the strength of binding to be determined. ESI mass spectra were obtained of reaction mixtures containing metal:duplex DNA ratios ranging from 1:1 up to a maximum of 6:1, using the Q-ToF2 mass spectrometer (see section 2.2.1). Under the experimental conditions used, the most abundant ions observed corresponding to ruthenium/DNA complexes were 5- and 6- ions, with 7- ions also observed in some spectra. Assignments for these ions are listed in Appendix 1. The main variable that altered the appearance of the spectra was cone voltage. However the changes observed were minor and did not significantly alter the relative abundances of complexes present in the spectra. The cone voltage that gave the best quality spectra was 50 V. Lower or higher cone voltages resulted in spectra of poorer quality (low signal-to-noise ratio). Figure 3.12 illustrates the ESI mass spectra of solutions containing D2 and $[Ru(phen)_2(dppz)]^{2+}$, where the Ru:D2 ratio was 0:1, 1:1, 3:1 and 6:1.

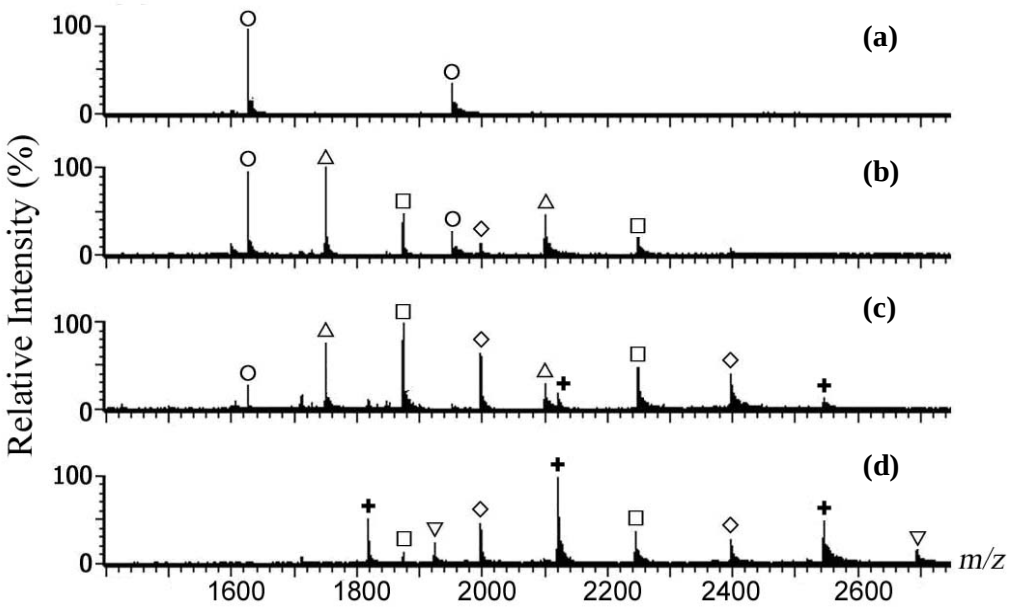


Figure 3.12 Negative ion ESI mass spectra of reaction mixtures containing different ratios of $[Ru(phen)_2(dppz)]^{2+}$ and D2. The ratios of metal:dsDNA are (a) 0:1, (b) 1:1, (c) 3:1 and (d) 6:1. ○ dsDNA; △ dsDNA + $[Ru(phen)_2(dppz)]^{2+}$; □ dsDNA + $2[Ru(phen)_2(dppz)]^{2+}$; ◇ dsDNA + $3[Ru(phen)_2(dppz)]^{2+}$; + dsDNA + $4[Ru(phen)_2(dppz)]^{2+}$; ▽ dsDNA + $5[Ru(phen)_2(dppz)]^{2+}$.

When there was no ruthenium compound in the solution only ions from free duplex D2 at *m/z* 1626.4 and 1952.0 were observed (Figure 3.12 (a)). These ions were still prominent in the spectrum of the solution with a Ru:DNA ratio of 1:1 (Figure 3.12 (b)). The latter spectrum also contained ions of medium to high abundance assigned to non-covalent complexes containing either one or two $[Ru(phen)_2(dppz)]^{2+}$ molecules bound to D2, and ions of low abundance owing to a complex containing three ruthenium molecules bound to dsDNA. As the Ru:D2 ratio was increased further, ions from free D2 decreased further in abundance and eventually disappeared when the ratio was 6:1, while the abundance of ions from complexes containing two or more $[Ru(phen)_2(dppz)]^{2+}$ molecules bound to D2 increased. Figure 3.12 (d) shows the ESI mass spectrum obtained when the Ru:D2 ratio was 6:1. The most abundant ion observed was that from a complex containing four $[Ru(phen)_2(dppz)]^{2+}$

molecules bound to D2. Ions of low to medium abundances from non-covalent complexes containing two, three or five $[Ru(phen)_2(dppz)]^{2+}$ molecules bound to D2 were also present.

Figure 3.13 shows negative ion ESI mass spectra of reaction mixtures containing a 6:1 ratio of different ruthenium compounds and D2. Examination of these spectra suggests that the six ruthenium compounds have significantly different abilities to non-covalently bind to duplex DNA. An ion of high abundance at *m/z* 1626.4 and an ion of lower abundance at *m/z* 1952.0 were observed in the spectra of reaction mixtures containing $[Ru(phen)_3]^{2+}$, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$ (Figure 3.13 (a)-(c)). These ions are attributable to free $[D2\text{-}6H]^{6-}$ and $[D2\text{-}5H]^{5-}$, respectively. These ions from free DNA were not present in the spectra of reaction mixtures containing $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$ or $[Ru(phen)_2(dppzMe_2)]^{2+}$ (Figure 3.13 (d)-(f)). This suggests that the latter three ruthenium compounds bind to duplex D2 with greater affinity than the former three. The term "greater affinity" used here indicates that under the same experimental conditions, that a greater number of molecules of one ruthenium compound can bind to DNA than for another ruthenium compound.

Since the DNA used in these experiments was 16 base pairs long, it was expected to have more than one site at which an intercalating ligand can bind. It is important to remember that the binding of the first ruthenium molecule to DNA will most likely affect subsequent binding events through steric hindrance and/or alterations to the conformation of the dsDNA.

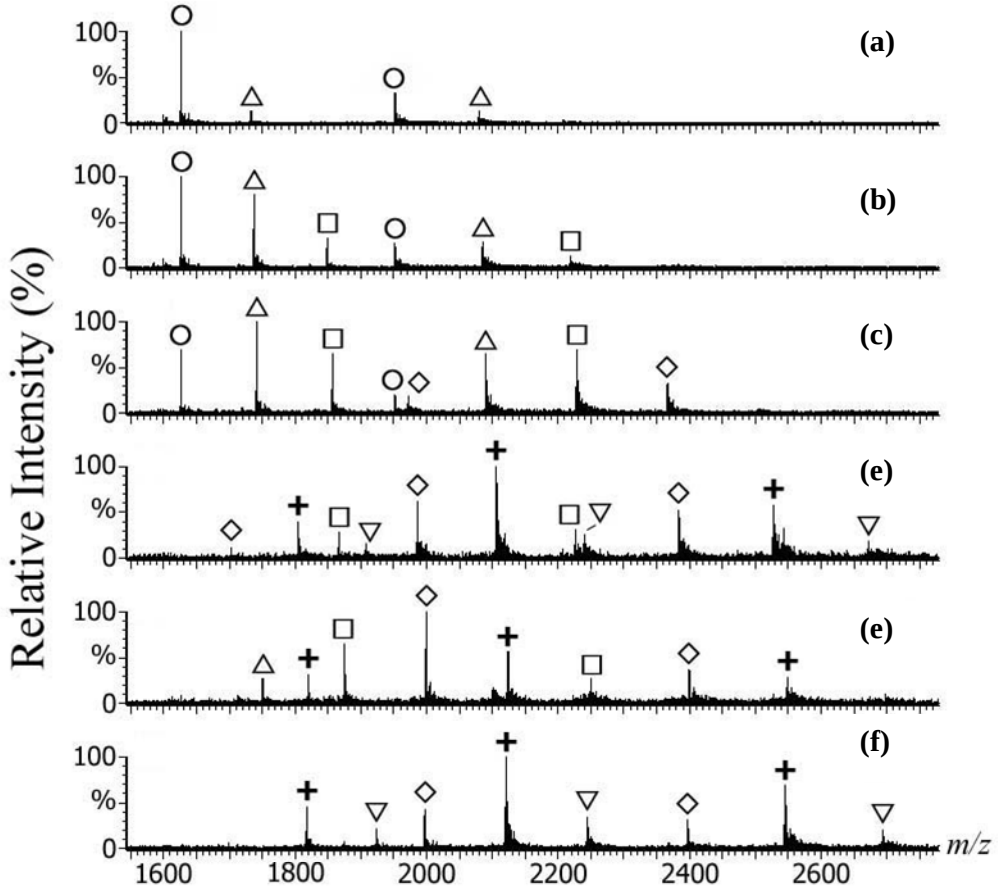


Figure 3.13 Negative ion ESI mass spectra of reaction mixtures containing a 6:1 ratio of ruthenium compound and duplex D2, obtained using a Q-Tof2™ ESI-mass spectrometer. (a) $[Ru(phen)_3]^{2+}$, (b) $[Ru(phen)_2(pda)]^{2+}$, (c) $[Ru(phen)_2(dpq)]^{2+}$, (d) $[Ru(phen)_2(dpqMe_2)]^{2+}$, (e) $[Ru(phen)_2(dpqC)]^{2+}$ and (f) $[Ru(phen)_2(dppz)]^{2+}$. ○ dsDNA; △ dsDNA + $[Ru(phen)_2(L)]^{2+}$; □ dsDNA +2$[Ru(phen)_2(L)]^{2+}$; ◇ dsDNA + 3$[Ru(phen)_2(L)]^{2+}$; + dsDNA + 4$[Ru(phen)_2(L)]^{2+}$; ▽ dsDNA + 5$[Ru(phen)_2(L)]^{2+}$. There were no ions of significant abundance at values of *m/z* < 1600 or > 2800.

By using a combination of equilibrium dialysis experiments and nonlinear least-squares analysis governed by an equation reflecting non-cooperative binding to the DNA helix, Barton *et al*. showed that one molecule of $[Ru(phen)_3]^{2+}$ binds on average to every four base pairs in calf thymus DNA.[253] In another report using fluorescence spectroscopy, the average number of calf thymus DNA base pairs involved in interactions with one molecule of $[Ru(bpy)_2(dpq)]^{2+}$ or $[Ru(bpy)_2(dpqC)]^{2+}$ were 8 and 4, respectively,[274] suggesting that a maximum of 2

and 4 ruthenium molecules can bind to a 16-mer dsDNA. The results obtained using ESI-MS (Figure 3.13 (c) and (e)) are consistent with these previous fluorescence studies.

The spectrum in Figure 3.13 (a) contained ions of low abundance from a non-covalent complex containing one $[Ru(phen)_3]^{2+}$ bound to DNA. However, when $[Ru(phen)_2(pda)]^{2+}$ or $[Ru(phen)_2(dpq)]^{2+}$ was present in the reaction mixture with D2, the abundance of ions assigned to a non-covalent complex containing one ruthenium molecule bound to D2 was considerably greater (Figure 3.13 (b) and (c)). In addition, these spectra also contained ions assigned to non-covalent complexes containing two ruthenium molecules bound to DNA. This suggests that both $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$ have a greater affinity for D2 than $[Ru(phen)_3]^{2+}$.

There have been no binding constants reported for any of the ruthenium complexes studied here with D2. However, values of 0.7×10^3 M^{-1} and 5.9×10^4 M^{-1} were obtained for $[Ru(bpy)_3]^{2+}$ and $[Ru(bpy)_2(dpq)]^{2+}$, respectively with calf thymus DNA using equilibrium dialysis and luminescence titrations methods.[274,292] A value of 5.4×10^4 M^{-1} for $[Ru(phen)_2(dpq)]^{2+}$ binding to calf thymus DNA, measured by fluorescence spectroscopy, has also been reported, and is similar to the value for $[Ru(bpy)_2(dpq)]^{2+}$.[293] Taken together these binding constants are consistent with the results of the current ESI-MS study, which suggests that $[Ru(phen)_2(dpq)]^{2+}$ has significantly greater affinity towards D2 than $[Ru(phen)_3]^{2+}$. This may be attributed to the presence of the dpq ligand in the former compound, which confers on it a significantly greater ability to intercalate into DNA. Likewise the greater DNA

binding affinity of $[Ru(phen)_2(pda)]^{2+}$ compared to $[Ru(phen)_3]^{2+}$ observed in the study reported here may be attributed to the presence of the pda ligand. The difference between the phen and pda ligands is that the latter coordinates to metal ions via exocyclic amine groups rather than endocyclic amines (compare Figures 3.8 (a) and (e)), resulting in the intercalating ligand being located further from the central ruthenium ion in the case of the latter ligand. This may facilitate greater hydrophobic interactions between the pda ligand and non-polar sites in the DNA base stack, resulting in a stronger overall binding interaction.

Figure 3.13 (b) and (c) allow a comparison to be made of the strength of the binding interaction between $[Ru(phen)_2(pda)]^{2+}$ and D2 on the one hand, and $[Ru(phen)_2(dpq)]^{2+}$ and D2 on the other. Figure 3.13 (c) contains ions from free D2 of slightly lower abundance, and ions of greater abundance from ruthenium/DNA complexes, compared to Figure 3.13 (b). This suggests that $[Ru(phen)_2(dpq)]^{2+}$ has a slightly greater affinity for D2 than $[Ru(phen)_2(pda)]^{2+}$. This proposal is also supported by the presence of ions of low abundance assigned to non-covalent complexes containing three $[Ru(phen)_2(dpq)]^{2+}$ bound to D2 at *m/z* 2367.3 in Figure 3.13 (c), as there are no corresponding ions in the spectrum of the reaction mixture containing $Ru(phen)_2(pda)]^{2+}$ (Figure 3.13 (b)).

The importance of hydrophobic interactions between the ruthenium compounds and DNA may be seen by comparing the spectra of reaction mixtures containing $[Ru(phen)_2(dpq)]^{2+}$ and D2, and $[Ru(phen)_2(dpqMe_2)]^{2+}$ and D2. Figure 3.13 (d) shows that the most abundant ions present in the spectrum of the reaction mixture containing the latter ruthenium compound are from a non-covalent complex

containing four ruthenium molecules bound to D2. Ions with medium abundance assigned to complexes containing three and five ruthenium molecules bound to D2 were also observed, whereas ions assigned to complexes containing two ruthenium molecules bound to D2 were of low abundance. No ions assignable to complexes containing one ruthenium molecule bound to D2 or free D2 were present in this spectrum. In contrast, in the spectrum of the reaction mixture containing $[Ru(phen)_2(dpq)]^{2+}$ and D2 (Figure 3.13 (c)), the most abundant ions are from complexes containing only one ruthenium molecule bound to D2. Ions from complexes containing two and three ruthenium molecules bound to D2 were also present at medium and low abundance, respectively. However, there were no ions assignable to complexes containing higher numbers of ruthenium molecules bound to D2. This suggests that $[Ru(phen)_2(dpqMe_2)]^{2+}$ has a significantly greater binding affinity for D2 than $[Ru(phen)_2(dpq)]^{2+}$. The only difference between the two ruthenium compounds is the presence of two methyl groups on the leading edge of the $dpqMe_2$ ligand, which would not be expected to increase the ability of $[Ru(phen)_2(dpqMe_2)]^{2+}$ to intercalate more deeply into DNA than $[Ru(phen)_2(dpq)]^{2+}$. These methyl groups, however, may participate in additional hydrophobic interactions with non-polar sites in the DNA base stack, resulting in greater binding affinity towards D2.

The most abundant ion present in the spectrum of the mixture containing D2 and $[Ru(phen)_2(dpqC)]^{2+}$ (Figure 3.13 (e)) is at *m/z* 1999.4, which is assigned to a complex containing three $[Ru(phen)_2(dpqC)]^{2+}$ molecules bound to D2. Ions from complexes containing four ruthenium molecules bound to D2 were also present but at a lower abundance, and there were no ions from complexes containing five

ruthenium molecules bound to D2. In contrast, Figure 3.13 (f) shows ions assigned to complexes containing four and five $[Ru(phen)_2(dppz)]^{2+}$ molecules bound to D2 at high and medium abundance, respectively. This suggests that both $[Ru(phen)_2(dpqC)]^{2+}$ and $[Ru(phen)_2(dppz)]^{2+}$ also have a greater binding affinity towards D2 than $[Ru(phen)_2(dpq)]^{2+}$, most likely because of the additional ring systems present in both of the former compounds, which enables additional intercalation interactions with the DNA base stack. It also appears that $[Ru(phen)_2(dppz)]^{2+}$, which contains the completely aromatic dppz ligand, shows a greater ability to bind to D2 than $[Ru(phen)_2(dpqC)]^{2+}$, containing dpqC. The latter ligand differs from dppz in having one saturated ring system, which may reduce the strength of intercalative interactions owing to slight deviations from co-planarity with the rest of the ligand.

The above differences in binding affinities displayed by the different ruthenium complexes towards DNA can be summarised by plotting the relative abundances of ions assigned to different non-covalent complexes as a function of the number of ruthenium molecules bound to DNA.[294] Relative abundances were obtained by summing the total ion intensities for all 5-, 6- and 7- ions assigned to individual ruthenium-dsDNA complexes, and then dividing by the total ion intensity of all ions in each spectrum, and expressing the result as a percentage. Figure 3.14 shows the relative abundances of non-covalent complexes formed in reaction mixtures containing a 6:1 ratio of ruthenium compound and D2. Comparison of the graphs readily reveals that $[Ru(phen)_2(dpqMe_2)]^{2+}$ and $[Ru(phen)_2(dppz)]^{2+}$ have significantly greater binding affinities than the other four ruthenium compounds. This is shown by the fact that the ions with greatest abundance in the spectra of

reaction mixtures containing the former two compounds were those containing four ruthenium molecules bound to DNA. In addition these spectra contained ions assigned to non-covalent complexes containing three and five ruthenium molecules with similar relative abundances.

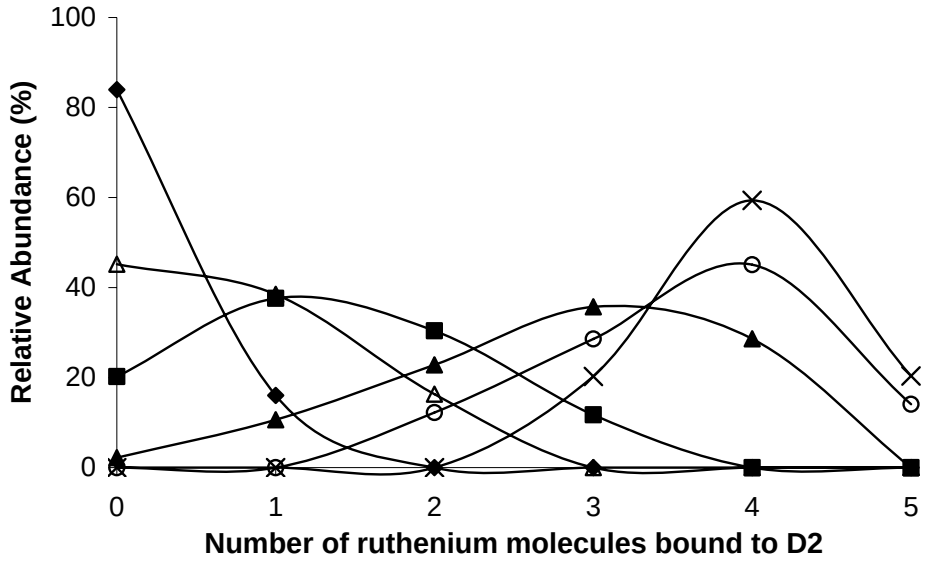


Figure 3.14 Relative abundances of non-covalent complexes obtained from reaction mixtures containing a 6:1 ratio of ruthenium compound and duplex D2. ◆ $[Ru(phen)_3)]^{2+}$; ■ $[Ru(phen)_2(dpq)]^{2+}$; ▲ $[Ru(phen)_2(dpqC)]^{2+}$; × $[Ru(phen)_2(dppz)]^{2+}$; ○ $[Ru(phen)_2(dpqMe_2)]^{2+}$; △ $[Ru(phen)_2(pda)]^{2+}$.

Figure 3.14 also shows that the ruthenium compound with the next highest binding affinity towards D2 is $[Ru(phen)_2(dpqC)]^{2+}$. This is borne out by the much higher relative abundance of ions containing three and four ruthenium molecules bound to D2, compared to that of the corresponding ions containing the other three ruthenium compounds. Further comparison of the results presented graphically in Figure 3.14 shows the overall order of binding affinity to be $[Ru(phen)_2(dppz)]^{2+} \geq [Ru(phen)_2(dpqMe_2)]^{2+} > [Ru(phen)_2(dpqC)]^{2+} > [Ru(phen)_2(dpq)]^{2+} > [Ru(phen)_2(pda)]^{2+} > [Ru(phen)_3]^{2+}$. Similar trends in relative binding affinity were

revealed by graphical analysis of spectra of reaction mixtures containing the same ruthenium compounds and either D1 or D3 (see Appendix 3).

3.7.1.2 Competition experiments between ruthenium compounds

In order to test the order of relative binding affinities of ruthenium compounds for DNA, obtained by comparison of the reaction mixtures described above, a series of competition experiments using solutions containing pairs of ruthenium compounds and a single DNA duplex (either D1, D2 or D3) were performed. In each reaction mixture the ratio of complex1:complex2:DNA was 3:3:1. Overall the results obtained for the different duplex DNA sequences were very similar. That is, the relative affinity of each of the ruthenium complexes was independent of the DNA sequence used. In addition, the results obtained from these competition experiments supported the order of relative binding affinities described in section 3.7.1.1.

ESI mass spectra of competition mixtures containing $[Ru(phen)_3]^{2+}$ and a second ruthenium compound were dominated by ions assigned to non-covalent complexes containing the latter compounds (data not shown). This confirms that $[Ru(phen)_3]^{2+}$ has the lowest affinity towards each duplex DNA examined. ESI mass spectra of most other mixtures of ruthenium compounds were typically more complex, owing to the presence of ions from complexes containing one or more of either or both ruthenium molecules.

$[Ru(phen)_2(dpq)]^{2+}$ and $[Ru(phen)_2(pda)]^{2+}$

Preliminary studies (section 3.7.1.1) suggested that $[Ru(phen)_2(dpq)]^{2+}$ and $[Ru(phen)_2(pda)]^{2+}$ both had DNA binding affinities greater than that of

$[Ru(phen)_3]^{2+}$, but less than that of the other three ruthenium compounds. The latter was confirmed by competition experiments in which $[Ru(phen)_2(dpq)]^{2+}$ or $[Ru(phen)_2(pda)]^{2+}$ was allowed to compete with $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$ or $[Ru(phen)_2(dpqMe_2)]^{2+}$. For example, Figure 3.15 (a) shows the ESI mass spectrum of a reaction mixture containing D1, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpqC)]^{2+}$. Ions at *m/z* 1751.4, 1875.8, 2000.2, 2101.7 and 2250.9 are assigned to complexes containing one or more $[Ru(phen)_2(dpqC)]^{2+}$ bound to D1 (▲). These ions are of greater abundance than those at *m/z* 1738.3, 1849.8, 2086.1 and 2219.5, which are assigned to complexes containing one or more $[Ru(phen)_2(pda)]^{2+}$ bound to D1 (■). This supports the earlier conclusion that $[Ru(phen)_2(dpqC)]^{2+}$ has greater affinity towards D1 than $[Ru(phen)_2(pda)]^{2+}$.

The spectrum in Figure 3.15 (a) also shows other ions, including those at *m/z* 1862.5, 1974.1, 1986.8, 2235.3 and 2384.5 which are assigned to complexes containing one or more of both ruthenium molecules bound to D1 (❖). Figure 3.15 (b) shows the ESI mass spectrum of a solution containing D1, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$. Ions assigned to non-covalent complexes containing one or two $[Ru(phen)_2(dpq)]^{2+}$ bound to D1 are higher in abundance than those for complexes of D1 with $[Ru(phen)_2(pda)]^{2+}$. This experiment therefore provides further evidence that $[Ru(phen)_2(dpq)]^{2+}$ has a greater DNA binding affinity than $[Ru(phen)_2(pda)]^{2+}$. Furthermore when the duplex present in the competition reaction mixture was either D2 or D3, ions assigned to complexes containing $[Ru(phen)_2(dpq)]^{2+}$ bound to dsDNA were more prominent, providing more support for this conclusion, and suggesting that there were no substantial variations in DNA sequence specificity between these two ruthenium molecules.

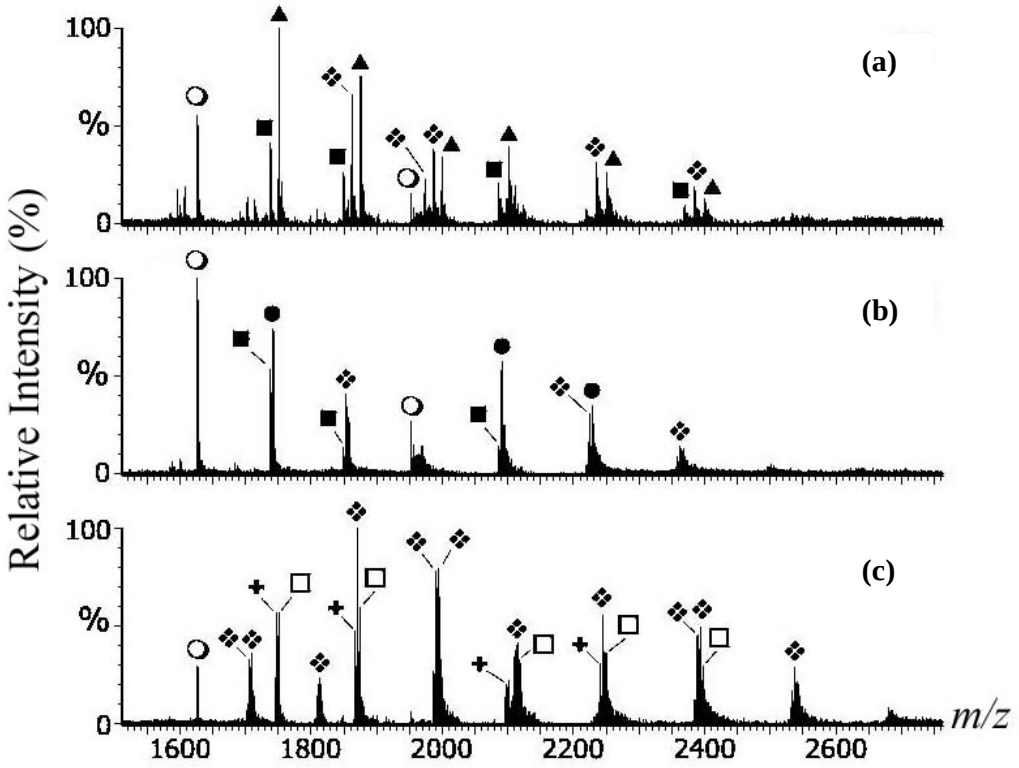


Figure 3.15 Negative ion ESI mass spectra of reaction mixtures containing a 3:3:1 ratio of two ruthenium compounds and D1. (a) solution containing $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpqC)]^{2+}$; (b) solution containing $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$ and (c) solution containing $[Ru(phen)_2(dpqMe_2)]^{2+}$ and $[Ru(phen)_2(dppz)]^{2+}$. ○ dsDNA; ▲ dsDNA + $x[Ru(phen)_2(dpqC)]^{2+}$, x = 1-3; ■ dsDNA + $x[Ru(phen)_2(pda)]^{2+}$, x = 1-2; ● dsDNA + $x[Ru(phen)_2(dpq)]^{2+}$, x = 1-2; □ dsDNA + $x[Ru(phen)_2(dppz)]^{2+}$, x = 1-4; + dsDNA + $x[Ru(phen)_2(dpqMe_2)]^{2+}$, x = 1-2; ❖ dsDNA + one or more of both ruthenium molecules.

$[Ru(phen)_2(dppz)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$

Competition experiments conducted using reaction mixtures containing $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dpqMe_2)]^{2+}$ and $[Ru(phen)_2(dppz)]^{2+}$ suggested that these ruthenium complexes have very similar DNA affinities greater than that of the remaining ruthenium compounds. For example, Figure 3.15 (c) shows the ESI mass spectrum of a solution containing D1, $[Ru(phen)_2(dppz)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$. The abundances of ions assigned to non-covalent complexes containing one or two ruthenium molecules bound to D1 are very similar for both

ruthenium compounds. The ions with the highest abundances are those at *m/z* 1870.9, 1990.3 and 1994.6. The first ion (*m/z* 1870.9) is assigned to $[D1+Ru(phen)_2(dppz)+Ru(phen)_2(dpqMe_2)-10H]^{6-}$, which corresponds to dsDNA bound to one molecule of both ruthenium compounds. The second ion (*m/z* 1990.3) is assigned $[D1+Ru(phen)_2(dppz)+2Ru(phen)_2(dpqMe_2)-10H]^{6-}$, which contains one molecule of $[Ru(phen)_2(dppz)]^{2+}$ and two molecules of $[Ru(phen)_2(dpqMe_2)]^{2+}$ bound to D1. The last ion (*m/z* 1994.6) is assigned to $[D1+2Ru(phen)_2(dppz)+Ru(phen)_2(dpqMe_2)-12H]^{6-}$, which contains two molecules of $[Ru(phen)_2(dppz)]^{2+}$ and only one $[Ru(phen)_2(dpqMe_2)]^{2+}$ bound to D1. The similarity in abundance of the last two ions further supports the conclusion that $[Ru(phen)_2(dppz)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$ have very similar DNA affinities for D1.

The results obtained from these competition experiments suggests that the relative DNA affinities of the ruthenium compounds follows the order: $[Ru(phen)_2(dppz)]^{2+}$ ~ $[Ru(phen)_2(dpqMe_2)]^{2+}$ ~ $[Ru(phen)_2(dpqC)]^{2+}$ > $[Ru(phen)_2(dpq)]^{2+}$ > $[Ru(phen)_2(pda)]^{2+}$ > $[Ru(phen)_3]^{2+}$. To date, only binding constants for $[Ru(phen)_2(dppz)]^{2+}$ (K_b = 3.6 x 10^6 M^{-1}, obtained from fluorescence titrations with $d(CGCGATCGCG)_2$) and $[Ru(phen)_2(dpq)]^{2+}$ (K_b = 5.4 x 10^4 M^{-1}, obtained from fluorescence titrations with calf thymus DNA) have been reported for members of this series of ruthenium compounds.[293,295] However, binding constants for complexes of the series $[Ru(bpy)_2(L)]^{2+}$, where L = bpy, dpq, dpqC and dppz, are known. These binding constants were obtained from luminescence titrations of the ruthenium compounds with calf thymus DNA.[274,292] The binding constant for $[Ru(bpy)_2(dppz)]^{2+}$ (K_b = 8.8 x 10^6 M^{-1}) is much greater than that of either

$[Ru(bpy)_2(dpqC)]^{2+}$ (K_b = 8.5 x 10^4 M^{-1}) or $[Ru(bpy)_2(dpq)]^{2+}$ (K_b = 5.9 x 10^4 M^{-1}), which are in turn significantly greater than that for $[Ru(bpy)_3]^{2+}$ (K_b = 0.07 x 10^4 M^{-1}). The trend in relative DNA binding affinities for the series $[Ru(bpy)_2(L)]^{2+}$ is therefore in broad agreement with the trend in relative binding affinities obtained from ESI mass spectra of competition mixtures involving the series $[Ru(phen)_2(L)]^{2+}$ presented here. This provides support for the use of ESI-MS as a tool for rapidly assessing relative DNA binding affinities of metallointercalators.

3.7.1.3 DNA selectivity

The most important mode of binding to DNA for the ruthenium compounds investigated here is probably intercalation. This conclusion is based on the results of earlier studies involving these and similar compounds with different DNA sequences.[161,239,247,265,296,297] In many of these studies, the intercalators were shown to bind preferentially to duplex DNA in GC rich regions.[160,229,247,296,298-300] For example, Figure 3.16 shows a high-resolution X-ray crystal structure (1.2 Å) of the rhodium complex, Δ-α-$[Rh[(R,R)\text{-}Me_2trien]phi]^{3+}$, intercalating into the oligonucleotide 5′-G(dIU)TGCAAC-3′ (dIU, 5-iodo-deoxy-uridine), specifically in the GC region in the major groove.[300]

Among the duplex DNA sequences examined here, it was expected that the ruthenium compounds would show the highest affinity towards D1 (a GC-rich sequence) and the lowest affinity towards D3 (an AT-rich sequence). It was therefore surprising that all of the ruthenium compounds examined in this study displayed greater affinity towards D2 compared to the other two duplexes.

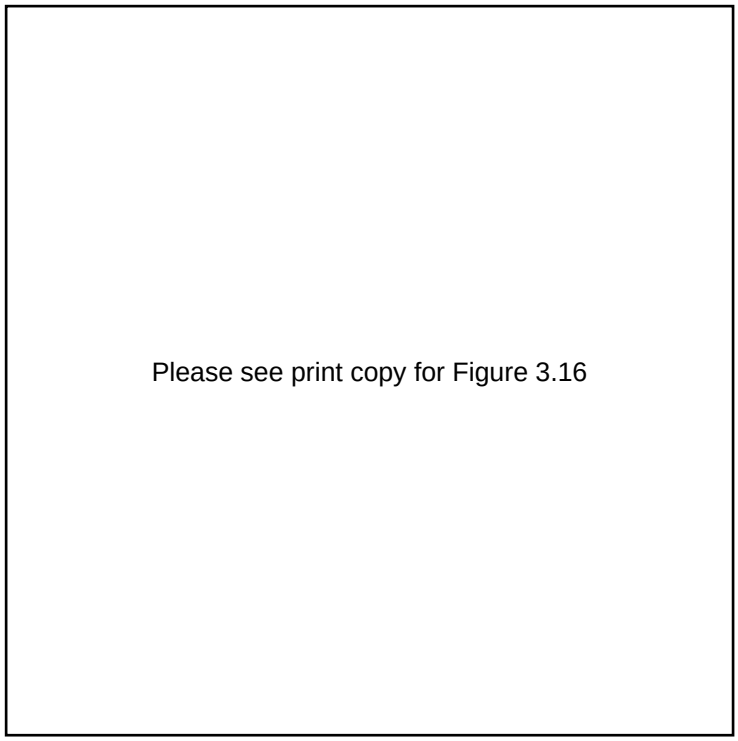


Figure 3.16 Crystal structure of Δ-α-[Rh[(R,R)-Me_2trien]phi]$^{3+}$ bound to the major groove of 5′-G(dIU)TGCAAC-3′ (dIU, 5-iodo-deoxy-uridine) in the GC region.[300] One of the regions where the rhodium complex is specifically bound to DNA is shown in blue spacefill (G4, C13, G12 and C5). The rhodium atoms are in red spacefill and ligands are in green (ball and stick). From coordinates in Kielkopf *et al*..[300]

For example, Figure 3.17 illustrates the DNA selectivity of [Ru(phen)$_2$(dpqMe$_2$)]$^{2+}$, using relative abundances obtained from reaction mixtures containing a 6:1 ratio of [Ru(phen)$_2$(dpqMe$_2$)]$^{2+}$ and duplex D1, D2 or D3. The relative abundances of ions from non-covalent complexes containing four and five ruthenium molecules bound to dsDNA are significantly greater for D2 than for D1 or D3. Similar results were obtained when relative abundances were plotted in the same fashion for reaction mixtures containing [Ru(phen)$_2$(dpq)]$^{2+}$, [Ru(phen)$_2$(dpqC)]$^{2+}$, [Ru(phen)$_2$(dppz)]$^{2+}$ and [Ru(phen)$_2$(pda)]$^{2+}$ (data not shown). The differences in relative abundances between ions containing the same number of ruthenium molecules bound to different

DNA duplexes were, however, much smaller for the latter compound. This reflects the lower DNA affinity of this ruthenium compound compared to all of the others studied here, with the exception of $[Ru(phen)_3]^{2+}$.

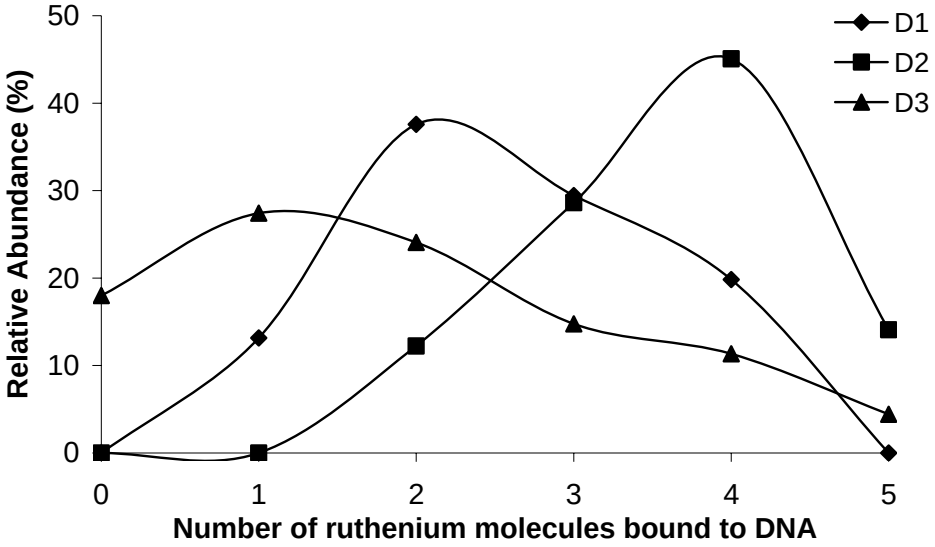


Figure 3.17 DNA sequence selectivity of $[Ru(phen)_2(dpqMe_2)]^{2+}$. Relative abundances of complexes in reaction mixtures containing a 6:1 ratio of $[Ru(phen)_2(dpqMe_2)]^{2+}$ and duplex D1, D2 or D3. ○ Ruthenium molecules bound to D1; ■ ruthenium molecules bound to D2; ▲ ruthenium molecules bound to D3.

For $[Ru(phen)_3)]^{2+}$, there was very little difference between the relative abundances of ions containing the same number of ruthenium molecules bound to different duplexes (Figure 3.18). This is consistent with $[Ru(phen)_3]^{2+}$ having the lowest DNA selectivity (and affinity) amongst the ruthenium compounds examined. The observation that most of the ruthenium compounds preferred to bind to D2, compared to D1 and D3, was surprising since D2 does not contain the highest GC content amongst the duplexes being studied. This suggests that either intercalation does not totally dominate interactions between ruthenium compounds and DNA, or that the increase in GC content (from D2 to D1) does not necessarily result in more, or stronger binding interactions. It should also be remembered that the proposal that

greater GC content favours intercalation for the ruthenium compounds and for organic intercalators,[296,298] is a generalisation that does not take account of the subtle effects of: (i) interactions involving substituents (e.g. ancillary ligands on metal ions) that may enhance or decrease binding; (ii) changes in the local conformation of DNA after initial binding by an intercalator at the highest affinity site, or (iii) the inherent lower stability of AT-rich sequences that might skew measurements of the relative abundances of complexes. A full understanding of the reasons why most ruthenium complexes prefer to bind to D2 will require a detailed structural investigation of these interactions using NMR spectroscopy or X-ray crystallography. These experiments will be challenging, as the current ESI-MS studies have shown that some reaction mixtures even with a 1:1 metal:DNA ratio contain non-covalent complexes with a range of stoichiometries.

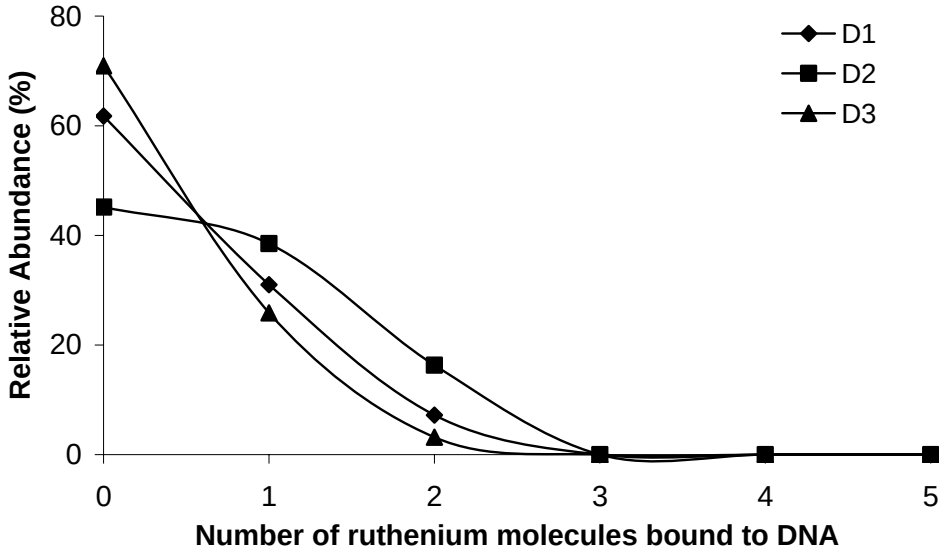


Figure 3.18 DNA sequence selectivity of $[Ru(phen)_3]^{2+}$. Relative abundances of complexes in reaction mixtures containing a 6:1 ratio of $[Ru(phen)_3)]^{2+}$ and duplex D1, D2 or D3. ○ Ruthenium molecules bound to D1; ■ ruthenium molecules bound to D2; ▲ ruthenium molecules bound to D3.

In analysing equilibrium mixture, the question must be asked as to whether the method of observation used affects the relative abundances of species in the mixture. In the case of ESI-MS there are two effects to consider. The first is that there might be non-specific associations of molecules in the ionisation source (e.g. positively charged ruthenium molecules binding to negatively charged DNA). Ding and Anderegg[152] were the first to address this question by comparing ESI mass spectra of self-complementary and nonself-complementary DNA strands. They found that there were negligible amounts of ions from non-specific complexes in ESI mass spectra when the concentration of DNA in solution was ≤ 100 μM. In the current work the concentration of DNA in solution was 10 μM. Furthermore, the instrument conditions used in the current work did not favour the maintenance of electrostatic interactions. Increasing the cone voltage accelerates non-specific complexes through the source, making them more likely to dissociate when they collide with nitrogen molecules (collision-induced dissociation, CID) when their internal energy increases. The cone voltage used in this work was chosen to minimise non-specific interactions such as those involving Na^+ or K^+ and DNA. Although there may be some contribution from electrostatic forces to the strength of binding of the ruthenium compounds to DNA, it is likely that these complexes were stabilised predominantly by intercalation.

The second effect to be considered is the possibility that the non-covalent complexes might dissociate in the gas phase. However, there have been several studies into the gas phase stability of drug-DNA complexes performed using ESI-MS in this and other laboratories,[128,155,164,165,294,301,302] which showed that relative binding affinities measured by ESI-MS agree with those obtained by other methods.[36,48,71,303] If this is

also true for the $[Ru(phen)_2L]^{2+}$ complexes examined in the present study, it would be expected as the ratio of drug:DNA in solution is increased, the number of drug molecules binding to DNA would eventually reach a maximum value. The experiments described in the following section were performed to determine if this is what is observed.

3.7.1.4 Saturation experiments

In previous experiments, reaction mixtures contained up to a maximum of a 6:1 ratio of ruthenium compound to DNA. ESI mass spectra of these solutions showed that the extent of DNA binding depended on the identity of the ruthenium compound used and also the ruthenium:DNA ratio. Complexes containing up to five ruthenium molecules bound to dsDNA were observed in some instances. The neighbour exclusion principle[304] states that simple molecules such as the classical intercalator ethidium bromide can only bind to every second base pair in DNA. According to this principle only eight molecules can bind to a 16-mer duplex. This is supported by a recent ESI-MS study which showed that a maximum number of eight molecules of ethidium bromide could bind to a 16-mer dsDNA molecule.[305] However, since the ruthenium compounds examined here are greater in size compared to ethidium ion, fewer ruthenium molecules would be expected to bind to dsDNA of this length if non-specific gas phase associations are not important, even at high ruthenium:DNA ratios. In order to address this issue, which bears directly on the question whether complexes observed in ESI mass spectra accurately reflect solution composition, mass spectra were obtained of reaction mixtures containing from a 1:1 up to a 25:1 ratio of $[Ru(phen)_2(dpqC)]Cl_2$ and D2. The higher solubility of the chloride salt of this compound compared to the corresponding hexafluorophosphate salt used

initially, allowed the preparation of reaction mixtures containing much higher ratios of ruthenium:DNA. Figure 3.19 shows the relative abundances of complexes containing different numbers of $[Ru(phen)_2(dpqC)]^{2+}$ molecules bound to D2 in the reaction mixtures. Ions assigned to free D2 decreased quickly as the ruthenium:DNA ratio was increased, reflecting the high affinity of $[Ru(phen)_2(dpqC)]^{2+}$ molecules for D2. The relative abundances of non-covalent complexes containing one and two $[Ru(phen)_2(dpqC)]^{2+}$ molecules bound to D2 increased to a maximum value at the relatively low ruthenium:D2 ratio of 5:1, and then decreased as the ruthenium:D2 ratio was increased further. Figure 3.19 also shows that the relative abundances of non-covalent complexes containing three $[Ru(phen)_2(dpqC)]^{2+}$ molecules bound to D2 reached a maximum value at a ruthenium:DNA ratio of between 10:1 and 15:1, and then also decreased as the ratio was increased further.

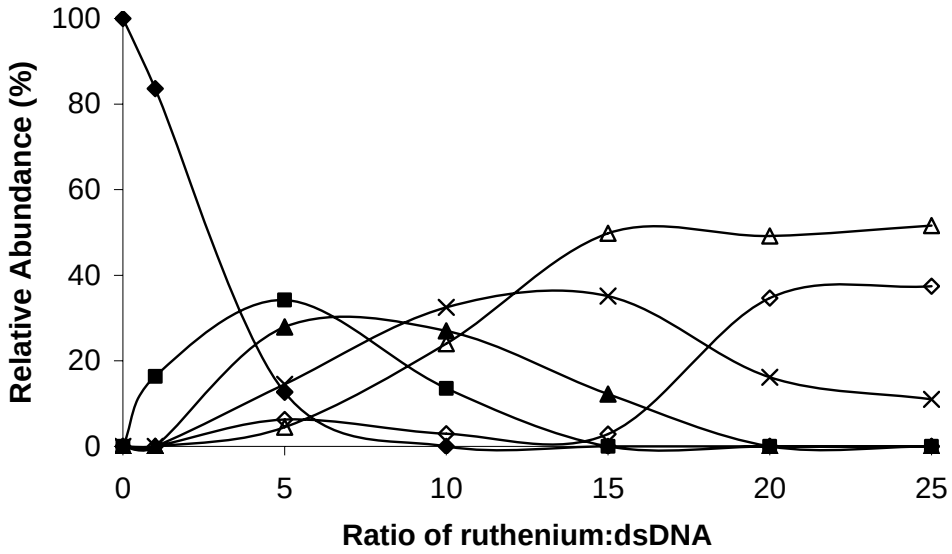


Figure 3.19 Relative abundances of ions assigned to non-covalent complexes present in ESI mass spectra of reaction mixtures containing $[Ru(phen)_2(dpqC)]Cl_2$ and D2. ◆ 0 Ruthenium molecules bound; ■ 1 ruthenium molecule bound; ▲ 2 ruthenium molecules bound; × 3 ruthenium molecules bound; △ 4 ruthenium molecules bound; ◇ 5 ruthenium molecules bound.

Complexes containing four and five $[Ru(phen)_2(dpqC)]^{2+}$ molecules bound to D2 grew as the ratio of ruthenium:D2 was increased, and were the dominant species in the reaction mixtures with the highest ratios. The changes in relative abundances of non-covalent complexes illustrated in Figure 3.19 reflect the variations in concentrations expected for metal complexes involved in stepwise complex formation equilibria.[306] This supports the view that ESI mass spectra faithfully reflect solution equilibria. In addition, the absence of ions assigned to complexes containing more than five $[Ru(phen)_2(dpqC)]^{2+}$ molecules bound to D2, even at very high metal:DNA ratios, suggests that this is the maximum number of ruthenium molecules that can bind to this 16-mer duplex. This observation is in reasonable agreement with the results of binding studies involving $[Ru(bpy)_2(dpqC)]^{2+}$ and calf thymus DNA, which showed that one ruthenium molecule binds on average to every four DNA base pairs.[253,274]

3.7.1.5 DNA melting experiments

The effect that binding by different ruthenium compounds has on the stability of DNA was investigated by monitoring its melting temperature. Previous studies have shown that intercalation of small molecules into DNA increases the thermal stability of the duplex, resulting in an increase in DNA melting temperature.[307-309] It is possible to follow the DNA melting process by observing the change in absorbance at 260 nm, where ultraviolet light is absorbed strongly by the purine and pyrimidine bases. The resulting plot of absorbance at 260 nm (A_{260}) versus temperature is called a DNA melting curve. DNA melting curves are generally sigmoidal in shape, with the point of inflection corresponding to the situation where 50% of all dsDNA molecules in solution have undergone strand separation to form ssDNA. This

temperature is referred to as the DNA melting temperature (T_m), and can be used to rapidly compare the stability of different DNA molecules or the same DNA molecule in different environments.

Figure 3.20 shows DNA melting curves for D2 **(■)** and for a reaction mixture containing a 3:1 ratio of $[Ru(phen)_2(dpq)]^{2+}$ and D2 **(◆)**. The melting temperatures (T_m) obtained from these plots were 63.9 and 66.3 °C, respectively. The increase in DNA melting temperature for the solution containing the ruthenium complex suggests that it has intercalated into the double helical DNA to a significant extent, stabilising it and resulting in more energy being required to separate the two strands. Similar experiments were performed using solutions containing a 3:1 ratio of the other ruthenium compounds and D2, and the results are summarised in Table 3.1.

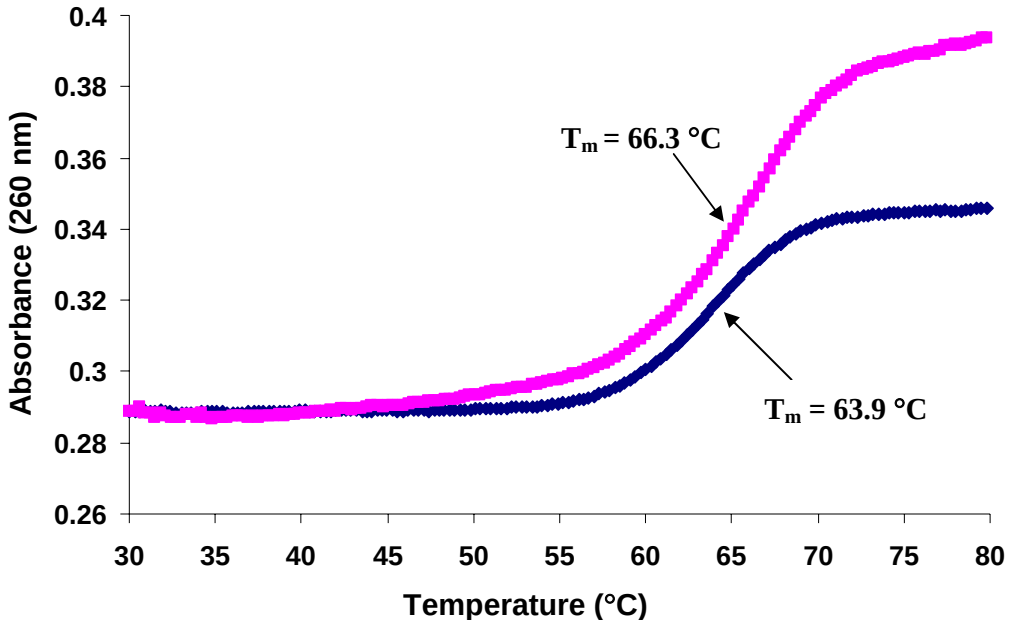


Figure 3.20 DNA melting curves for D2 (◆) and for a reaction mixture containing a 3:1 ratio of $[Ru(phen)_2(dpq)]^{2+}$ and D2 (■) in 100 mM NH_4OAc, pH 8.5, the same conditions used for ESI-MS experiments. Melting points (T_m), were calculated from the inflection points of the curves using Cary WINUV software.

Table 3.1 DNA melting temperatures obtained from reaction mixtures containing D2 and different ruthenium compounds.

Reaction mixtures	Melting Temperature, T_m, (°C)	ΔT_m (°C)*
D2 alone	63.9	-
D2 + $[Ru(phen)_3]^{2+}$	64.9	1.0
D2 + $[Ru(phen)_2(pda)]^{2+}$	65.5	1.6
D2 + $[Ru(phen)_2(dpq)]^{2+}$	66.3	2.4
D2 + $[Ru(phen)_2(dpqC)]^{2+}$	69.2	5.3
D2 + $[Ru(phen)_2(dpqMe_2)]^{2+}$	69.7	5.8
D2 + $[Ru(phen)_2(dppz)]^{2+}$	73.2	9.3

* ΔT_m is the difference between T_m for D2 and that for D2 with each ruthenium compound.

The higher melting temperatures observed in each case for solutions containing DNA and ruthenium compounds suggest that the duplex is stabilised when the latter compounds bind. Furthermore, the extent of the change in melting temperature (ΔT_m) can be used as a measure of the relative binding affinity of the different ruthenium compounds. The highest melting temperature observed was for a solution containing $[Ru(phen)_2(dppz)]^{2+}$ (T_m = 73.2 °C). This was followed closely by the T_m for $[Ru(phen)_2(dpqMe_2)]^{2+}$ (69.7 °C) and $[Ru(phen)_2(dpqC)]^{2+}$ (69.2 °C). Competition experiments performed using ESI-MS and solutions containing $[Ru(phen)_2(dppz)]^{2+}$ and $[Ru(phen)_2(dpqMe_2)]^{2+}$ (section 3.7.1.2) could not unambiguously decide which compound has the highest affinity for D2. However, the results of the DNA melting temperature experiments clearly show that binding of $[Ru(phen)_2(dppz)]^{2+}$ caused a greater increase in the stability of DNA than $[Ru(phen)_2(dpqMe_2)]^{2+}$, which suggests that the former compound has the greater DNA binding affinity of the two. Table 3.1

also shows that $[Ru(phen)_2(dpqMe_2]^{2+}$ only increased T_m by a slightly greater amount than $[Ru(phen)_2(dpqC]^{2+}$. This is consistent with the former compound having a slightly greater DNA affinity, as was also revealed by the ESI-MS results shown in Figure 3.13. The three remaining ruthenium compounds, $[Ru(phen)_3]^{2+}$, $[Ru(phen)_2(pda)]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$, all had a much smaller effect on T_m than the first three. Overall the results shown in Table 3.1 suggests the following order of relative binding affinities: $[Ru(phen)_2(dppz)]^{2+}$ > $[Ru(phen)_2(dpqMe_2)]^{2+}$ > $[Ru(phen)_2(dpqC)]^{2+}$ > $[Ru(phen)_2(dpq)]^{2+}$ > $[Ru(phen)_2(pda)]^{2+}$ > $[Ru(phen)_3]^{2+}$. This order is in good agreement with that obtained earlier by ESI-MS.

3.7.2 Competition experiments involving ruthenium compounds and organic drugs

The above experiments demonstrate that ESI-MS is a useful tool that can provide information about the number, relative amounts and stoichiometry of complexes present in reaction mixtures containing ruthenium compounds and DNA. Furthermore it also provides information about DNA selectivity and the relative binding affinities of individual ruthenium compounds. In order to obtain information on the DNA binding modes of these ruthenium compounds and their preferred binding sites, a series of competition experiments were performed in which each ruthenium compound competed for binding sites on duplex DNA with the well-characterised organic DNA-binding compounds daunomycin (an intercalator) and distamycin (a minor groove binder). Distamycin A (Figure 3.4 (a)) shows a preference for binding to AT-rich regions in the minor groove of DNA,[199,235,310] while daunomycin (Figure 3.6 (b)) binds preferentially to GC-rich DNA via intercalation.[246,297,311] Therefore it was expected that distamycin would bind

preferentially to D3 in its minor groove in an analogous fashion to that shown in Figure 3.5 (a) for the binding of distamycin to the sequence d(GGCCAATTGG)$_2$.[227] Figure 3.5 (b) shows intercalation of two disaccharide anthracyclines, which are structurally similar to daunomycin with an extra sugar unit attached, between GC base pairs in the hexamer d(CGATCG)$_2$.[236]

Two types of experiments were performed, the first involved incubation of either excess daunomycin or distamycin with DNA prior to reaction with the ruthenium compounds. The purpose of this set of experiments was to determine whether the binding of the organic intercalator or minor groove binder would prevent the ruthenium compounds from binding to DNA. Since the binding constants for both organic compounds with DNA are at least equal to or much greater than that for the ruthenium compounds,[245,312] it would be expected that the ruthenium compounds could only bind to DNA at base sequences different from those used by the organic drug. For the second type of experiment, individual ruthenium compounds were first incubated with DNA prior to reaction with either daunomycin or distamycin. If this resulted in the displacement of ruthenium molecules from DNA then this would provide evidence that they bind to the same regions of DNA as the organic drugs.

3.7.2.1 Competition between daunomycin and ruthenium compounds

In preliminary experiments, D2 and D3 were titrated with daunomycin and distamycin, respectively. The purpose of these titration experiments was to determine how much organic drug needed to be added to occupy all the intercalation or minor

groove binding sites. The results showed that all binding sites were occupied (became saturated) at a 10:1 ratio of organic drug to dsDNA, i.e. complexes with a higher number of drug molecules bound to DNA were not observed even when the drug:DNA ratio was increased further.

Figure 3.21 (a) shows the ESI mass spectrum obtained from a reaction mixture containing a 10:1 ratio of daunomycin and duplex D2. The two most abundant ions are at *m/z* 2066.0 and 2153.9, assigned to complexes containing five and six daunomycin molecules bound to D2, respectively. Ions corresponding to complexes containing different numbers of daunomycin molecules bound to DNA are also observed. These include ions at *m/z* 2241.9 and 2329.9 from complexes containing seven and eight daunomycin molecules bound to D2, respectively. This observation suggests that duplex D2 can readily bind six daunomycin molecules, and up to eight molecules in total. Such a conclusion is consistent with previous studies which showed for DNA molecules with different lengths that the average binding site size for daunomycin is every two base pairs.[81,313]

Figure 3.21 (b)-(d) show the ESI mass spectra obtained from reaction mixtures initially containing a 10:1 ratio of daunomycin and duplex D2, and subsequently treated with 30 equivalents of $[Ru(phen)_3]^{2+}$, 6 equivalents of $[Ru(phen)_2(dpq)]^{2+}$ and 6 equivalents of $[Ru(phen)_2(dppz)]^{2+}$, respectively. Assignments for all major ions observed in these experiments can be found in Appendix 2. Figure 3.21 (b) shows that the majority of ions assigned to non-covalent complexes containing daunomycin molecules and D2 that were present prior to the addition of $[Ru(phen)_3]^{2+}$ were now either absent or considerably reduced in abundance. In addition many new ions were

now also present. The most abundant of these new ions are those at *m/z* 2085.0 and 2172.6, which are assigned to [D2+4daunomycin+Ru(phen)$_3$-8H]$^{6-}$ (●) and [D2+5daunomycin+Ru(phen)$_3$-8H]$^{6-}$ (▲), respectively.

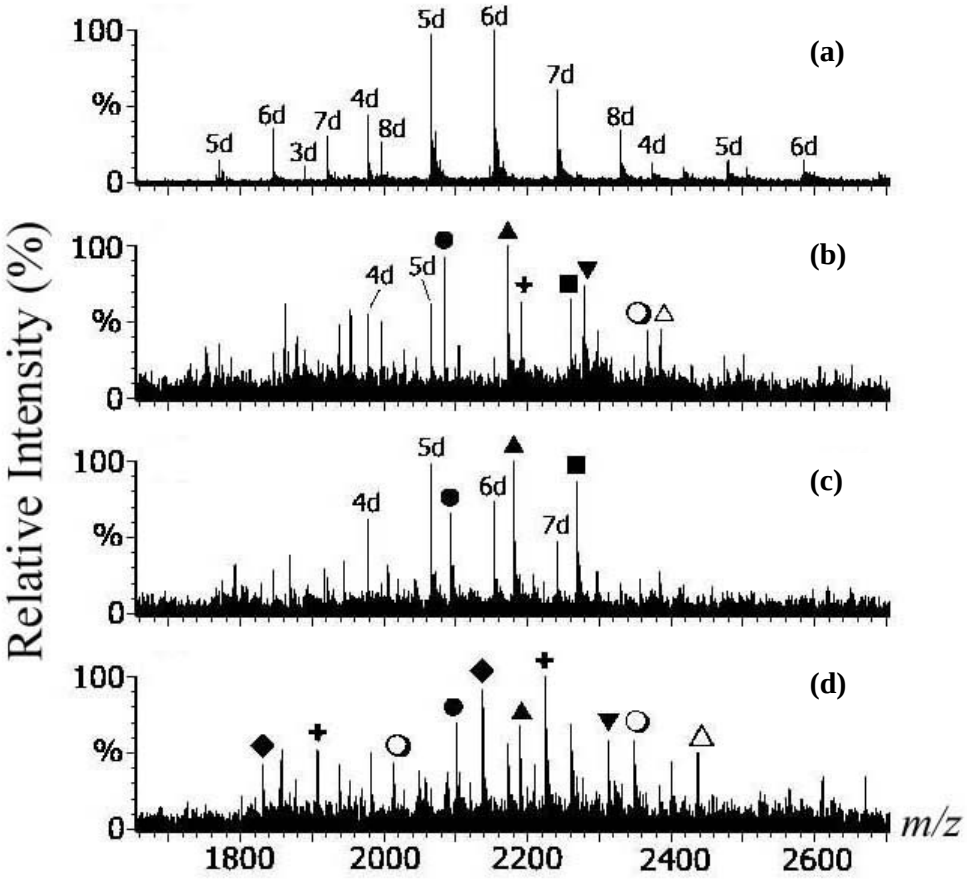


Figure 3.21 Negative ion ESI mass spectra of reaction mixtures containing D2 and: (a) 10 equivalents of daunomycin; (b) 10 equivalents of daunomycin and 30 equivalents of [Ru(phen)$_3$]$^{2+}$; (c) 10 equivalents of daunomycin and 6 equivalents of [Ru(phen)$_2$(dpq)]$^{2+}$; (d) 10 equivalents of daunomycin and 6 equivalents of [Ru(phen)$_2$(dppz)]$^{2+}$. 3d = dsDNA + 3daunomycin; 4d = dsDNA + 4daunomycin; 5d = dsDNA + 5daunomycin; 6d = dsDNA + 6daunomycin; 7d = dsDNA + 7daunomycin; 8d = dsDNA + 8daunomycin; ● dsDNA + 4daunomycin + 1[Ru(phen)$_2$(L)]$^{2+}$; ▲ dsDNA + 5daunomycin + 1[Ru(phen)$_2$(L)]$^{2+}$; ■ dsDNA + 6daunomycin + 1[Ru(phen)$_2$(L)]$^{2+}$; ◆ dsDNA + 3daunomycin + 2[Ru(phen)$_2$(L)]$^{2+}$; + dsDNA + 4daunomycin + 2[Ru(phen)$_2$(L)]$^{2+}$; ▼ dsDNA + 5daunomycin + 2[Ru(phen)$_2$(L)]$^{2+}$; ❍ dsDNA + 6daunomycin + 2[Ru(phen)$_2$(L)]$^{2+}$, △ dsDNA + 5daunomycin + 3[Ru(phen)$_2$(L)]$^{2+}$.

These complexes may arise via two pathways. The first involves addition of $[Ru(phen)_3]^{2+}$ to non-covalent complexes containing four or five daunomycin molecules already bound to D2. This would suggest that $[Ru(phen)_3]^{2+}$ has the ability to bind to different sites on D2 than daunomycin. However, the relatively high abundance of the ion assigned to $[D2+4daunomycin+Ru(phen)_3-8H]^{6-}$ in Figure 3.21 (b), compared to the medium abundance of the ion assigned to $[D2+4daunomycin–8H]^{6-}$ in Figure 3.21 (a), suggests that this is not the main pathway for formation of the former ion. The second possible pathway for formation of $[D2+4daunomycin+Ru(phen)_3-8H]^{6-}$ and $[D2+5daunomycin+Ru(phen)_3-8H]^{6-}$, is the replacement of one daunomycin from $[D2+5daunomycin+Ru(phen)_3-8H]^{6-}$ and $[D2+6daunomycin+Ru(phen)_3-8H]^{6-}$, respectively, by a $[Ru(phen)_3]^{2+}$ molecule. This pathway may appear to be unlikely at first glance since the reported binding constant for binding of daunomycin to calf-thymus DNA[245,312] is considerably greater than that for $[Ru(phen)_3]^{2+}$.[262] However, the concentration of $[Ru(phen)_3]^{2+}$ in this particular reaction mixture is much greater than that of daunomycin, and not all of the daunomycin bound to D2 would necessarily be bound to high affinity sites. Furthermore, the similarity in relative intensities of ions assigned to $[D2+5daunomycin-8H]^{6-}$ and $[D2+6daunomycin-8H]^{6-}$ in Figure 3.21 (a), compared to those assigned to $[D2+4daunomycin+Ru(phen)_3-8H]^{6-}$ and $[D2+5daunomycin+Ru(phen)_3-8H]^{6-}$ in Figure 3.21 (b), also provides evidence for the pathway involving displacement of daunomycin molecules.

Ions of low abundance at *m/z* 2367.4 and 2385.9 in Figure 3.21 (b) are assigned to $[D2+6daunomycin+2Ru(phen)_3-10H]^{6-}$ (❍) and $[D2+5daunomycin+3Ru(phen)_3-12H]^{6-}$ (△), respectively. Both complexes contain a total of eight molecules non-

covalently bound to D2, which is the same as the maximum number of daunomycin bound to D2 in the spectrum shown in Figure 3.21 (a). This is also consistent with the proposal that $[Ru(phen)_3]^{2+}$ cannot find alternative binding sites on duplex D2 molecules that have been saturated with daunomycin, but can instead displace some of the more weakly bound daunomycin molecules.

Analysis of ESI mass spectra of solutions obtained by adding 6 equivalents of $[Ru(phen)_2(dpq)]^{2+}$, $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$, $[Ru(phen)_2(pda)]^{2+}$ or $[Ru(phen)_2(dpqMe_2)]^{2+}$, to solutions already containing 10 equivalents of daunomycin and D2, gave similar results to those discussed above. In all cases there were no ions in the final solution that could be assigned to complexes containing more than a total of eight daunomycin and ruthenium molecules bound to D2. For example, Figure 3.21 (c) shows ions assigned to complexes containing four, five and six daunomycin molecules bound to D2 in addition to one $[Ru(phen)_2(dpq)]^{2+}$. Ions containing four to seven molecules of daunomycin bound to D2 are also present in medium to high abundance. This suggests that the overall degree of replacement of daunomycin by ruthenium molecules, in this case $[Ru(phen)_2(dpq)]^{2+}$, is not as great as that by $[Ru(phen)_3]^{2+}$. This is due largely to the lower ratio of ruthenium:duplex DNA used (6:1) for $[Ru(phen)_2(dpq)]^{2+}$ compared to $[Ru(phen)_3]^{2+}$ (30:1). Another explanation is that the greater size of the intercalating dpq ligand in $[Ru(phen)_2(dpq)]^{2+}$, compared to that of the phen ligand in $[Ru(phen)_3]^{2+}$, hinders binding by additional ruthenium molecules.

ESI mass spectra of solutions containing D2, 10 equivalents of daunomycin, and 6 equivalents of $[Ru(phen)_2(dpqC)]^{2+}$, $[Ru(phen)_2(dppz)]^{2+}$ or $[Ru(phen)_2(dpqMe_2)]^{2+}$

were very similar to each other, and accurately reflected the greater DNA binding affinity of these three ruthenium compounds compared to $[Ru(phen)_3]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$. For example, Figure 3.21 (d) shows an ESI mass spectrum of a reaction mixture prepared by adding 10 equivalents of daunomycin to D2, and subsequently adding 6 equivalents of $[Ru(phen)_2(dppz)]^{2+}$. Ions from complexes containing only daunomycin molecules bound to D2 are totally absent. The two most abundant ions were observed at *m/z* 2137.0 and 2224.9, which correspond to $[D2+3daunomycin+2Ru(phen)_2(dppz)-10H]^{6-}$ (◆) and $[D2+4daunomycin+2Ru(phen)_2(dppz)-10H]^{6-}$ (+), respectively. The most likely mechanism of formation of these complexes is replacement of two daunomycin molecules by $[Ru(phen)_2(dppz)]^{2+}$ from non-covalent complexes containing five and six daunomycin molecules bound to D2. Ions of low to medium abundance at *m/z* 1938.5, 2261.0, 2013.8, 2349.1 and 2437.5 are assigned to non-covalent complexes containing three, four or five daunomycin molecules as well as three $[Ru(phen)_2(dppz)]^{2+}$ molecules bound to D2. These observations support the earlier conclusion that $[Ru(phen)_2(dppz)]^{2+}$ displays a greater affinity for D2 compared to $[Ru(phen)_3]^{2+}$ and $[Ru(phen)_2(dpq)]^{2+}$. It is noteworthy that even in reaction mixtures containing the high affinity ruthenium compounds $[Ru(phen)_2(dppz)]^{2+}$ or $[Ru(phen)_2(dpqMe_2)]^{2+}$, there were no ions from non-covalent complexes containing more than a total number of eight daunomycin and ruthenium molecules. This set of experiments therefore suggests that the ruthenium molecule(s) can displace one or more daunomycin molecules from D2, suggesting that the two types of drugs display some similarities in their modes of DNA binding and/or their binding site preferences.

3.7.2.2 Competition between distamycin and ruthenium compounds

The D3 DNA sequence was chosen for competition experiments involving ruthenium compounds and the minor groove binding agent distamycin, since the latter is known to preferentially bind to DNA with AT-rich sequences. It was proposed that if ruthenium compounds could displace distamycin from D3 in these experiments, then this would imply that they can bind to the same DNA sequences. Such a result would not imply that the ruthenium complexes should be classed as minor groove binding agents. However, it would suggest that the ruthenium compounds can bind at the same regions of DNA.

Figure 3.22 (a) shows the ESI mass spectrum of a reaction mixture containing a distamycin:D3 ratio of 10:1. Ions of medium to high abundances from non-covalent complexes containing four, five, six and seven distamycin bound to D3 are present, with the most abundant being that containing six distamycin molecules. The number of distamycin molecules bound is reasonable since D3 contains two AAAA/TTTT base sequences, which are suitable for binding of distamycin. Furthermore it has been shown that distamycin molecules can lie side by side in the minor groove of DNA containing appropriate base sequences.[228,231,314] It is therefore likely that four distamycin molecules are tightly bound in the minor groove at the AAAA/TTTT binding sites, with the remaining distamycin molecules bound less tightly somewhere else. The ESI mass spectrum obtained after the addition of 30 equivalents of $[Ru(phen)_3]^{2+}$ to a reaction mixture already containing 10 equivalents of distamycin and D3 is shown in Figure 3.22 (b). Ions from complexes containing five to seven distamycin molecules bound to D3 are absent, with the exception of that at *m/z* 1946.7 assigned to D3 containing four distamycin molecules.

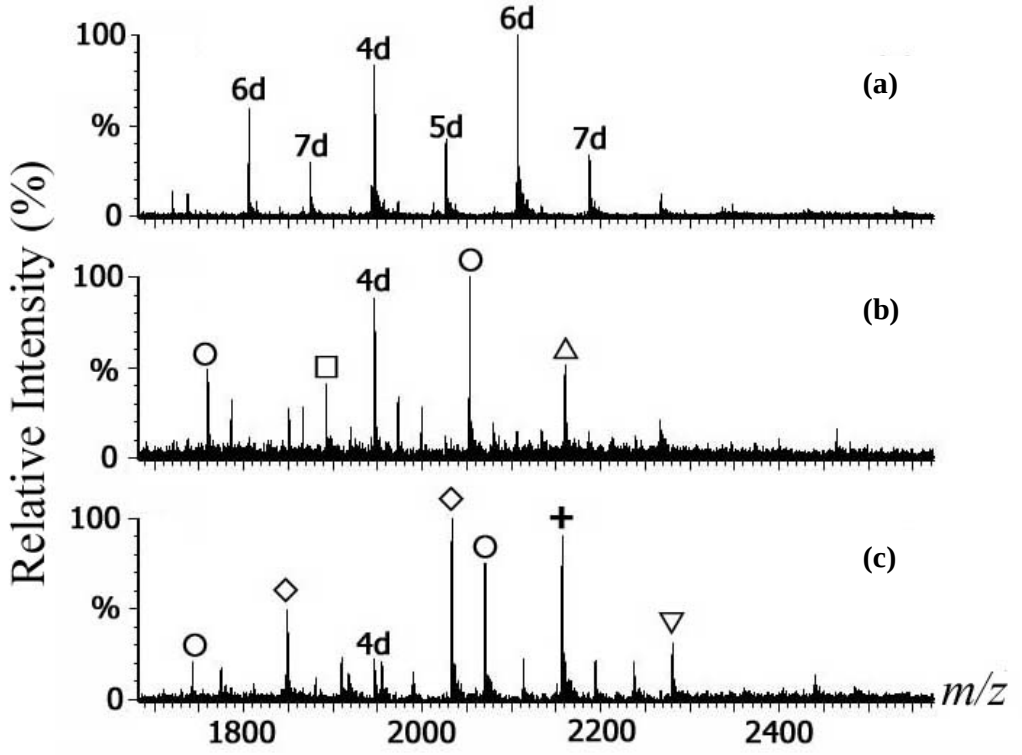


Figure 3.22 Negative ion ESI mass spectra of reaction mixtures containing D3 and: (a) 10 equivalents of distamycin; (b) 10 equivalents of distamycin and 30 equivalents of $[Ru(phen)_3]^{2+}$; (c) 10 equivalents of distamycin and 6 equivalents of $[Ru(phen)_2(dppz)]^{2+}$. 4d = dsDNA + 4distamycin; 5d = dsDNA + 5distamycin; 6d = dsDNA + 6distamycin; 7d = dsDNA + 7distamycin; ○ dsDNA + 4distamycin + 1$[Ru(phen)_2(L)]^{2+}$; △ dsDNA + 4distamycin + 2$[Ru(phen)_2(L)]^{2+}$; □ dsDNA + 2distamycin + 1$[Ru(phen)_2(L)]^{2+}$; ◇ dsDNA + 2distamycin + 2$[Ru(phen)_2(L)]^{2+}$; + dsDNA + 2distamycin + 3$[Ru(phen)_2(L)]^{2+}$; ▽ dsDNA + 2distamycin + 4$[Ru(phen)_2(L)]^{2+}$.

The most abundant ion observed in Figure 3.22 (b) is that at *m/z* 2053.5, assigned to $[D3+4distamycin+Ru(phen)_3-8H]^{6-}$. The 7- ion corresponding to this complex is also present in medium abundance at *m/z* 1760.0. The next most abundant ion is that at *m/z* 2160.2, assigned to $D3+4distamycin+2Ru(phen)_3-10H]^{6-}$. There are no ions from complexes containing more than a total of six distamycin and ruthenium molecules bound to D3. It is also important to note that the abundance of the ion at *m/z* 1946.7 assigned to a non-covalent complex containing four distamycin molecules bound to D3 is similar in both Figure 3.22 (a) and (b). The persistence of this ion is consistent

with the statement above, that four distamycin molecules are probably bound tightly to the minor groove of D3 at the AAAA/TTTT binding sites. The major mechanism of formation of ions containing both types of drug molecules bound to D3 is probably by displacement of one or more of the more loosely held distamycin molecules from ions such as $[D3+6distamycin-6H]^{6-}$ and $[D3+7distamycin-6H]^{6-}$. Furthermore, the absence of ions containing five distamycin and one or two ruthenium molecules bound to D3 in Figure 3.22 (b), suggests that one $[Ru(phen)_3]^{2+}$ is capable of displacing two loosely bound distamycin molecules.

Even stronger evidence suggesting that ruthenium molecules can bind to the minor groove of duplex DNA (but not necessarily as a classical minor groove binder) was provided by studying the effect of adding 6 equivalents of either $[Ru(phen)_2(dpqC)]^{2+}$ or $[Ru(phen)_2(dppz)]^{2+}$, to a reaction mixture already containing 10 equivalents of distamycin and D3. Figure 3.22 (c) shows the ESI mass spectrum of the reaction mixture containing $[Ru(phen)_2(dppz)]^{2+}$. The only ion assigned to a complex containing only distamycin molecules bound to D3 is again that at *m/z* 1946.7, assigned to $[D3+4distamycin-6H]^{6-}$. The persistence of this ion even in the presence of the ruthenium compound with the highest DNA affinity, suggests that it has a very stable structure. The ion is present in significantly reduced abundance compared to the spectra in Figure 3.22 (a) and (b), suggesting that $[Ru(phen)_2(dppz)]^{2+}$ is not only capable of binding to D3 at base sequences in the minor groove, but also has a comparable DNA binding affinity to distamycin.

An ion of medium abundance at *m/z* 2070.0, assigned to $[D3+4distamycin+1Ru(phen)_2(dppz)-8H]^{6-}$ (○), is most likely formed by

displacement of the weakly held fifth and sixth distamycin molecules from ions such as $[D3+6distamycin-6H]^{6-}$. In Figure 3.22 (c) the two most abundant ions, at *m/z* 2034.0 and 2157.2, are assigned to $[D3+2distamycin+2Ru(phen)_2(dppz)-10H]^{6-}$ (◇) and $[D3+2distamycin+3Ru(phen)_2(dppz)-12H]^{6-}$ (+), respectively. These as well as other ions at *m/z* 1848.8 and 2280.6, assigned to $[D3+2distamycin+2Ru(phen)_2(dppz)-11H]^{7-}$ and $[D3+2distamycin+4Ru(phen)_2(dppz)-14H]^{6-}$ (▽), respectively, were most probably formed by displacement of three or four distamycin molecules from ions such as $[D3+6distamycin-6H]^{6-}$, $[D3+6distamycin-7H]^{7-}$ and $[D3+7distamycin-6H]^{6-}$. While the first two or three distamycin molecules displaced from the latter ions may have been weakly bound to D3, the last two would probably have been bound more tightly to the AT-rich sequences in the minor groove. This provides further evidence that $[Ru(phen)_2(dppz)]^{2+}$ can bind to AT-rich base sequences.

3.8 Conclusions

There have been many studies of the binding of metallointercalators to DNA, using a variety of techniques including chemical, biochemical and spectroscopic methods. There are, however, many questions that still need to be answered, for instance, identifying the actual binding sites on large, heterogeneous molecules such as calf-thymus DNA. In addition, many spectroscopic techniques are not well suited to analysing solutions containing mixtures of non-covalent complexes. However, the results presented here clearly show that ESI-MS can readily provide information on the number, relative amounts and stoichiometry of non-covalent complexes present in solutions containing up to five different metal-DNA complexes. This is a result of

the combination of extremely high sensitivity inherent to ESI-MS, and the simplicity of the resulting spectra. In addition, speed of analysis makes ESI-MS an extraordinarily attractive choice for large-scale screening of the DNA binding properties of metal complexes. For the same reasons, ESI-MS is also suited to monitoring competition experiments involving metal complexes and organic drugs. The results obtained from competition experiments provided evidence that some of the ruthenium compounds studied probably bind to DNA at similar base sequences to where daunomycin and distamycin bind. However, there are alternative explanations for the observed results that do not rely on the ruthenium compounds directly displacing the organic drugs from their DNA binding sites. For instance, the binding of a ruthenium compound somewhere else along the DNA helix might induce a conformational change that results in less tightly bound organic drug molecules. It is also possible that electrostatic interactions between positively-charged ruthenium and distamycin molecules located in close proximity to each other lead to the displacement of the latter molecules.

Information about the relative binding affinities and sequence selectivities of the six ruthenium compounds examined towards three different duplexes was also obtained using ESI-MS. The relative order of binding affinities was found to be: $[Ru(phen)_2(dppz)]^{2+} \geq [Ru(phen)_2(dpqMe_2)]^{2+} > [Ru(phen)_2(dpqC)]^{2+} > [Ru(phen)_2(dpq)]^{2+} > [Ru(phen)_2(pda)]^{2+} > [Ru(phen)_3]^{2+}$. A very similar order was obtained from DNA melting curve experiments, providing support for the use of ESI-MS in analysing non-covalent interactions between metal complexes and DNA. The order of relative binding affinities obtained from both techniques is consistent with

the proposal that the strength of binding is heavily dependent on the size, and planarity of the unique intercalating ligand.

Most ruthenium compounds examined here showed a greater capability to form non-covalent complexes with duplex D2 than either D1 or D3. This was most evident with those ruthenium compounds with high DNA affinity, and suggests that D2 has either a greater number of binding sites and/or its binding sites are more attractive to these compounds. This was quite interesting as it was expected that ruthenium compounds with intercalating ligands such as dpqC and dppz would prefer the DNA sequence with greater GC content (D1). One possible explanation is that although D1 contains a greater number of potential binding sites, their close proximity to each other results in unfavourable steric and/or electrostatic interactions between ruthenium compounds. On the other hand D2 contains a smaller number of suitable binding sites that are sufficiently far apart to result in very strong binding interactions.

A final question still to be answered is where exactly do these ruthenium compounds bind to these DNA duplexes. This particular question cannot easily be answered by ESI-MS alone. However, ESI-MS may be able to provide some clues, for example by performing additional binding studies using other 16-mer dsDNA molecules with different base sequences, or by performing partial enzymatic digestion of reaction mixtures containing ruthenium-DNA complexes using exonucleases. However, the latter experiment may be extremely difficult due to the length of the DNA duplexes used and the number of ruthenium molecules bound to DNA. Comparison of the results presented here with those obtained from binding studies involving smaller,

related DNA duplexes might provide further insight into the metal binding sites. These studies are in progress, in parallel with other spectroscopic investigations to determine binding constants for the interactions of these ruthenium compounds with D1, D2 and D3.

Chapter 4
Investigation of Interactions of Metal Ions with the Exonuclease Subunit of *E. coli* DNA Polymerase III

4.1 Introduction

This chapter presents the results of an ESI-MS investigation to determine the binding affinities of various metal ions for the catalytic N-terminal domain of the exonuclease proofreading subunit of *Escherichia coli* DNA polymerase III, ε. Epsilon (ε) has an important role in replication, preventing accumulation of mutations in the *E. coli* genome. An overview of the process of DNA replication is given below, in addition to a description of DNA polymerases, the structure and function of the epsilon (ε) subunit of DNA polymerase III, and an introduction to the roles of metal ions in enzymes.

DNA replication involves three steps: initiation at the origin(s) of replication, synthesis or elongation of DNA at replication forks, and finally, termination of replication. Each of these steps is mediated by multiple protein-protein and protein-DNA interactions, which in *E. coli* involves subassemblies of around thirty different replication proteins.[315] The entire multiprotein-nucleic acid complex is called the replisome. Studies of the mechanism of DNA replication have made extensive use of *E. coli* proteins. This is because they can be purified in large quantities from overproducing strains. Over the last ten years, high-resolution structures of many of the individual proteins (or protein domains) and some large complexes of the *E. coli*

replisome have become available.[177,316-318] The replisome consists of a replicase (DNA polymerase) and a primosome (helicase/primase). In this chapter the focus is on the DNA polymerase, while the primosome will be the focus in Chapter 5. To date, an intact replisome has not yet been isolated in sufficient quantity for biophysical measurements. However, functional and structural studies have been possible by *in vitro* assembly of individual subunits.

4.2 Replication in Escherichia coli

The *E. coli* replisome is the most extensively studied biochemically and genetically,[319-323] and is therefore used as a model for replication in other organisms. This is reasonable because proteins involved in the replication process in all organisms are involved in complex macromolecular assemblies that have highly conserved functions. For example, the protein in *E. coli* that is responsible for holding the replication fork to the DNA polymerase (the beta (β) sliding clamp protein) has a very similar overall three dimensional structure to the analogous protein (the proliferating cell nuclear antigen (PCNA)) in archaea and eukaryotes [319,324,325].

Replication is initiated at a 260-bp sequence located in the 4.6 million base pairs (bp) of the circular double-stranded (ds) DNA genome of *E. coli*. This region is referred to as the origin, *ori*C. During the replication process, dsDNA is separated into two replication forks, running in two opposite directions from the origin, and the replisome components are loaded onto single-stranded (ss) DNA templates. In order to initiate replication, the protein DnaA must first recognise five copies of a 9-bp AT-rich DnaA-box within *ori*C.[326,327] When small basic histone-like proteins (HU

and/or IHF) which stabilise DNA also bind to *ori*C the DNA is separated into two strands.[327,328] In the next step of initiation, two molecules of the hexameric DnaB ($(DnaB)_6$) helicase are loaded onto the single-stranded (ss) DNA with the help of another protein, DnaC, resulting in unwinding of DNA.[326] Interactions of DnaB and DnaC will be discussed in further detail in chapter 5. Polymerases cannot start synthesising polynucleotide chains until each nascent DNA fragment is primed by synthesis of short RNA primers by the primase protein, DnaG.[329]

The next step is to load the replicative polymerase at the primer termini. On the leading strand, the primer stays attached to the parental ssDNA allowing the polymerase to extend the length of DNA by addition of complementary nucleotides. In contrast, replication of the lagging strand occurs discontinuously producing Okazaki fragments.[330] The primase stays attached to the lagging strand at the end of each fragment until a new primer is made. Finally, the former RNA primer is removed by DNA polymerase I, which also fills in the gap between each fragment and is finally joined by DNA ligase.[331] Termination of DNA replication in *E. coli* is mediated by a tight complex of Tus protein and a 23-bp *Ter* DNA sequence in the region of the chromosome opposite *oriC*.[332]

4.3 DNA Polymerases

The biochemical elucidation of the mechanism of DNA polymerisation began in the mid-1950s with the discovery of DNA polymerase I (Pol I) in *E. coli* by Arthur Kornberg.[333] Soon after, DNA polymerases II and III (Pol II and III, respectively) were discovered.[334] Studies on these polymerases have provided critical information

that is important for understanding the mechanism of replication in different organisms.

Not long after its discovery, it was shown that the properties of the Pol I enzyme were inconsistent with those expected for an enzyme catalysing chromosomal DNA replication, suggesting that Pol I could not be the primary *E. coli* polymerase. For example, there are about 400 molecules of this enzyme per cell, making it an abundant protein in vast excess over the small number of replication forks present (<10).[331,335] Secondly, the enzyme only catalyses the addition of about 20 nucleotides per second,[336] while the replicative chain growth *in vivo* occurs at around 800 nucleotides per second.[337] Many more studies have also subsequently shown that DNA Pol I is not suitable for rapid, efficient and accurate DNA replication.[320,331,338,339]

Based on amino acid sequence comparisons and crystal structures, DNA polymerases have been classified into six different families.[338,340] These are the A, B, C, X, RT (reverse transcriptase), and UmuC/DinB families. Family A polymerases are found primarily in organisms related to prokaryotes, and include prokaryotic DNA polymerase I (Klenow fragment of *E. coli*), mitochondrial polymerase γ and odd-numbered bacteriophages such as T3, T5 and T7. Family B polymerases are present in bacteriophages, viruses, archaea and eukaryotes. Many of these polymerases function to replicate the host genome, and include those from even-numbered phages such as T4 and T6, herpes viruses, archaeal pol "Vent", and mammalian pol α, δ and ε. Family C polymerases encompass those that replicate the majority of bacterial genomes, including the subject of the current work, *E. coli* DNA polymerase III.

Family X contains mammalian pol β, λ and μ, which function during DNA repair. The reverse transcriptase family contains RTs from retroviruses as well as eukaryotic telomerases. The recently discovered UmuC/DinB family includes pol η, i and κ, and deoxycytidyl transferase.[341]

4.4 DNA Polymerase III Holoenzyme

The function of DNA polymerase III (Pol III) holoenzyme is replication of the *E. coli* chromosome. DNA polymerase III contains ten different subunits: alpha (α), epsilon (ε), theta (θ), tau (τ), gamma (γ), delta (δ), delta prime (δ′), chi (χ), psi (ψ) and beta (β), working together as an efficient, processive and high fidelity holoenzyme.[320] The probable stoichiometric assembly of the Pol III subunits is $(\alpha\varepsilon\theta)_2$-$(\delta'\gamma\tau_2\delta\chi\Psi)$-$(\beta_2)_2$.[315,319] Figure 4.1(a) shows a model of protein-protein and protein-DNA interactions in the *E. coli* replisome. The DNA replicase shown in Figure 4.1(b) (circled in Figure 4.1 (a)) consists of three separate subassemblies: the core $(\alpha\varepsilon\theta)$,[342-344] the β_2 DNA sliding clamp and the clamp loader or DnaX complex $(\delta'\gamma\tau_2\delta\chi\Psi)$.[319,345,346]

DNA polymerase III contains two core polymerase assemblies, $(\alpha\varepsilon\theta)_2$, which are responsible for simultaneously replicating the leading and lagging strands. The large 130 kDa α subunit, the product of the *dnaE* gene, contains the polymerase active site,[347] while the 27.5 kDa ε subunit, the product of the *dnaQ* gene, contains the 3′→5′ exonuclease that serves as a proofreader for replication errors.[337,348] The function of the smaller 9 kDa θ subunit, the product of the *holE* gene, is as yet

undetermined.[337] A slight stimulation of ε activity on a mismatched T-G base pair[349] has been observed, suggesting that θ may function to stabilise the ε subunit.[350,351]

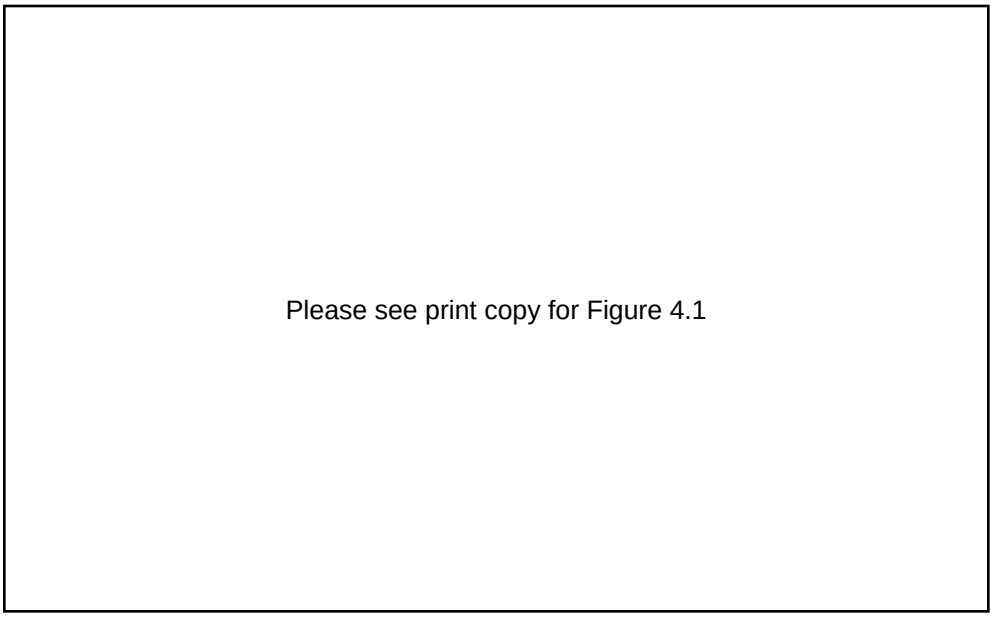


Figure 4.1 Structural model showing the stoichiometry of DNA polymerase III holoenzyme subunits. (a) Composition of the *E. coli* replisome, comprising the replicase (DNA polymerase III holoenzyme; circled) and primosome (DnaB and DnaG) and (b) close up of the DNA polymerase III holoenzyme; circled in (a). Taken from Schaeffer *et al*..[315]

4.4.1 Epsilon (ε)

The very high fidelity of Pol III is achieved through the exonuclease activity of ε which serves as a DNA proofreader. The overall error rate of replication in *E. coli* is ~10^{-10} per 4.6 million base pairs,[322,352] of which α contributes about 10^{-5}-10^{-7}, and proofreading contributes ~10^{-2}-10^{-3}.[322,353] The ε subunit is composed of two domains. The first of these is a 20.5 kDa N-terminal domain (residues 2-186), called ε186, which contains the exonuclease active site and forms a stable 1:1 complex with θ.[177,321,323,349] The second domain is a small C-terminal domain (residues 187-243) that interacts with α.[323,354] Unlike most of the other proofreading polymerases that

belong to the same family,[355] the polymerising and exonucleolytic activities of Pol III are present on two separate subunits. The polymerase activity is contained in the α subunit, while the 3′→5′ exonuclease activity is contained in the ε subunit. Amino acid alignments among polymerase-associated 3′→5′ exonucleases have revealed homologous regions containing conserved amino acid residues (Exo motifs), designated Exo I, Exo II and Exo III, with each motif also containing several highly conserved residues.[356] However, studies on the 3′-exonuclease of the *Bacillus subtilis* DNA polymerase III suggested that while the Exo III motif is missing, an alternative motif called Exo III-ε was present.[357] This name was chosen as amino acid alignments identified a similar motif in the *E. coli* Pol III ε subunit.[321,356-358] These conserved motifs in the ε subunit of DNA Pol III contain carboxylate residues presumed by analogy with the structure of the corresponding domain of DNA Pol I,[356,359] to interact with two divalent metal ions that participate in phosphodiester bond cleavage.[360,361]

Many research groups have focused on the N-terminal domain of ε, ε186, since it contains the exonuclease active site,[177,323,362] and since full length ε has proven to be unsuitable for NMR analysis as a result of protein precipitation even at low temperatures and concentrations.[351] However, obtaining structural information on ε186 still poses a number of difficulties, owing to the limited stability of the isolated catalytic domain at high concentrations,[351] and a tendency to aggregate at elevated temperatures.[177] A structure of ε186 was determined using a combination of X-ray crystallography and NMR spectroscopy.[177] More recently NMR data were used in conjunction with molecular modelling techniques to obtain a model of the structure of the ε186 domain,[351] which is in agreement with the X-ray structure.[177,362]

The X-ray crystallographic structure of ε186 treated with Mn^{2+} and the inhibitor thymidine-5′-monophosphate (TMP) revealed that the active site contained a complex with two Mn(II) ions and a molecule of TMP. The structure also revealed that ε186 has a topology similar to that of seven other DNA polymerase proofreading domains from the Pol A and Pol B families, despite having low sequence homology overall. All have a central five stranded β sheet with a long C-terminal helix packed against it, which is consistent with the structure obtained from NMR studies.[351]

4.5 Metal Ions in Proteins and Enzymes

Metal ions have numerous important roles in biological systems. It has been estimated that around 40% of all proteins and enzymes contain metal ions in their structures.[363,364] Metal ions are important for many metabolic processes, such as biological energy conversion in photosynthesis and respiration. Metal ions in proteins are involved as catalysts (in substrate binding and activation), or have roles in transport (e.g. O_2 bound to iron in haemoglobin) and storage (e.g. iron storage in ferritin).[365,366] A variety of metal ions can bind to proteins or enzymes. However, the strongest binding interactions involve Mg^{2+} and transition metals such as Fe^{2+}, Mn^{2+} and Zn^{2+}, owing to their high charge density, and ability to form strong coordinate and electrostatic interactions with functional groups on amino acids.[364,365] Transition metals present at the active site of enzymes are usually involved in the catalysis of redox reactions (e.g. Fe^{3+}/Fe^{2+} in cytochromes), hydrolysis reactions (e.g. Zn^{2+} in carbonic anhydrase) or phosphoryl transfer reactions (e.g. Mn^{2+} in sweet potato purple acid phosphatase).[367]

Phosphoryl transfer reactions are ubiquitous in biological systems and the enzymes that catalyse these types of reactions often require divalent metal ions for maintenance of correct tertiary structure and/or catalytic activity.[368] Many phosphatases catalyse the cleavage of phosphate ester bonds using the metal ion as a Lewis acid, which polarises protein functional groups to which it binds. In addition, the metal often lowers the pK_a of a bound water molecule thereby creating an effective nucleophile, or directly coordinates a nucleophile at the active site.[369,370] X-ray crystallography[177,371-373] and NMR spectroscopy[362,371,374-376] have been employed to obtain information about how metal ions bind to proteins. Other techniques such as UV spectrophotometry,[377,378] fluorescence spectroscopy,[379,380] electron paramagnetic resonance (EPR),[381-383] calorimetry[384] and circular dichroism (CD) spectroscopy[385] have also been used to examine the interactions of metal(s) and proteins. In order to help gain greater understanding of how metal(s) and protein interact, dissociation/binding constants can be determined. These values have generally been obtained by spectroscopic methods such as visible and fluorescence spectroscopies.[379,383,386] These techniques, however, cannot give an unequivocal determination of the metal binding stoichiometry.

Some metal ions (such as Mg^{2+}, Ca^{2+} and Zn^{2+}) are difficult to study by optical spectroscopic, electrochemical or magnetic resonance methods. These spectroscopically and redox silent cations must therefore be substituted by other metal ions. The ability of some lanthanide ions, such as Tb^{3+} and Eu^{3+}, to luminesce in aqueous solution at room temperature allows them to be used as sensitive spectroscopic probes for determining similarities and differences in the metal ion binding sites of different proteins and enzymes.[387] Lanthanide ions have long been

used to examine the structure, function and metal-binding properties of proteins or enzymes,[388-391] especially those which bind Ca^{2+}.[392-394] They can be used as an alternative means to gain information on the metal-binding properties of both wild type and mutant enzymes.[387] The trivalent lanthanide ions and Ca^{2+} are very similar in size,[387] and hence have similar chemical properties, making the lanthanides useful Ca^{2+} analogues. For example, Atreya *et al*. used Yb^{3+} to displace the Ca^{2+} in a protozoan protein, and compared the relative binding specificities and affinities for the two metal ions. The study showed that Yb^{3+} sequentially displaced Ca^{2+} from the four metal binding sites of the protein.[394] Ye *et al*. employed La^{3+} and Tb^{3+} to obtain information about the binding affinity of calcium for individual Ca^{2+}-binding loops in calmodulin.[392]

Lanthanide ions such as La^{3+} and Dy^{3+} have also been used successfully in NMR studies as paramagnetic probes of metal binding sites.[387,395,396] The successful application of lanthanide ions as NMR probes relies on their extremely short electron relaxation times, which minimise broadening effects on NMR spectra of small molecules. In a recent study of protein-protein interactions in the θ-ε186 complex of Pol III of *E. coli*, a single lanthanide ion, either La^{3+} or Dy^{3+}, replaced two Mn^{2+} ions at the metal binding site.[396] The effects of metal substitution were observed using ^{15}N HSQC (heteronuclear single quantum correlation) NMR spectroscopy. The NMR data obtained from spectra of the θ-ε186 complex containing these paramagnetic lanthanide ions provided reliable and detailed structural information about the complex.[396]

4.5.1 Metal ion involvement in exonuclease activities of Pol I and Pol III

Unlike Pol III, the polymerase and exonuclease activities of Pol I are present on the same polypeptide chain in two distinct domains which are physically separated by ~30 Å.[387,397] Studies of the 3′→5′ exonuclease from the large Klenow fragment (KF exo) of DNA polymerase I from *E. coli* revealed the enzyme contains two divalent metal ions. These two ions are in close proximity to one another, ~4 Å apart, and are both essential for catalysis.[360,361,398] The proposed mechanism of catalysis appears to be analogous to that found for a number of different enzymes that catalyse similar phosphoryl transfer reactions, including other DNA polymerases,[361,399,400] alkaline phosphatase[401] and the RNaseH domain of HIV-1 reverse transcriptase.[402]

The structure and biochemical properties of KF exo have been examined extensively. The crystal structure of the enzyme has been obtained in the presence of both a single-stranded DNA substrate and the deoxynucleotide thymidine monophosphate (dTMP) product.[361,397,398] Complexes with dTMP bound to the exonuclease active site have shown that two metal ions bind and are in contact with the phosphate and several acidic amino acid residues.[400] One of the two divalent metal ions is proposed to help in orientating the substrate and in activation of an incoming nucleophile, OH^-. The nucleophile attacks the scissile phosphate, generating a pentacoordinate transition state, which is stabilised by both metal ions. Mutant proteins that cannot bind the second metal ion show considerably lower activity.[398] While the identity of the metal ions in the native protein has still not been determined, it has been shown that binding of Mg^{2+}, Mn^{2+}, Zn^{2+} and Co^{2+} to the active site of KF exo makes the protein active in DNA hydrolysis.[361,398,403] It has also been shown by NMR

spectroscopy and molecular modelling[351] that the KF exonuclease and the N-terminal domain of the ε subunit, ε186, have very similar amino acid sequences. Furthermore, ε186 also requires two divalent metal ions to catalyse hydrolysis of nucleotide substrates such as the *p*-nitrophenyl ester of thymidine-5′-monophosphate (*p*NP-TMP).[173,362]

In order to determine the identity of the native metal(s) in a protein it is necessary to purify it under conditions where the metal(s) remain bound. However, in the case of ε from *E. coli* Pol III, this is a difficult task since there are only 10-20 molecules of Pol III per cell.[320] Hamdan *et al.* prepared crystals of recombinant ε186 in the presence of Mn^{2+} to determine the X-ray structure of the ε186·$Mn(II)_2$·TMP complex, and showed that the architecture of its active site is closely related to that of Pol I.[362] Catalysis of the hydrolysis reaction by ε most likely involves the two divalent metal ions and a histidine residue (His162), which takes the place of Tyr497 in Pol I.[362] A mechanism for hydrolysis of phosphodiester bonds by ε, proposed by Hamdan *et al.*[362] is shown in Figure 4.2. The diagram shows that one phosphate oxygen atom of the substrate (*p*NP-TMP) bridges the two manganese ions (labelled A and B in the diagram). The metal ions polarise the P-O bond, orienting the phosphate group and allowing in-line attack by hydroxide ion coordinated to one Mn^{2+} ion (Mn_A). The hydroxide ion is generated by deprotonation of a water molecule by the basic His162 residue in the active site. The 3′-oxygen of the ester coordinates to the second Mn^{2+} ion (Mn_B) in the trigonal bipyramidal transition state, assisting the 3′-OH of the nucleotide product to leave.

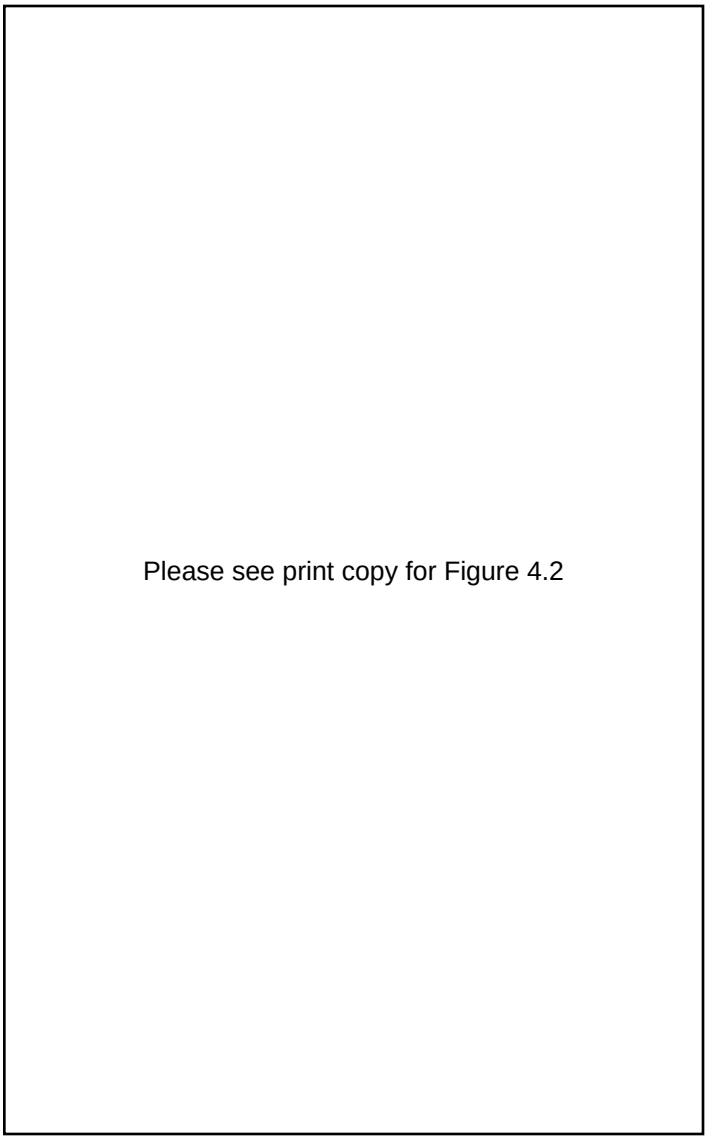


Figure 4.2 Proposed mechanism for hydrolysis of phosphodiester bonds by the ε subunit of DNA polymerase III. The enzyme-substrate, ε-*p*NP-TMP, complex (a) is based on the structure of the ε186·Mn(II)$_2$·TMP complex at pH 8.5, whereas the enzyme-product, ε-TMP, complex (c) is based on the complex structure at pH 5.8. In addition a proposed structure for the transition state is shown in (b). The two manganese ions in the active site of the enzyme are labelled A and B. Taken from Hamdan *et al.*.[362]

The ability of different metal ions to promote the exonuclease activity of ε186 has been investigated using a continuous spectrophotometric assay. Hamdan *et al.* investigated phosphate ester hydrolysis catalysed by the ε186 subunit using the *p*-

nitrophenyl ester of thymidine 5′-monophosphate (*p*NP-TMP) as the substrate.[173] Hydrolysis of *p*NP-TMP to *p*-nitrophenol and TMP was monitored to determine the rates of nucleotide phosphodiester hydrolysis by ε186. It was shown that hydrolysis rates were dependent on the presence of metal ions such as Mn^{2+} and Mg^{2+}. In addition, the rates were inhibited by TMP, a nucleotide product of the exonuclease reaction.[173]

There are few methods available that can determine the precise stoichiometry and binding affinity of metal ions for proteins. However, in recent years, ESI-MS has shown great promise for these types of studies.[133,404-406] For example, Zhu and co-workers investigated a protein-metal complex using a combination method called "PLIMSTEX" (protein-ligand interaction using mass spectrometry, titration and hydrogen/deuterium exchange). They were able to detect conformational changes, and determine the binding stoichiometry and binding/dissociation constants for a complex of calmodulin and Ca^{2+} under various conditions including in the presence or absence of Li^{+}, Na^{+} and K^{+} cations.[407] Several other studies have shown that ESI-MS is an effective technique for examining protein conformational changes and quantifying protein-ligand interactions.[63,65,404] In addition, ESI-MS has been used to analyse metal binding selectivities by examining the direct competition between different metal ions for the metal binding sites of a metalloenzyme.[408,409]

4.6 Scope of This Chapter

The aim of the work described in this chapter was to use ESI-MS to examine interactions between ε186 and three different metal ions (Mn^{2+}, Zn^{2+} and Dy^{3+}). The

manganese(II) ion was chosen as it is likely to be the native metal ion owing to a higher rate of enzyme-catalysed hydrolysis observed by Hamdan *et al*. when it was present, compared to the rate when Mg^{2+} was present.[173] The zinc(II) ion was studied because it is a common cofactor in many enzymes,[367,410] while the dysprosium(III) ion was examined because it has been used to assist structural analysis of ε186 by NMR spectroscopy.[396] From the ESI-MS data, information regarding the relative binding affinities of the different metal ions for the protein was obtained. The relative abilities of these metal ions to promote hydrolysis of *p*NP-TMP by ε186 were compared using the spectrophotometric assay developed by Hamdan *et al.*.[173]

4.7 Results and Discussion

4.7.1 Binding of metal ions (Mn^{2+}, Zn^{2+} and Dy^{3+}) to ε186

In the following experiments, ESI-MS was used to estimate dissociation constants for the binding of different metal ions to ε186. In preliminary experiments, mixtures of ε186 and a single metal ion were analysed by ESI-MS at various time points. ESI mass spectra acquired after several minutes were essentially the same as those acquired after longer periods of time (data not shown). Therefore, in all subsequent experiments, mixtures containing protein and metal ions were allowed to react for 15 minutes (at 0 °C) prior to analysis in order to maintain the stability of ε186.

Figure 4.3 shows positive ion ESI mass spectra of ε186 (2 μM) in the absence and presence of increasing concentrations of Mn^{2+}.

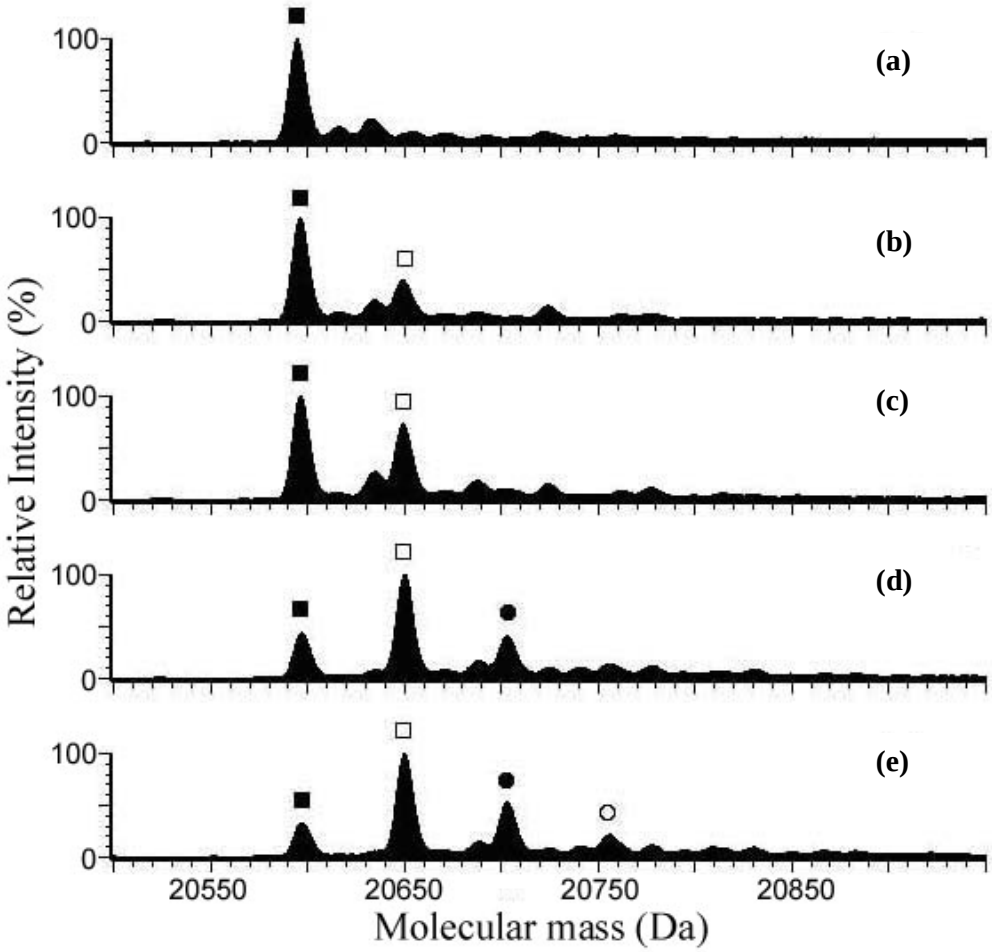


Figure 4.3 Positive ion ESI mass spectra (transformed to a mass scale using MassLynx software™) of ε186 (2 μM) with increasing Mn^{2+} concentrations. (a) ε186 with no Mn^{2+} present (control); (b) ε186 with 20 μM Mn^{2+}; (c) ε186 with 40 μM Mn^{2+}; (d) ε186 with 80 μM Mn^{2+} and (e) ε186 with 120 μM Mn^{2+}. ■ ε186 alone; □ ε186 with 1 Mn^{2+} bound; ● ε186 with 2 Mn^{2+} bound; ○ ε186 with 3 Mn^{2+} bound.

The peak in Figure 4.3 (a) corresponds to a mass of 20857 Da (■) which is the mass of ε186 with no metal bound (calculated M_r 20856). Peaks in Figure 4.3 (b)-(e) at 20642 (□), 20697 (●) and 20752 (○) correspond to ε186 with one, two, and three Mn^{2+} ions bound, respectively. As the concentration of metal ion was increased, a reduction in the amount of free ε186 was observed, indicating that the metal binding sites in ε186 were being occupied by Mn^{2+} ions. The relative intensity of the peak

assigned to ε186 with one Mn^{2+} bound (ε186 + 1 Mn^{2+}) continued to increase until the Mn^{2+} concentration was 80 μM (Figure 4.3 (d)). At this Mn^{2+} concentration a peak from an ε186 + 2 Mn^{2+} complex was also present in the ESI mass spectrum.

As the Mn^{2+} concentration was increased further (up to 600 μM), complexes of ε186 with up to five Mn^{2+} were present, albeit at low abundance (data not shown). The crystal structure of ε186 obtained by Hamdan *et al*. showed that two metal ions were present in the active site of ε186.[362] It has also been reported that when Mn^{2+} is present at high concentrations (mM) a third metal ion can bind to the protein (unpublished; Dixon NE, personal communication). The presence of a third metal ion was also observed in the X-ray crystal structure of the functionally analogous KF exonuclease domain of *E. coli* DNA polymerase I,[403] which has a similar structure to ε186. Since complexes with three to five metal ions bound to ε186 were detected only at high metal concentrations in the ESI-MS experiments reported here, this suggests that binding of the third, fourth and fifth metal ions only occurs as a result of non-specific interactions between negatively charged residues (e.g. aspartate and glutamate) of ε186 and the positively charged metal ions. Non-specific binding of metal ions to proteins when the metal ions were present in excess has been observed for other proteins. For example, calmodulin is known to bind to four Ca^{2+} ions. However, at high metal concentrations, protein complexes containing five Ca^{2+} were detected using ESI-MS.[84]

If the third, fourth and fifth metal ions are bound non-specifically to ε186, it is likely that they will be bound less tightly than the first two metal ions. An attempt was therefore made to compare the stabilities of complexes of ε186 with the first two

metal ions bound, and with the third, fourth and fifth metal ions bound, by in-source and collision cell collision-induced dissociation (CID),[13,32,411-413] and by attempting thermal dissociation through increasing the desolvation temperature.[32,164] Increasing the cone voltage from 100 to 700 V, desolvation temperature from 100 to 500 °C and/or collision energy from 2 to 30 V, did not dissociate all five metal ions from the protein. These conditions resulted in spectra of poor quality, with total ion counts reduced by a factor of up to 20. Therefore the stabilities of these protein-Mn^{2+} adducts were examined by comparing ESI mass spectra before and after dialysis of the reaction mixtures against 100 mM NH_4OAc, pH 8.0 at 4 °C for three hours. Figure 4.4 shows ESI mass spectra of ε186 that had been treated with a 500-fold excess of Mn^{2+} before and after dialysis. The mass spectrum obtained before dialysis (Figure 4.4 (a)) showed up to five Mn^{2+} ions were bound to ε186. In contrast, the spectrum obtained after the mixture was dialysed for 3 hours (Figure 4.4 (b)) showed there were no ions corresponding to ε186 with more than two Mn^{2+} bound. The major peaks in the latter spectrum were from, ε186, ε186 + 1 Mn^{2+} and ε186 + 2 Mn^{2+}. The amount of metal-free ε186 also had increased, suggesting that some Mn^{2+} was lost from the first and second metal-binding sites. There were also ions from ε186-Na^+ adducts presumably arising from the introduction of adventitious Na^+ during the dialysis step. These observations suggest that the first two Mn^{2+} ions were bound more tightly to ε186, whereas the third, fourth and fifth Mn^{2+} were more weakly bound to the protein since they were not stable to dialysis. This is consistent the NMR spectroscopic and X-ray crystallographic data that showed two metal ions are bound to the active site of ε186 subunit[177,362] and the observation that non-specific interactions can occur in the ESI source at high analyte concentrations.[13,84,168,414]

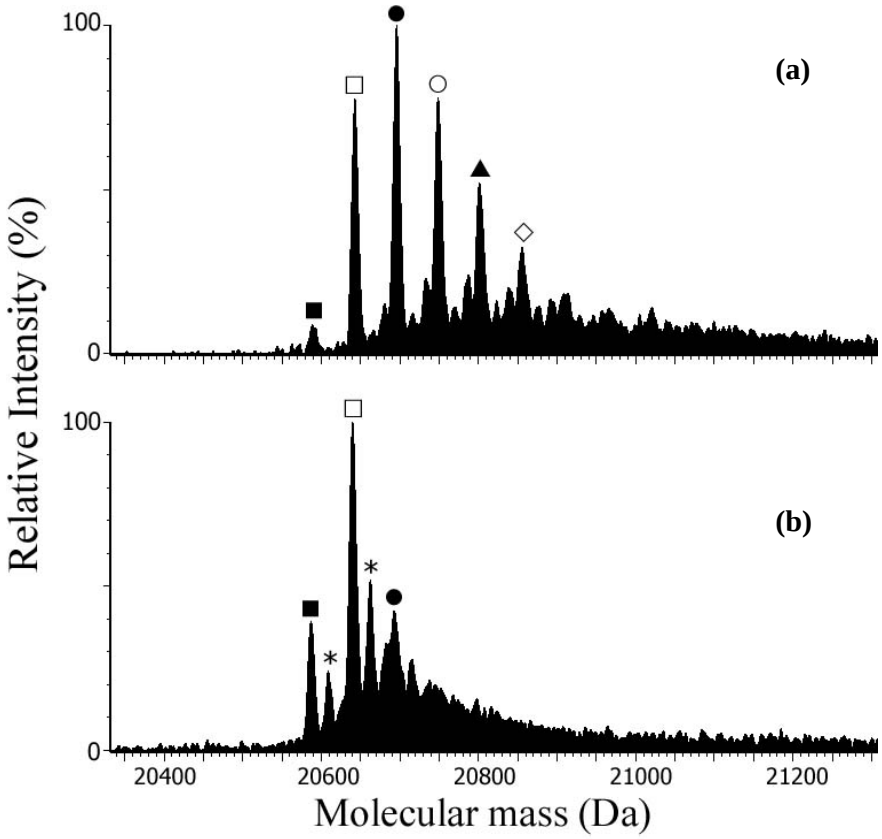


Figure 4.4 Positive ion ESI mass spectra (transformed to a mass scale using MassLynx software™) of a 1:500 mixture of ε186:Mn^{2+}. (a) Before dialysis, and (b) after dialysis in 100 mM NH_4OAc, pH 8.0, for three hours. ■ ε186 alone; □ ε186 with 1 Mn^{2+} bound; ● ε186 with 2 Mn^{2+} bound; ○ ε186 with 3 Mn^{2+} bound; ▲ ε186 with 4 Mn^{2+} bound; ◇ ε186 with 5 Mn^{2+} bound; * complexes containing Na^+.

In order to facilitate data analysis, the relative abundance of each ion observed in a mass spectrum was expressed as a percentage of the total abundance of all ions present. Figure 4.5 shows a plot of relative abundances of ε186 alone and ε186 bound to between one and three Mn^{2+} ions, as a function of Mn^{2+} concentration. The plot in Figure 4.5 (b) highlights the abundances of ions in reaction mixtures containing low concentrations of Mn^{2+} (0-25 μM).

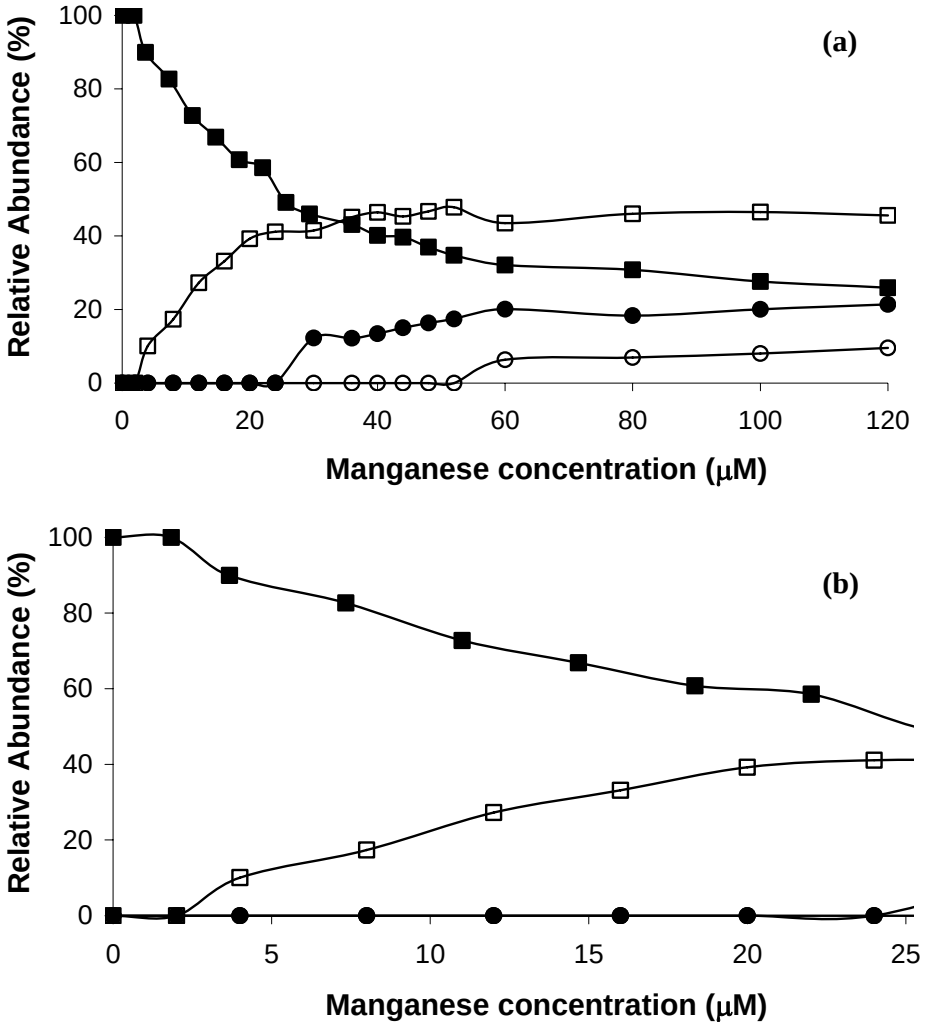


Figure 4.5 Relative abundances of ε186, and complexes of ε186 with different numbers of bound Mn^{2+} ions in ESI mass spectra. The total concentration of ε186 in each reaction mixture was 2 μM. (a) Relative abundances obtained at Mn^{2+} concentrations between 0-120 μM and (b) relative abundances obtained at Mn^{2+} concentrations between 0-25 μM. ■ ε186 alone; □ ε186 with 1 Mn^{2+} bound; ● ε186 with 2 Mn^{2+} bound; ○ ε186 with 3 Mn^{2+} bound.

To obtain the dissociation constant (K_d) for the binding of the first Mn^{2+} to ε186, the range of the Mn^{2+} concentrations where only one metal was bound to the protein was examined. The K_d for the first Mn^{2+} ion bound to ε186 was therefore obtained from spectra of reaction mixtures where the Mn^{2+} concentration was 4, 8, 12, 16, 20 and 24 μM. For each of these reaction mixtures it was possible to determine the

equilibrium amounts of ε186, ε186 + 1 Mn^{2+} and Mn^{2+}. The dissociation constants determined at each of the above six Mn^{2+} concentrations were 4.0 x 10^{-5}, 4.0 x 10^{-5}, 4.5 x 10^{-5}, 4.1 x 10^{-5}, 4.3 x 10^{-5} and 3.6 x 10^{-5} M, respectively. A detailed calculation is shown in Appendix 4. These values are in good agreement considering the error associated with the experiment. The experiment was performed three times, allowing the average K_d for the first Mn^{2+} ion (Mn_A) bound to ε186 to be obtained from a total of eighteen data points. This value was determined as 3.9 x 10^{-5} (± 0.2 x 10^{-5}) M. The K_d values determined by other techniques (e.g. fluorescence and CD spectroscopies) for other Mn-containing enzymes such as phosphoenolpyruvate carboxykinase,[415] pigeon liver malic enzyme[416] and the manganese containing water-splitting-enzyme in photosystem II[417] are in the μM range. This suggests that the value measured here by ESI-MS is reasonable. Calculation of the K_d for the second metal binding site is more complicated as there is no evidence to determine whether the second metal ion binds independently to the protein. Independent binding means that the presence of the first metal ion neither enhances nor inhibits the binding of the second metal ion.[418]

Calculation of a dissociation constant from the ESI-MS data assumes that the mass spectrometer is faithfully sampling the relative amounts of ε186 and ε186 + 1 Mn^{2+} present in solution. In other studies, K_d values/binding affinities measured by ESI-MS have been shown to be in good agreement with data from other solution techniques.[409,419-421] For example, ESI-MS has been used to examine non-covalent interactions of the Src SH2 protein and a range of ligands with different affinities towards the protein.[422] The dissociation constants of these compounds determined by

ESI-MS were consistent with those obtained using an equilibrium fluorescence polarisation assay.

Although these studies showed agreement between K_d values obtained by ESI-MS and other methods, differences should sometimes be expected since there is usually a requirement for different buffers to be used. In ESI-MS, volatile salts (e.g. NH_4OAc, NH_4HCO_3) must be used whereas most other techniques employ phosphate buffers. There have also been examples when, although K_d values measured in solution and by ESI-MS differ, the relative binding affinity for a series of complexes were the same.[71]

Zinc was the second metal ion whose interaction with ε186 was investigated. Zinc is present in other phosphoryl transfer enzymes such as the purple phosphatase from *Phaseolus vulgaris* (red kidney bean), and the *Saccharomyces cerevisiae* (yeast) RNA triphosphatase.[386] Figure 4.6 shows a plot of relative abundances of ions observed in ESI mass spectra of solutions containing 2 μM ε186 and different concentrations of Zn^{2+}. The plot in Figure 4.6 (b) highlights the abundances of ions in reaction mixtures containing low concentrations of Zn^{2+} (0-16 μM). Ions assigned to complexes of ε186 with up to five bound Zn^{2+} were observed in some reaction mixtures. Although the spectra do not prove that Zn^{2+} is binding to the same sites as Mn^{2+}, this would not be surprising since they have similar ionic radii, 74 pm and 83 pm, respectively.[423]

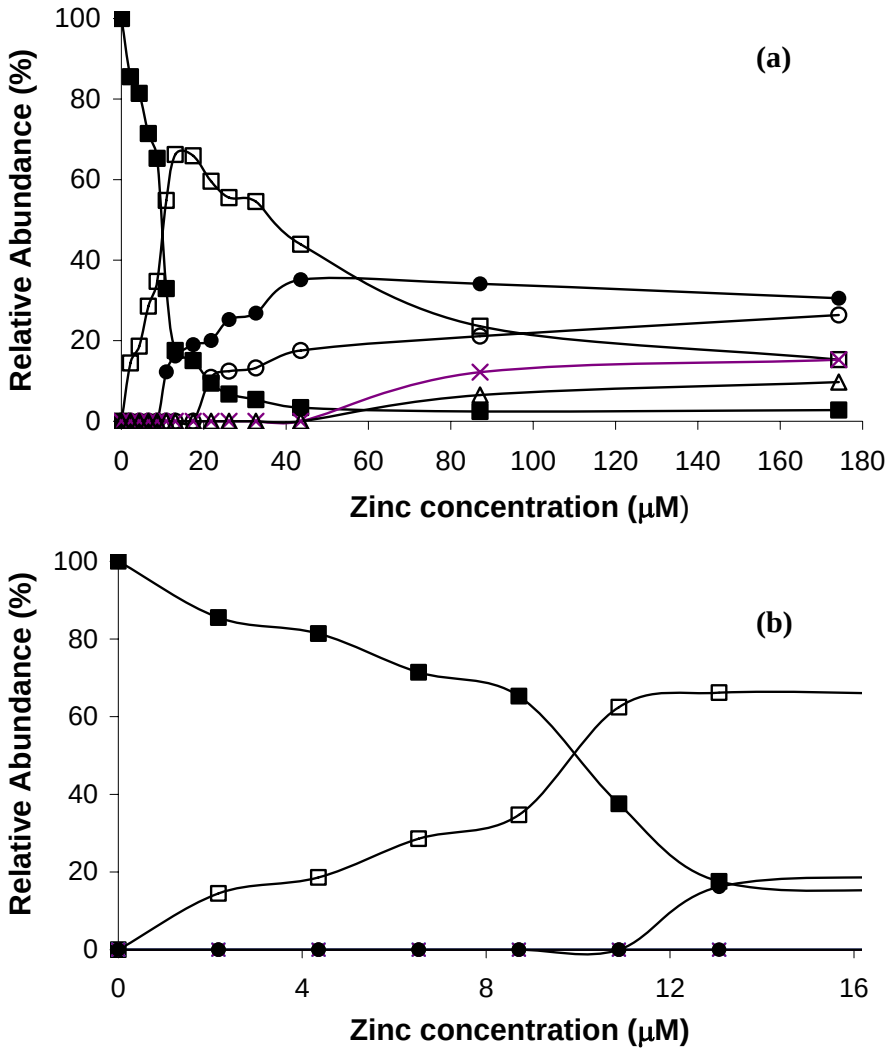


Figure 4.6 Relative abundances of ε186, and complexes of ε186 with different numbers of bound Zn^{2+} ions in ESI mass spectra. The total concentration of ε186 in each reaction mixture was 2 μM. (a) Relative abundances obtained at all Zn^{2+} concentrations examined (0-175 μM) and (b) relative abundances obtained at Zn^{2+} concentrations between 0-16 μM range. ■ ε186 alone; □ ε186 with 1 Zn^{2+} bound; ● ε186 with 2 Zn^{2+} bound; ○ ε186 with 3 Zn^{2+} bound; × ε186 with 4 Zn^{2+} bound; △ ε186 + 5 Zn^{2+} bound.

The amount of free ε186 decreased more rapidly as the concentrations of Zn^{2+} was increased, compared to when Mn^{2+} was present in the reaction mixtures. Complexes of ε186 with one and two Zn^{2+} ions were observed when the metal concentration was ≥12 μM, whereas the complex of ε186 with two Mn^{2+} ions was not apparent until the

metal concentration reached 25 μM (Figure 4.5). Furthermore, the complex of ε186 with three Zn^{2+} ions was of significant abundance at high concentrations of Zn^{2+}. In contrast, the complex with three Mn^{2+} bound to ε186 was of relatively low abundance at all Mn^{2+} concentrations examined. These observations suggest that Zn^{2+} has a greater binding affinity for ε186 than Mn^{2+}. The K_d for binding of the first zinc ion (Zn_A) to ε186 was calculated from the relative amounts of ε186 and ε186 + 1 Zn^{2+} in reaction mixtures where the metal concentration was 2.2, 4.4, 6.5, 8.7 and 11 μM. At these concentrations there was only one Zn^{2+} ion bound to the protein. The experiment was repeated, providing a total of ten individual values of K_d from which an average value of 3.7×10^{-6} ($\pm 0.2 \times 10^{-6}$) M was obtained. Therefore, the K_d for binding of Zn_A to ε186 is ~10 times lower than for Mn_A. This clearly shows that the first zinc ion binds more tightly to the protein than the first manganese ion.

The other metal ion whose binding interaction with ε186 was studied was the lanthanide ion Dy^{3+}. This metal had been used by other workers in NMR studies of the θ-ε186 complex.[396] In the latter study 1.2 equivalents of the metal ion (as $DyCl_3$) were added to the protein to give θ-ε186 with one Dy^{3+} bound. $Dy(OAc)_3$ was used in the ESI-MS titration experiments described here, and added to ε186 using the same experimental procedures used for experiments with Mn^{2+} and Zn^{2+}. Figure 4.7 shows a plot of the relative abundances of ε186, and the complex ε186 + 1 Dy^{3+}, calculated from the relative abundances of ions in ESI mass spectra of solutions with different concentrations of Dy^{3+}.

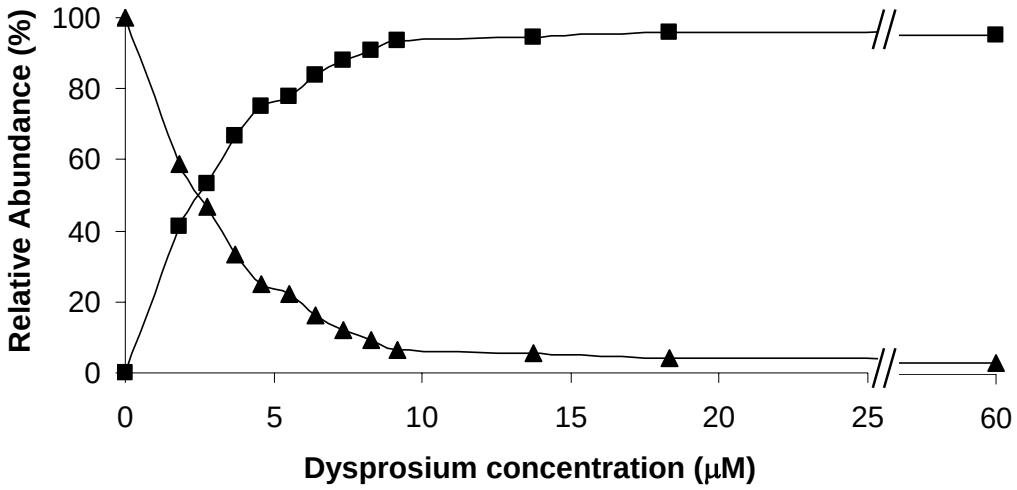


Figure 4.7 Relative abundances of ε186 and ε186 + 1 Dy^{3+} in ESI mass spectra of solutions containing 2 μM ε186 and different concentrations of Dy^{3+}. ▲ ε186 alone; ■ ε186 + 1 Dy^{3+} bound. There were no ions from ε186 + 2 Dy^{3+} in any reaction mixture.

In contrast to the results obtained with Mn^{2+} and Zn^{2+}, at high metal concentrations (e.g. 60 μM) only one Dy^{3+} ion was found to bind to ε186. This observation is in agreement with studies of the binding of lanthanide ions to the exonucleolytic active site of DNA polymerase I.[387,424] Brautigam *et al*. found that only one Eu^{3+} ion bound to KF exo.[424] This is mostly likely because of the larger size of the Dy^{3+} ion (ionic radius 103 pm), compared to Mn^{2+} and Zn^{2+} (83 and 74 pm, respectively).[423] The plots of relative abundance versus metal concentration (Figures 4.5-4.7) show that the first metal binding site was fully occupied (saturated) at metal concentrations of approximately 20, 15 and 10 μM for Mn^{2+}, Zn^{2+} and Dy^{3+}, respectively. An average K_d for binding of the first Dy^{3+} ion to ε186 was obtained using eight data points from each of three separate experiments, and found to be 2.0×10^{-6} ($\pm 0.9 \times 10^{-6}$) M. This value is significantly smaller than that for Mn^{2+} (38.5×10^{-6} M) and slightly lower than that for Zn^{2+} (3.7×10^{-6} M). The order of binding affinities at the first binding site for the three metal ions was therefore $Dy^{3+} \sim Zn^{2+} > Mn^{2+}$. Since Zn^{2+} binds more

tightly to ε186 than Mn^{2+}, it raises the question that Zn^{2+} might be the native metal ion in the protein. However, the intracellular concentration of free Mn^{2+} in bacteria is higher than the concentration of free Zn^{2+}, leaving this an open question.[425] The preferred coordination mode of Mn^{2+} ($3d^5$) and Zn^{2+} ($3d^{10}$) are different. This is evident by comparison of the ligands that bind to these two metal ions in different proteins such as the zinc transporter protein, ZnuA, from *E. coli*, and the manganese transporter protein, PsaA, from *Streptococcus pneumoniae*.[425] These proteins are similar in overall structure,[425] but differ significantly in the metal ligation environments. In ZnuA, where three His residues and one water molecule comprise the metal binding site, Zn^{2+} coordination is favoured. In PsaA, an Asp and a Glu residue in combination with two His ligands results in a slight preference for Mn^{2+}.[425]

In previous work,[173] Mn^{2+} was shown to be more effective than Mg^{2+} in supporting hydrolysis by ε186 of the phosphate ester bond of *p*NP-TMP. Since Zn^{2+} has been shown here to bind more tightly to ε186, it was of interest to determine whether this correlates with a higher nuclease activity for Zn^{2+}-containing ε186 compared to Mn^{2+}-containing ε186. The following experiments were therefore carried out to compare the activities of these two metalloenzyme complexes.

4.7.2 Spectrophotometric assay of ε186 activity

The natural substrate of ε is the phosphodiester bond of a mismatched nucleotide on single-stranded DNA. Therefore in order to determine the effect of the different metal ions on the activity of ε186, an assay was used in which the release of

p-nitrophenolate ion from the 5′-*p*-nitrophenyl ester of thymidine-5′-monophosphate (*p*NP-TMP) was monitored. The method was based on a continuous spectrophotometric assay developed by Hamdan *et al*..[173] In this assay, the release of *p*-nitrophenolate ion by ε186 was monitored spectrophotometrically at pH 8.0 in the presence of 1 mM metal ion, 1 mM DTT and 3 mM *p*NP-TMP. In initial experiments where Zn^{2+} or Dy^{3+} ions were added to the assay mixture under these conditions, precipitation occurred. In order to overcome this problem, 1 mM DTT was subsequently excluded from the assay mixture and the metal concentration reduced to 0.5 mM. For each of the three metal ions this concentration is greater (Mn^{2+} 13 fold, Zn^{2+} 135 fold and Dy^{3+} 250 fold, respectively) than the values of the disociaton constants measured by ESI-MS for binding of the first metal ion to ε186. The concentration of ε186 in the ESI-MS experiments was 2 μM. In the enzymatic activity assays, the concentration of ε186 was 0.1 μM. Under these conditions, it was expected that the metal ions would be bound at both binding sites (for Mn^{2+} and Zn^{2+}) or at the single binding site for Dy^{3+}.

Figure 4.8 (a) shows the change in A_{420} with time for the hydrolysis of 10 mM *p*NP-TMP (pH 8.0 and 25 °C) in the presence of 0.5 mM Mn^{2+}, Zn^{2+} or Dy^{3+}. In Figure 4.8 (b), the scale on the y-axis has been changed to show the increase in A_{420} when Zn^{2+} and Dy^{3+} were present in reaction mixtures.

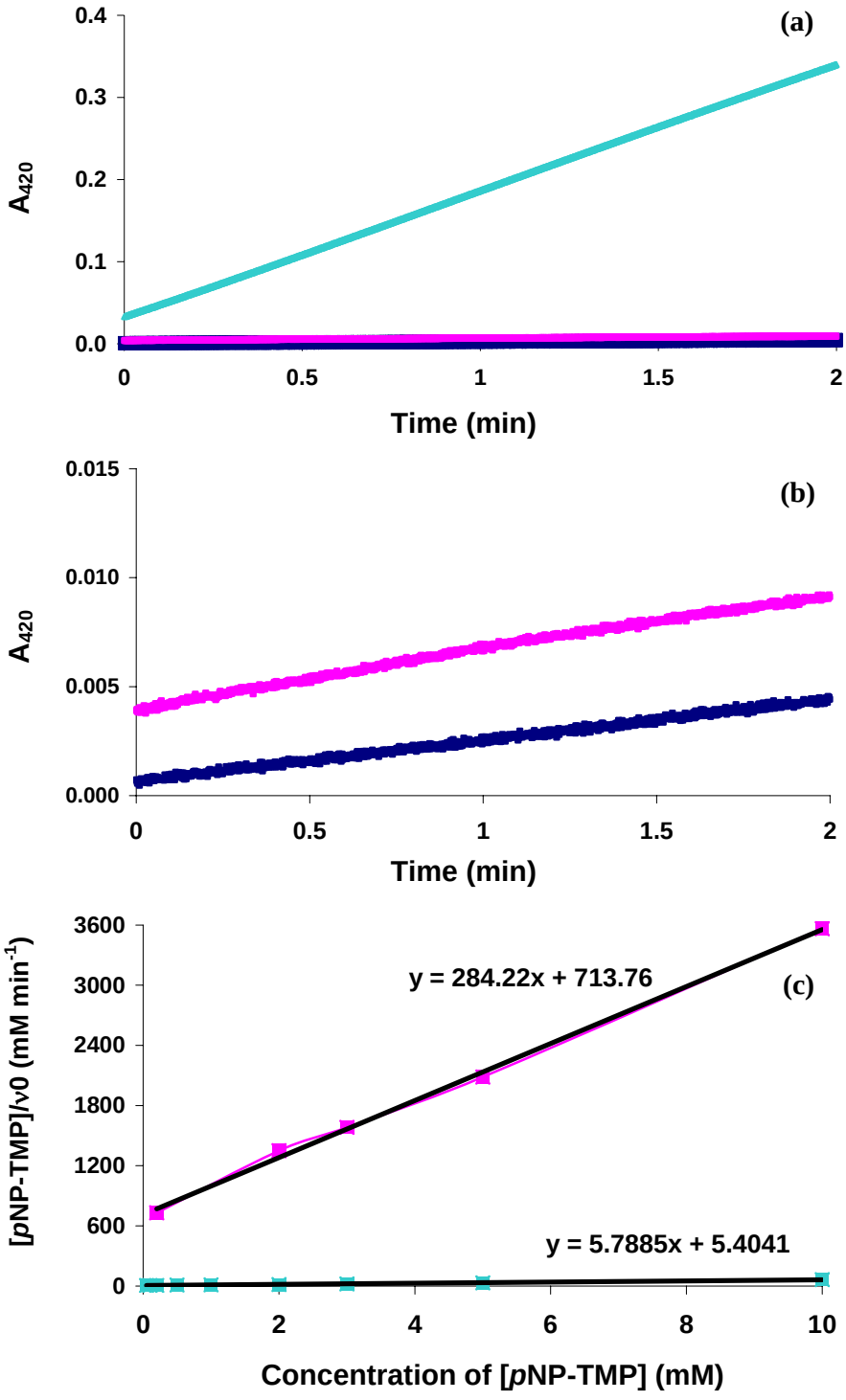


Figure 4.8 Hydrolysis of *p*NP-TMP by ε186, 25 °C, pH 8.0 in the presence of different metal ions. (a) Increase in A_{420} when [*p*NP-TMP] = 10 mM; [ε186] = 0.1 μM and [metal] = 0.5 mM, (b) expanded region of the plot shown in Figure 4.8 (a), showing the differences in variation in A_{420} with time when the metal was Zn^{2+} and Dy^{3+}, (c) Hanes-Woolf plots for the hydrolysis of *p*NP-TMP by ε186 in the presence of Mn^{2+} and Zn^{2+} when [metal] = 0.5 mM. Mn^{2+}; Zn^{2+}; Dy^{3+}.

The concentration of *p*NP-TMP used in the assay was 9 fold higher than the Michaelis constant (K_M)[426] of 1.1 mM determined by Hamdan *et al*. using Mn^{2+} as the metal in the reaction mixture.[173] Therefore, the substrate was not a limiting factor under these assay conditions. The rate of change observed at A_{420} was used to determine the initial velocity (v_0) for the hydrolysis of *p*NP-TMP. The rates were linear for the first few minutes and then decreased significantly as a result of inhibition of the enzyme by the product of the reaction, TMP, as previously observed by Hamdan *et al*..[173] Under these experimental conditions, the highest rate of hydrolysis was observed using Mn^{2+} as the activating metal (rate = 1.1 x 10^{-2} μmoles of *p*-nitrophenolate released per min; μmole min^{-1}). When the metal ion present was Zn^{2+}, the rate was ~75 fold slower (1.5 x 10^{-4} μmole min^{-1}), while the presence of Dy^{3+} in the reaction mixture resulted in a reaction rate that was indistinguishable from the spontaneous rate when no metal ions were added. The rates of reaction were used to calculate specific activities for the enzyme in the presence of difference metal ions. Table 4.1 presents these values, the K_d values for the binding of the first metal ion determined by ESI-MS, and some catalytic parameters (discussed below).

The effect of varying the substrate (*p*NP-TMP) concentration on the rates of ε186-catalysed hydrolysis in the presence of 0.5 mM Mn^{2+} and Zn^{2+} was investigated. These data were used to construct Hanes-Woolf plots (Figure 4.8 (c)) to determine k_{cat},[b] and the Michaelis constant, K_M, for the enzyme catalysed reactions.[427] The linearity of the Hanes-Woolf plots indicated that the hydrolysis of *p*NP-TMP followed Michaelis-Menten kinetics. Values of the slope and y-intercept from each

[b] Turn over number (per minute) is comparable to a first order rate constant when the enzyme active sites are saturated with substrate.

plot were used to determine k_{cat} and K_M when the different metal ions were present at the active site of ε186. For ε186 in the presence of 0.5 mM manganese and zinc (without DTT), k_{cat} was found to be 1730 ± 10 and 35 ± 0.1 min^{-1}, respectively, while K_M was 0.93 ± 0.02 and 2.5 ± 0.2 mM, respectively (Table 4.1). The K_M value when Mn^{2+} was the metal (0.93 mM) is in good agreement with the value determined by Hamdan *et al*. (1.1 mM).[173]

Table 4.1 Kinetics and equilibrium parameters for ε186 treated with Mn^{2+}, Zn^{2+} or Dy^{3+}.

Metal	**K_d (M)**[a]	**Specific ε186 activity (Umg^{-1})**[b]	**k_{cat} (min^{-1})**	**K_M (mM)**	**k_{cat}/K_M (min^{-1} mM^{-1})**
Mn^{2+}	38.5 x 10^{-6} (± 2 x 10^{-6})	5.3 ± 0.2,[c] 0.67 ± 0.03[d]	1730 ± 10	0.93 ± 0.02	1860
Zn^{2+}	3.7 x 10^{-6} (± 0.2 x 10^{-6})	0.070 ± 0.001,[c] 0.058 ± 0.004[d]	35 ± 0.1	2.5 ± 0.2	14
Dy^{3+}	2.04 x 10^{-6} (± 0.9 x 10^{-6})	Negligible	Negligible	Not determined	Not determined

[a] Measured by ESI-MS; first metal binding site only.
[b] [*p*NP-TMP] = 3 mM; [ε186] = 0.1 μM.
[c] [metal] = 0.5 mM; [d] [metal] = 0.05 mM.

These results indicate that the hydrolysis of *p*NP-TMP by ε186 is more efficient in the presence of Mn^{2+} than Zn^{2+}. This suggests that Mn^{2+} may be the native metal required at the active site of the ε subunit. However, more experiments are required to determine which of these (or other) metal ions are most effective for exonuclease activity against the natural DNA substrate.

The effectiveness of the metal ions in promoting ε186-catalysed hydrolysis activity was found to follow the order $Mn^{2+} > > Zn^{2+} > Dy^{3+}$. There was in fact no measurable activity in the presence of Dy^{3+}, which is consistent with the results of a study involving the binding of another lanthanide ion, Eu^{3+}, to KF exo of Pol I.[424] Binding of one lanthanide ion to Pol I or to several other dinuclear phosphoryl-transfer enzymes was shown to cause conformational changes that prevented a second Eu^{3+} from binding and inactivated the enzymes.[424] The low activity of the ε186-Dy enzyme noted here may also be attributed to the large ionic radius of the Dy^{3+} ion (103 pm), which prevented the binding of a second Dy^{3+} ion and most likely alters the geometry of the active site, preventing the substrate from binding.

The effect of varying the concentration of metal ion on the specific activities of the enzyme was also investigated. When the manganese concentration was reduced from 0.5 to 0.05 mM the average specific activity decreased from 5.30 ± 0.2 to 0.67 ± 0.03 μmole min^{-1} mg^{-1} (Umg^{-1}; detailed calculations are shown in Appendix 5). The lower activity observed at the lower concentration reflects that there were fewer ε186 molecules with bound Mn^{2+} ions. The average specific activities of ε186 obtained in the presence of 0.5 and 0.05 mM zinc were 0.070 ± 0.003 and 0.058 ± 0.004 Umg^{-1}, respectively. The comparatively small change in activity caused by a 10-fold reduction in metal concentration here is in agreement with the lower K_d value measured for Zn^{2+} by ESI-MS. In other words, because zinc ion binds more tightly to the enzyme than Mn^{2+}, when the Zn^{2+} concentration was lowered there was still a substantial percentage of ε186 molecules with bound Zn^{2+} ions that were able to catalyse the reaction. The enzymatic activity was, however, much lower than that when Mn^{2+} was bound to the enzyme.

4.8 Conclusions

ESI-MS proved to be a rapid and useful tool for examining interactions between metal ions (Mn^{2+}, Zn^{2+} and Dy^{3+}) and a protein (the N-terminal domain of the ε subunit of DNA polymerase III, ε186). ESI mass spectra showed that up to five Mn^{2+} or Zn^{2+} ions can bind to ε186, with the first two metal ions binding relatively tightly at the enzyme active site, and the others most likely involved in non-specific binding interactions. In contrast, only one Dy^{3+} ion was able to bind to ε186. ESI-MS also provided information on the relative binding affinity of each metal ion towards the protein, through determination of dissociation constants (K_d) for binding of the first metal ion to the protein. A comparison of the dissociation constants showed that Dy^{3+} has the highest relative binding affinity towards ε186, followed closely by Zn^{2+} and Mn^{2+}. Although Mn^{2+} binds the least tightly to ε186, it produced the highest enzyme activity. This supports the suggestion that Mn^{2+} is likely to be the native metal ion, especially considering that the concentration of free Mn^{2+} is higher than that of Zn^{2+} in the bacterial cell.[425]

Chapter 5
Oligomeric Forms of *Escherichia coli* Replicative Helicase DnaB and Complexes with Its Loading Partner DnaC

5.1 Helicases

In the processes of DNA replication, DNA repair, recombination and conjugation, double stranded (ds) DNA must first be unwound in order to provide a metabolically active single-stranded (ss) DNA intermediate. Strand separation of DNA duplexes is the key step in many cellular events. These reactions are catalysed by a class of enzymes called helicases. Helicases are ubiquitous enzymes with fundamental roles in nucleic acid metabolism.[428,429] Their activity leads to disruption of hydrogen bonds between the two strands of duplex DNA. These enzymes are motor proteins which use the free energy of nucleotide triphosphate hydrolysis to translocate along the nucleic acid molecule as it unwinds.[430,431] The unwinding DNA occurs with a specific polarity with respect to the strand on which the helicase is binding.[432] Interest in the study of helicases is increasing, as more mutated helicase genes have been shown to be presented in some serious inherited human diseases.[433]

5.1.1 DnaB helicase

There are at least 12 different enzymes in *E. coli* which exhibit helicase activity. Of these enzymes DnaB is the most important helicase.[434] DnaB is also the most

extensively studied; its primary function involves unwinding duplex DNA in the 5′→3′ direction in front of the replication fork during chromosomal DNA synthesis.[180,331,435] DnaB is also involved in both the initiation and elongation steps of DNA replication, and plays an important role in the replication of bacterial, phage and plasmid DNA.[331,436] Initiation of chromosomal replication requires correct delivery of DnaB to a DnaA-containing nucleoprotein complex at the origin of replication, *ori*C.[437] In order to achieve this, DnaB must interact with its loading partner, DnaC. DnaB also interacts with several other proteins of the replisome, including the tau (τ) subunit of the DNA polymerase III holoenzyme[438] and the replication terminator protein (Tus),[439] as well as nucleotide cofactors.[440-442] It is thought that functional DnaB is a hexamer made up of identical 52 kDa subunits ($(DnaB)_6$).[440,443] Proteolytic studies suggested that each monomer contains two domains, a 12 kDa N-terminal domain and a 33 kDa C-terminal domain, which are connected by a flexible hinge region of ~40 amino acid residues.[444,445] The N-terminal domain is critical for the DnaB helicase to be active and it has been shown to interact with the primase, DnaG.[446,447] The C-terminal domain contains binding sites for DNA, DnaC and nucleotides.[175,444,448] Both domains are necessary for helicase activity.[444,448]

At present, X-ray crystal structures of hexameric helicases are limited to the replicative hexameric replicase-primase of the gene 4 helicase of bacteriophage T7[449,450] and the N-terminal domain of monomeric DnaB.[451,452] There is no crystal structure of full-length DnaB as either a monomer or a hexamer. Sedimentation equilibrium and velocity analytical ultracentrifugation experiments showed that in the presence of magnesium ions (Mg^{2+}) DnaB exists as a stable hexamer $(DnaB)_6$

over a wide range of protein concentrations of 10^{-7}- 10^{-5} M (hexamer).[443] This study showed the important role of magnesium ions in stabilising the hexameric form of DnaB helicase. In the absence of Mg^{2+}, the DnaB protein formed a trimer which dissociated to a monomer at lower protein concentrations.[443] Studies using fluorescent nucleotide analogues showed that the DnaB hexamer binds to six nucleotide molecules, and displayed a preference for purine nucleotides.[441,453,454]

Electron microscopy (EM) studies of $(DnaB)_6$ suggested that it is a symmetric ring structure with a central channel with a diameter of about 3-4 nm.[455,456] In this study, the channel appeared to be completely open at both ends of the molecule with a length of 5.7 nm,[455] which is sufficient to accommodate 20 mer ssDNA (see Figure 5.1).

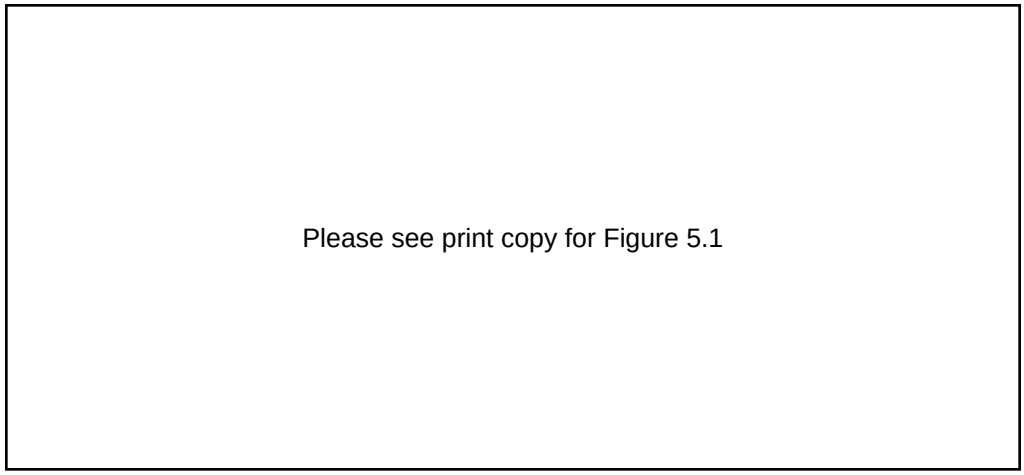


Figure 5.1 Model of the three dimensional structure of DnaB hexamer constructed from cryoelectron micrographs.[455]

Depending on the experimental conditions, the DnaB hexamer was observed in a sixfold (C_6) and/or threefold (C_3) symmetry state.[455-458] Donate *et al.* found that the quaternary state of the DnaB hexamer was dictated by the solution pH level, and not by the type of nucleotide cofactor present.[458] As shown in Figure 5.2, at pH ≥ 7.6, the

hexamer was exclusively observed in the C_3 symmetry state, whereas between pH 6.5-7.2 nearly equal amounts of hexamer with C_3 and C_6 symmetries were detected.[458] In addition these arrangements were also found to be fully reversible upon changing the pH of the solution. Similar quaternary polymorphism is observed in the *Bacillus subtilis* bacteriophage SPP1 gene 40 product (G40P) helicase,[459] the papilloma E1 helicase,[460] the *Thermolyticus aquaticus* RecA protein[461] and the RepA protein encoded by plasmid RSF1010.[462] This interconversion between different conformations may be of functional significance for the unwinding of DNA.[456]

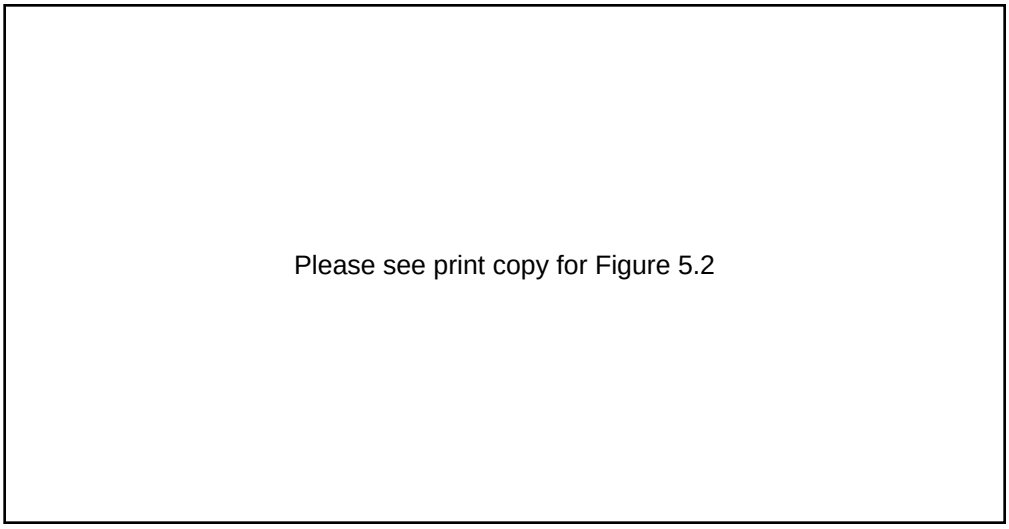


Figure 5.2 Electron micrographs after self-organising map algorithm analysis showing different quaternary structures of the DnaB helicase from pH values of 6.5, 7.2, 7.6 and 8.1. At pH ≥ 7.6 the DnaB hexamer exists only in the C_3 symmetry state, whereas at pH values between 6.5 and 7.2, the hexamer were present in C3 and C6 symmetry forms. Adapted from Donate *et al*..[458]

5.1.2 DnaC protein

DnaC is a 28 kDa monomeric protein which is essential for *E. coli* DNA replication.[331,463,464] DnaC is not necessary for the binding of the DnaB helicase to ssDNA,[465-467] but it is necessary to form the specific protein-protein complex,

$(DnaB)_6(DnaC)_6$ at o*ri*C.[331,463,464,468,469] Upon delivery of DnaB, ATP is hydrolysed and DnaC is released from the $(DnaB)_6(DnaC)_6$.[178,468] Figure 5.3 shows models developed from electron microscopic images of the $(DnaB)_6(DnaC)_6$ complex.[455,470] Each of the DnaC dimers is dumb-bell shaped, with two regions of different size (as shown in pink in Figure 5.3). It is believed that DnaC dimerises as it interacts with the hexameric DnaB and is responsible for locking the DnaB helicase into a C3 symmetry forrm.[470] DnaC dimers show extensive contacts with the DnaB hexamer. Each dimer associates with two different dimeric units of DnaB as indicated by the regions designated I, II and II in Figure 5.3 (c). The smaller lobe of the DnaC dimer is in contact with one subunit of a DnaB dimer (e.g. I), while the larger lobe of the DnaC dimer is in contact with both subunits of a neighbouring DnaB dimer (II and III).[470] In solution on its own DnaC exists as a monomer[471] and it binds one nucleotide molecule in the presence of Mg^{2+}.[472] The intrinsic binding affinities of ATP and ADP for DnaC are similar, and are ~3-4 times greater than those of other nucleotides.[472-474]

Please see print copy for Figure 5.3

Figure 5.3 Models of the $(DnaB)_6(DnaC)_6$ complex developed from electron micrographs.[470] (a) View from the top showing three DnaC dimers (pink) binding to the DnaB hexamer (blue), (b) angle between the DnaC and DnaB dimers and (c) side view showing three different contact areas (I, II and III) between each DnaC dimer and DnaB. Taken from Barcena *et al*..[470]

5.2 ESI-MS of Large Macromolecular Complexes

The analysis of large macromolecular complexes using ESI-MS has become possible as a result of the coupling of nanoESI with quadrupole time-of-flight mass analysers modified to increase transmission of high *m/z* ions. In nanoESI a much narrower orifice is used from which to spray ions than for ESI-MS.[475] The resulting smaller droplet size in nanoESI requires lower voltages to initiate the spray, and milder desolvation conditions for evaporation, favouring preservation of the non-covalent complexes during the ESI-MS process.[476]

In early experiments, analysis of non-covalent complexes by ESI-MS was limited to smaller complexes (M_r ~50-60 kDa),[477] since standard quadrupole mass analysers had a limited *m/z* range (typically *m/z* 2,000-4,000). Larger non-covalent complexes often exhibit ion charge states appearing well above the *m/z* 4,000 range. To overcome this limitation, ESI was coupled with time-of-flight (ToF) or quadrupole-ToF mass analysers.[478] Other modifications have increased the transmission of large ions in ESI-Q-ToF mass spectrometers. By manipulating the pressure gradients within the mass spectrometer by introducing collision gas at various stages during the path of the ions, Robinson and co-workers demonstrated that the intact 14 non-covalently bound subunits of chaperonin GroEL could be maintained in the gas phase.[479] There have been numerous studies that have shown that the transmission of high *m/z* ions is improved by increasing the pressure in the first vacuum stages by reducing the pumping speed or adding a collision gas.[57,479-482] This phenomenon is termed “collisional dampening/cooling”, whereby a bath gas absorbs the excess translational energy of the ions.[23,55,131,481] Using this approach, multimeric protein assemblies with molecular masses of more than 1 MDa have been observed.[479,482-488]

The stabilities of macromolecular complexes can be determined in CID experiments. In a Q-Tof mass spectrometer, an ion is selected in the quadrupole and the products of its dissociation detected in the ToF analyser.[54,55] For large complexes, the ions may have a greater *m/z* than is accessible by the quadrupole analysers of most commercial instruments. Robinson and co-workers used a custom-built quadrupole ToF instrument incorporating collision cooling and a quadrupole with reduced RF frequency that can operate up to *m/z* 32,000 (see Figure 5.4).[57] This allowed complexes in excess of 60 kDa to be dissociated enabling analysis of pathways for assembly and disassembly of subunits.[57,489]

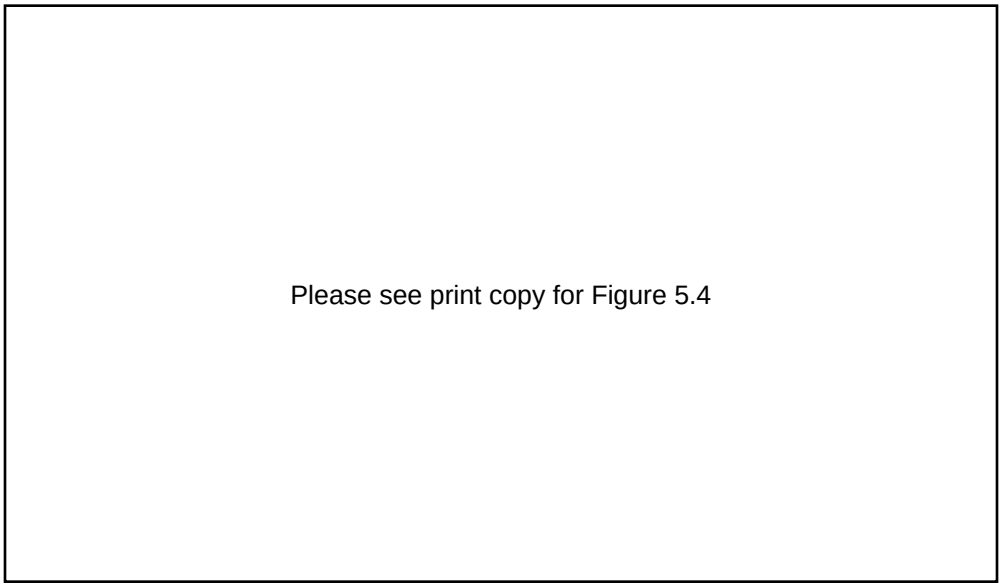


Figure 5.4 A schematic representation of the custom-built Waters Q-ToF Ultima™. Adapted from the Waters website (http://www.waters.com).

ESI-MS has also been used to analyse subunit exchange in oligomeric proteins. Sobott *et al*. performed real-time monitoring of subunit exchange among small heat shock proteins, *Ps*HSP18.1 from pea and *Ta*HSP16.9 from wheat, using nanoESI-MS.[490] The gentle conditions of ESI-MS kept the protein intact in the gas phase,

revealing that both proteins existed as dodecamers.[490] Mixtures of both proteins were studied at different ratios and exchange of subunits between the two proteins was monitored. The results showed that the composition of heterododecamers formed by subunit exchange was governed by the starting ratio of the two components rather than by an inherent preference for certain stoichiometries. The kinetics of the subunit exchange revealed that exchange occurred via sequential incorporation of subunits with dimeric species being the predominant units of exchange. This study illustrated the powerful application of the real-time analysis, where different species and their relative populations during the subunit exchange of multimeric protein complexes could be observed.[490] In the current work, a Q-ToF Ultima ESI mass spectrometer based on the instrument in the Robinson laboratory was used to examine complexes of $(DnaB)_6$ and $(DnaB)_6(DnaC)_6$.

5.3 Scope of This Chapter

There are different oligomeric states of helicases. Some helicases appear to function as monomers,[491] dimers[492] and others as hexameric rings.[430] DnaB is thought to function as a hexamer, which is stabilised by the binding of magnesium cations.[443] The X-ray structure of the isolated N-terminal domain of DnaB forms a symmetrical dimer interface (Figure 5.5), that is stabilised by interactions between Glu88 on one molecule and Lys110 on the other. There is also partial stacking of the two Phe102 residues from each molecule. Amino acid residue Asp82 may also be of some importance in stabilising the structure as it caps an α helix near the dimer interface (Figure 5.5 (b)). Since formation of the dimer might be an important step in assembly of the hexamer, several dimer interface (Phe102) and Asp82 mutants of full-length DnaB were prepared to determine the effect on the correct oligomeric assembly of

DnaB. The mutants examined were F102W, F102E, F102H and D82N. Furthermore, the mutants F102W and D82N were tested for their ability to form the helicase loading complex $(DnaB)_6(DnaC)_6$. For this work, a modified Waters Q-Tof Ultima mass spectrometer (with quadrupole to *m/z* 32,000) was used.

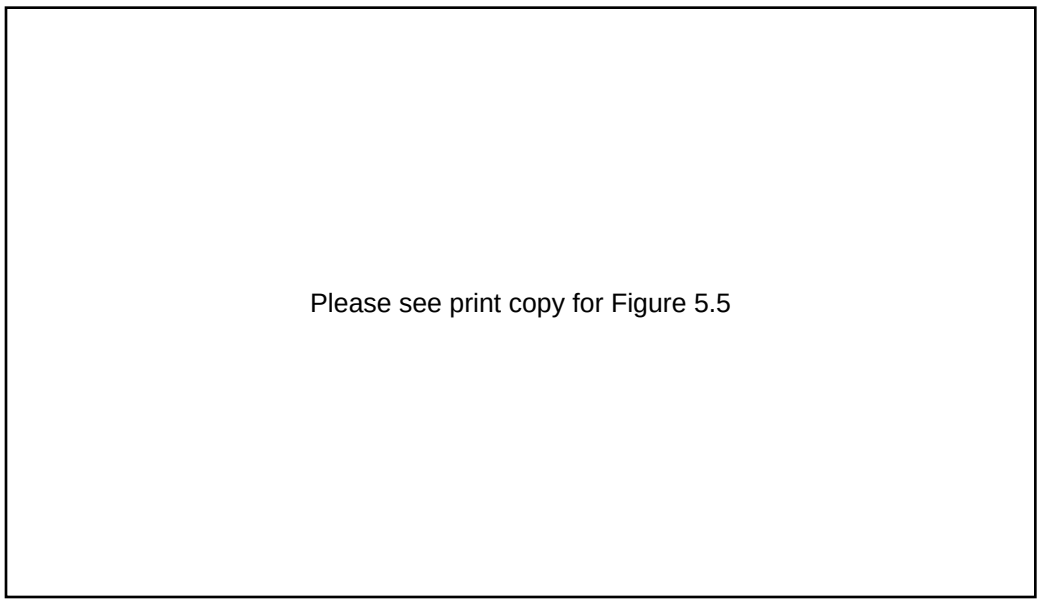


Figure 5.5 X-ray crystal structure of dimeric DnaB-N (Protein data bank 1B79). Stick and ribbon representations of the DnaB-N dimer interface are shown in (a) and (b), respectively. Amino acid residues involved in DnaB-N dimerisation are shown in yellow (Glu88), pink (Lys110) and red (Phe102). The green (D82N) residue caps an alpha helix.[452]

5.4 Results and Discussion

5.4.1 Oligomers of DnaB and DnaB mutants revealed by nanoESI-MS

In preliminary work, nanoESI mass spectra of DnaB were obtained from mixtures containing a range of NH_4OAc concentrations (50-1500 mM; data not shown). At concentrations lower than 750 mM, the peaks were not well resolved because they were broad, making it difficult to interpret the spectra. Spectra with narrower peaks

were obtained from mixtures containing 1 M NH_4OAc. While better resolution was also achieved using 1.5 M NH_4OAc, it was sometimes difficult to obtain a continuous spray. Therefore, subsequent experiments were carried out using solution containing 1 M NH_4OAc.

Since DnaB is a potent ATPase and because the hexamer has previously been shown to be stabilised by magnesium,[493] both 0.1 mM ATP and 1 mM Mg^{2+} were included with the 1 M NH_4OAc for preparation of oligomers. The calculated masses of monomeric DnaB, F102W, F102E, F102H and D82N are 52390, 52,429, 52372, 52380 and 52389, respectively. Figures 5.6 (a)-(d) show positive ion nanoESI mass spectra of full length DnaB and mutants F102W, F102H and D82N in 1 M NH_4OAc, 1 mM $Mg(OAc)_2$, 0.1 mM ATP, pH 7.6. The molecular mass of DnaB (and mutant) monomers was used to calculate the expected mass for oligomers. The most abundant ions in the spectrum of DnaB (Figure 5.6 (a)) are attributable to the hexamer $(DnaB)_6$. The calculated *m/z* for the $[M+35H]^+$ ion of $(DnaB)_6$, indicated by $[35^+]$, is 8960 (M_r 313,554). The peaks are broad most likely as a result of the adduction of magnesium, nucleotides and water to the protein as has been observed for several other large protein complexes such as transthyretin,[477] GroEL tetradecamer,[494] and intact ribosomes.[495] Close examination of the 35^+ ion reveals that it commences around 8960 and ends around 9080. Table 5.1 shows the calculated values of molecular mass and *m/z* for the 35^+ ion of complexes of $(DnaB)_6$ with 1 to 6 molecules of ADP and 4 Mg^{2+} ions bound.

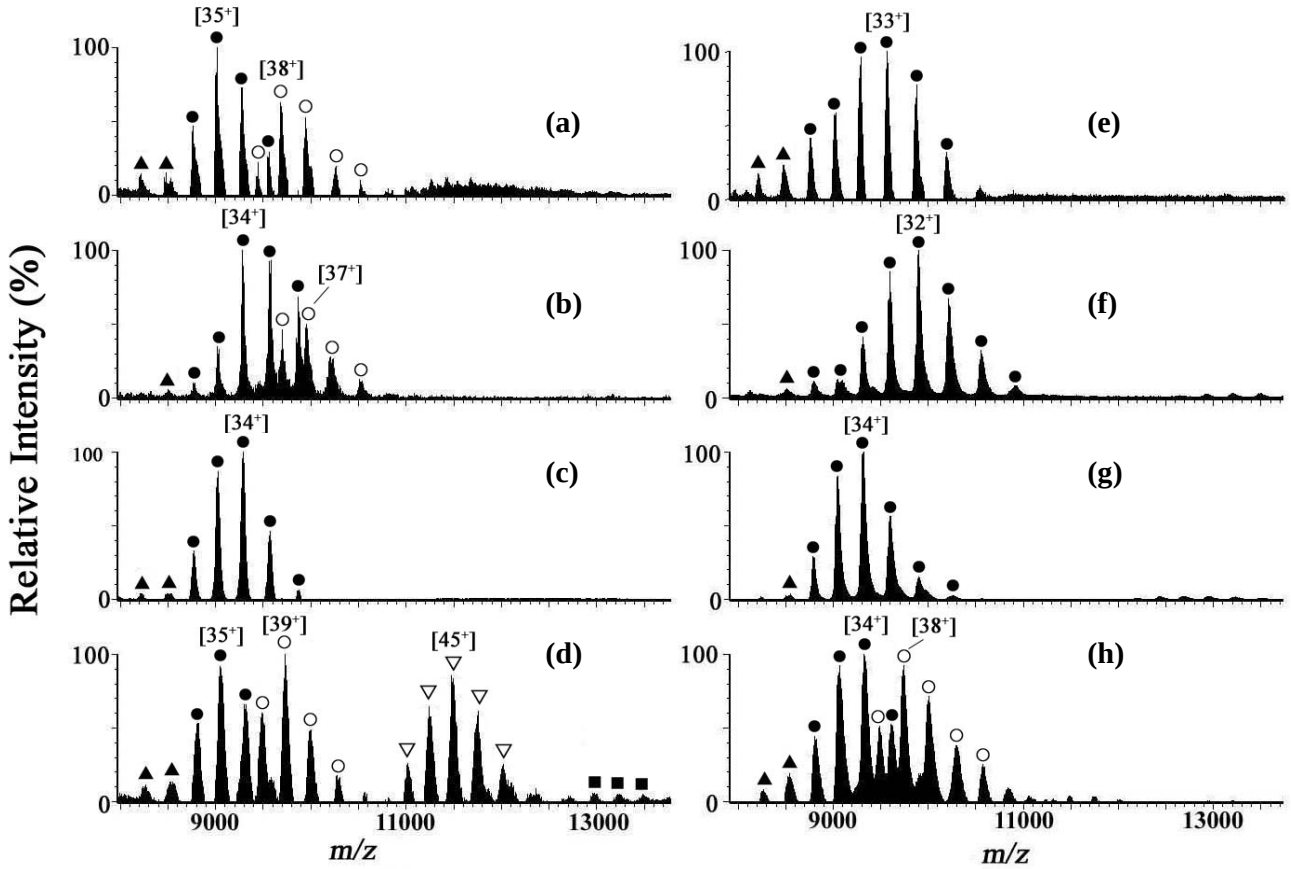


Figure 5.6 Positive ion nanoESI mass spectra of full length DnaB and mutants, F102W, F102H and D82N in 1 M NH_4OAc, 0.1 mM ATP, pH 7.6, with 1.0 mM $Mg(OAc)_2$ ((a) to (d), respectively), or with 0.1 mM $Mg(OAc)_2$ ((e) to (h), respectively). ▲ Pentamer; ● hexamer; ○ heptamer; ▽ decamer; ■ dodecamer. The numbers in square brackets indicate the number of charges for the most abundant ion present for each oligomeric form.

The range of *m/z* most likely results from protein complexes containing different numbers of bound ATP and Mg^{2+}. These results are in good agreement with other studies using techniques such as analytical ultracentrifugation, and fluorescence spectroscopy, where at least four magnesium ions and six molecules of nucleotide have been found to bind to the hexameric DnaB helicase.[443,496]

Table 5.1 Calculated values of *m/z* for the 35^{+} ion of hexameric DnaB ($(DnaB)_6$) and its complexes with ADP and magnesium.

Complex	Calculated mass (Da)	Calculated *m/z*
$(DnaB)_6$	313,554	8960
$(DnaB)_6$ + 1ADP	313,981	8972
$(DnaB)_6$ + 2ADP	314,408	8984
$(DnaB)_6$ + 3ADP	314,836	8996
$(DnaB)_6$ + 4ADP	315,263	9009
$(DnaB)_6$ + 5ADP	315,690	9021
$(DnaB)_6$ + 6ADP	316,118	9033
$(DnaB)_6$ + 1ADP + $4Mg^{2+}$	314,078	8975
$(DnaB)_6$ + 2ADP + $4Mg^{2+}$	314,506	8987
$(DnaB)_6$ + 3ADP + $4Mg^{2+}$	314,933	8999
$(DnaB)_6$ + 4ADP + $4Mg^{2+}$	315,360	9011
$(DnaB)_6$ + 5ADP + $4Mg^{2+}$	315,787	9023
$(DnaB)_6$ + 6ADP + $4Mg^{2+}$	316,249	9036

In addition to the presence of the hexameric DnaB in Figure 5.6 (a), a substantial amount of heptameric DnaB, $(DnaB)_7$, was also present. This is the first observation of $(DnaB)_7$ in aqueous solution. Previously, we observed $(DnaB)_7$ in 200 mM NH_4OAc, 0.1 mM ATP, 1.0 mM Mg^{2+}, pH 7.6, and 30% (*v/v*) methanol.[497]

The question of whether this heptameric form is the result of non-specific interactions in the ionisation source of the mass spectrometer requires consideration. Higher order oligomeric protein complexes have been observed in other ESI-MS studies when protein concentrations were high (e.g. millimolar insulin).[498] In the experiments here, however, the concentration of the hexamer was only 10 μM. In support of the ability of DnaB to form heptamers under conditions other than in the ESI source, heptameric forms have been observed previously for other helicases including the *Thermus thermophilus* RuvB,[499] *Methanobacterium thermoautotrophicum* MCM,[500] human Rad52[501] and bacteriophage T7 primase-helicase.[502] Early electron microscopy (EM) studies[503] and native gel electrophoresis[504] of the latter protein showed that it existed as a hexamer.[503,505,506] Later, the X-ray crystal structure revealed that the protein existed as a heptamer.[502] Furthermore, in the same work, EM revealed a mixture of hexameric (one-third) and heptameric (two-thirds) rings.[502] These observations prompted speculation as to whether the heptamer was a result of the different sample preparation procedures and/or whether the heptameric form had any biological relevance. Based on the crystal structure, it was suggested that the central channel of the heptamer was sufficiently large enough to accommodate double-stranded (ds) DNA, allowing translocation of the heptameric protein along the dsDNA, while the hexamer can only accommodate ssDNA.[502] Electron microscopic and X-ray crystallographic data showed that in the presence of nucleotides, only hexamer was observed when the protein oligomer was bound to ssDNA and a mixture of heptamer and hexamer when not bound to the ssDNA.[507] This suggested that the heptameric ring observed in these helicases may have a role in encircling DNA during loading, with loss of one subunit to form a tight complex in the active form.[500,507]

Previous NMR and X-ray crystallographic studies showed that dimerisation of the N-terminal domain of DnaB, DnaB-N, involves interactions between Glu88 and Lys110 and the partial stacking of Phe102 residues from each DnaB molecule.[451,452] In this work, the effects on formation of the hexamer and other oligomers resulting from altering interactions at the dimer interface were investigated. The role of F102 residues at the dimer interface of DnaB-N was examined by preparing full-length F102 mutants, F102W, F102E and F102H. NanoESI mass spectra of these mutants (Figures 5.6 (b)-(d)) were acquired under the same experimental conditions as those for wild-type DnaB (Figure 5.6 (a)). The nanoESI mass spectrum of F102W (Figure 5.6 (b)) is similar to that of DnaB; with ions consistent with both hexamer and heptamer present. There was, however, a small shift in charge state distribution of the ions observed. The most abundant ion for $(F102W)_6$, was the 34^+ ion, whereas that for $(DnaB)_6$ was the 35^+ ion. This may be indicative of small differences in conformation of the two proteins, which could be further investigated in ion mobility experiments.[508] The time taken for ions produced from nanoESI to travel a defined distance in the ion mobility mass spectrometer are measured and converted to a collision cross section value, which is directly related to the size and overall shape of an ion. The structure of the *trp* RNA binding attenuation protein (TRAP) was examined by determining collision cross sections of various TRAP complexes in the absence and presence of binding ligand.[509] The cross section values derived from ion mobility mass spectrometry data correlated well those determined by X-ray crystallography.[509] On the other hand, the difference in charge state distribution observed in Figure 5.6 (a) and (b) may be due to variations in the size of the orifice diameter of nanospray tips.[510] Li and Cole found that as the size of the orifice was decreased, the charge state distribution of peptides and proteins examined shifted

towards higher values.[510] In the current study, a subtle shift in charge state distribution was observed on several occasions where the nanoESI mass spectra were obtained from the same mixtures but using different tips, although the same oligomeric species were observed.

The only ions observed in the nanoESI mass spectrum of F102H were from the hexamer (Figure 5.6 (c)), with no heptamer detected under any of the experimental conditions examined. The mass spectrum of F102E (data not shown) was very similar to that of wild-type DnaB shown in Figure 5.6 (a). Dimerisation of DnaB-N has been shown in NMR experiments previously to be prevented by the F102E mutation.[511] However, the hexameric form of this protein was observed, suggesting that the F102E mutation does not prevent the formation of hexamer. Gel filtration experiments carried out by our collaborators (Dixon, NE; unpublished) also suggested that F102E, and all of the mutants examined here, associated as a hexamer. Taken together, these observations suggest that dimerisation as observed for the isolated N-terminal domain of DnaB, does not have a role in structural organisation of the hexamer. This is also consistent with the absence of DnaB-N dimer interactions in structural models of domain organisation in the C3 and C6 symmetry states of $(DnaB)_6$ derived from EM data.[512] There are, however, subtle effects on oligomerisation that affect the proteins ability to form higher order oligomers such as heptamers and dodecamers.

The Asp82 residue was also chosen for mutation, since this residue caps an α helix near the DnaB-N dimer interface (Figure 5.5). NanoESI mass spectra of the D82N mutant were significantly different from those of DnaB or the F102 mutants. Figure

5.6 (d) shows the nanoESI mass spectrum of D82N, where ions from hexamer and heptamer as well as that of a decamer $(D82N)_{10}$ were present. There were also some low-abundance ions around *m/z* 13,000 attributable to a dodecamer, which were also present for all the other proteins except for F102H. Dodecamers have been observed for other "hexameric" helicases. Using electron microscopy image analysis and three-dimensional reconstructions, Stasiak *et al*. observed a double hexameric ring (dodecamer) of *E. coli* RuvB helicase on DNA in the presence of ATP.[513] Dodecamers have also been observed for other helicases including the simian virus 40 large tumour antigen[514] and the archaeal *Methanobacterium thermoautotrophicum* MCM.[515,516] In the nanoESI mass spectrum of D82N, the decamer $(D82N)_{10}$ observed may be a result of dissociation of a dodecamer $(D82N)_{12}$, either in solution or in the gas phase. Dissociation of complexes may also account for the presence of small amounts of pentamers in all spectra, and monomers, dimers and trimers (data not shown) in some spectra. If open rings were present in solution as observed in electron micrographs of the MCM complex,[515] it is likely that these structures would be more susceptible to dissociation, even under gentle conditions in the mass spectrometer.

The high abundance of $(D82N)_7$ and $(D82N)_{10}$ suggests that the D82N mutation stabilises these forms. On the contrary, the absence of higher oligomers including heptamer under all experimental conditions examined for F102H suggest that this mutation markedly destabilised forms other than the hexamer. These residues are therefore important in enabling DnaB oligomers to assemble in the correct, functional oligomeric form(s), but the molecular basis for this will require X-ray crystal structures of these mutant oligomers and wild-type DnaB.

The question of whether the heptamers and dodecamers observed for other helicases using EM studies and here for the first time for *E. coli* DnaB helicase using nanoESI-MS are functionally relevant in the process of chromosomal replication remains to be determined. Nevertheless, these observations provide additional information about the physical properties of the helicase and highlight the role of mass spectrometry in the study of oligomeric protein complexes. This information may be useful for understanding the oligomerisation process of proteins in many cellular events, such as replication, transcription and translation, which are regulated by large multiple protein complexes.

5.4.2 Effect of Mg^{2+} concentration on oligomerisation of DnaB and mutants

It has been shown that Mg^{2+} is essential for formation of hexameric DnaB.[443] An analytical ultracentrifugation study revealed dimerisation of a trimer to a hexamer occurred upon addition of Mg^{2+} to 5 mM or an increase in NaCl concentration to approximately 0.9 M.[443] In the current work, oligomerisation of DnaB and the mutants were compared at two different Mg^{2+} concentrations: 1.0 and 0.1 mM. The nanoESI mass spectra of proteins prepared in a solution containing 1 M NH_4OAc, 1 mM $Mg(OAc)_2$, 0.1 mM ATP, pH 7.6 are shown in Figure 5.6 (a)-(d) (discussed above), and those prepared under the same conditions except for the presence of 0.1 mM Mg^{2+} in solution are shown in Figure 5.6 (e)-(h). In all cases, the resolution of the mass spectra of proteins prepared in the solution containing the lower magnesium concentration was improved, most likely because there were fewer non-specific Mg^{2+} adducts present.

The nanoESI mass spectra of DnaB, F102W, and F102H shown in Figure 5.6 (e), (f), and (g), respectively, and that of F102E (not shown) show that the hexameric proteins were the most abundant, with very little, if any, heptamer being present. The mass spectrum of D82N (Figure 5.6 (h)) differs from those of DnaB and the F102 mutants as $(D82N)_7$ was still present at approximately equal abundance to $(D82N)_6$. The distributions of charges present in the spectra of DnaB, F102W and D82N were different, depending upon the concentration of Mg^{2+}. At the higher Mg^{2+} concentration, higher numbers of charges for these proteins were observed (compare Figure 5.6 (a) with (e), (b) with (f) and (d) with (h)). It is not possible to determine from these experiments whether this is the result of the binding of Mg^{2+} ions contributing to a higher number of charges, a change in the protein conformations, the effect of variations of the size of nanospray tips or a combination of these factors. For F102H, the distribution of oligomeric forms and their charge states were unchanged when the Mg^{2+} concentration decreased from 1 to 0.1 mM (compare Figure 5.6 (c) and (g)). Overall, these observations (Figure 5.6) suggest that the hexameric forms are stabilised even at the low Mg^{2+} concentrations (0.1 mM). This is in contrast to investigations of the oligomerisation of DnaB using analytical ultracentrifugation where concentrations of Mg^{2+} greater than 3 mM were required to observe $(DnaB)_6$.[443] In the absence of Mg^{2+}, a trimer was observed which dissociated into monomers at low protein concentrations.[443] In the presence of 5 mM Mg^{2+}, the hexameric DnaB was stable over a wide range of hexamer concentrations of 0.1 to 10 μM.[443] The higher stability of the hexamer observed in the current nanoESI-MS experiments at lower Mg^{2+} concentrations (0.1 and 1 mM) is most likely the result of the different buffer systems used for preparation of the samples. In the nanoESI-MS

experiments (Figure 5.6), the higher order oligomeric forms, the heptamer, and in the case of D82N, the decamer, were stabilised at the higher Mg^{2+} concentration.

5.4.3 Titration of DnaB, F102W and D82N with DnaC

DnaB hexamers form complexes with six molecules of DnaC in the presence of ATP or ADP.[175,517] Previous studies using a DNA replication assay and protein overlay analysis of genetic mutations of DnaC,[518] and cryo-EM data,[470] showed that in the $(DnaB)_6(DnaC)_6$ complex, the N-terminal domain of DnaC interacted with the C-terminal domain of DnaB in $(DnaB)_6$, fixing the complex in C3 symmetry. In our laboratory, we previously developed conditions for the detection of the $(DnaB)_6(DnaC)_6$ complex by nanoESI-MS (Figure 5.7 (a)).[519] In this work, complexes of $(F102W)_6$ or $(D82N)_6$ with DnaC were examined using nanoESI-MS.

Preliminary studies showed that DnaB/DnaC mixtures were prone to precipitation when the concentration of ammonium acetate was greater than 500 mM. Furthermore, peaks in the nanoESI mass spectra of $(DnaB)_6(DnaC)_6$ were broad when 1 mM Mg^{2+} was present. Therefore, the complexes were prepared in 300 mM NH_4OAc, 0.1 mM $Mg(OAc)_2$, 0.1 mM ATP, pH 7.6.

Figure 5.7 shows a series of nanoESI mass spectra of $(DnaB)_6$ (Figure 5.7 (a)), $(F102W)_6$ (Figure 5.7 (b)) or $(D82N)_6$ (Figure 5.7 (c)) treated with three, six and eight molar equivalents of DnaC. The spectrum at the top of each panel shows the spectrum of $(DnaB)_6$, $(F102W)_6$, or $(D82N)_6$ with no DnaC added. In each case, the highest oligomeric forms (heptamer for $(DnaB)_6$ and $(F102W)_6$, decamer for $(D82N)_6$) were not present since the Mg^{2+} concentration was low (cf. Figure 5.6 (a),

(b) and (d) with Figure 5.7 (a), (b) and (c)). In each of the titrations of $(DnaB)_6$, $(F102W)_6$ or $(D82N)_6$ with DnaC, there were no complexes of these hexamers with DnaC of significant abundance until six molar equivalents of DnaC were added. At this ratio, a mixture of hexamers of $(DnaB)_6$, $(F102W)_6$ or $(D82N)_6$ with different numbers of DnaC molecules bound were observed. In the mass spectrum of $(DnaB)_6$ (Figure 5.7 (a)), the most abundant ion was from the complex $(DnaB)_6(DnaC)_5$, and a substantial amount of $(DnaB)_6(DnaC)_6$ complex was also present. As the amount of DnaC was increased to eight molar equivalents, the abundance of $(DnaB)_6(DnaC)_5$ reduced significantly and the most abundant ions were from the $(DnaB)_6(DnaC)_6$ complex. The mass spectra of $(F102W)_6$ with six and eight equivalents of DnaC added (Figure 5.7 (b)) were similar to those for $(DnaB)_6$ except that a $(DnaB)_6(DnaC)_4$ complex was present when the $(F102W)_6$:DnaC ratio was 1:6.

In the case of $(D82N)_6$ (Figure 5.7 (c)) the addition of three molar equivalents of DnaC resulted in a decrease in the abundance of the heptameric form. Furthermore, no complexes of DnaC with $(D82N)_7$ were observed in any of the mass spectra. These results suggest that the presence of DnaC in the mixture shifted an equilibrium between the hexamer and heptamer towards the hexamer.

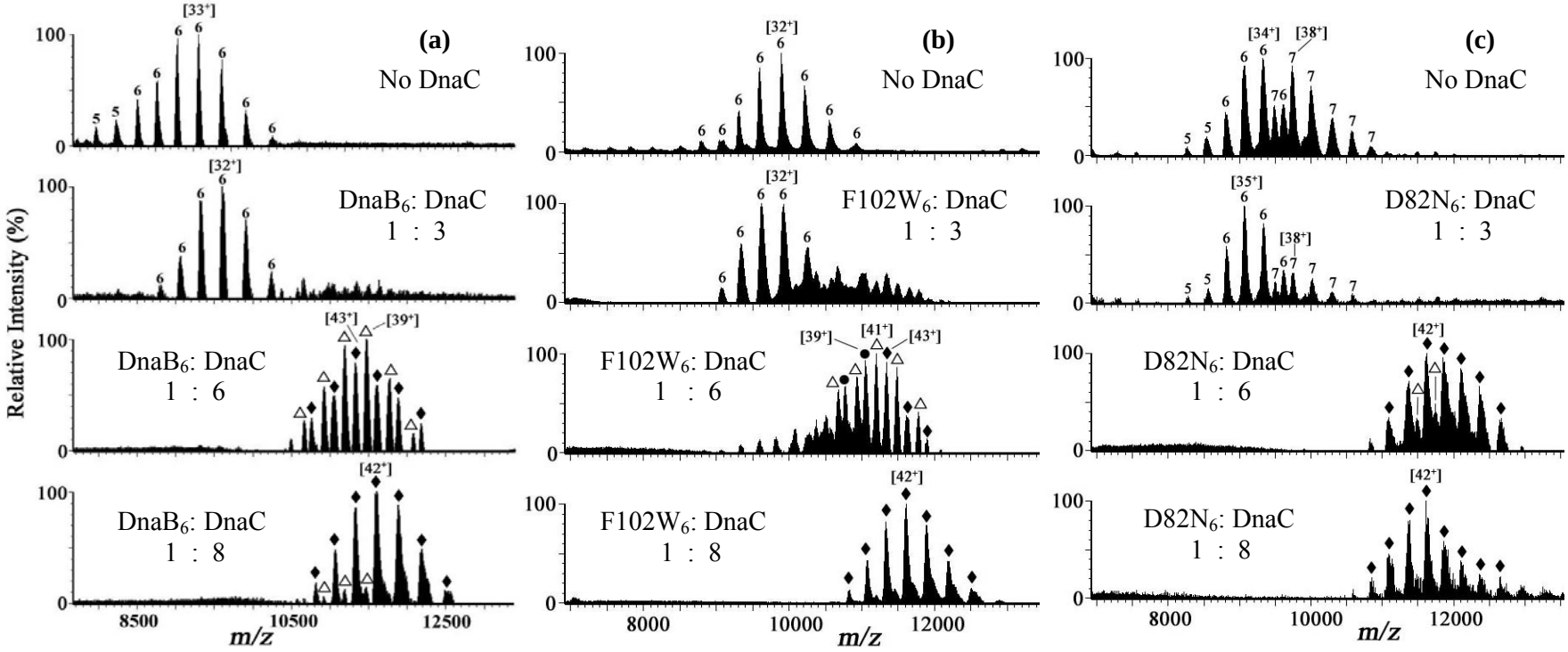


Figure 5.7 Positive ion nanoESI mass spectra of (a) $(DnaB)_6$, (b) $(F102W)_6$ and (c) $(D82N)_6$ treated with 0, 3, 6 or 8 equivalents of DnaC. The final concentration of $(DnaB)_6$, $(F102W)_6$ and $(D82N)_6$ was 10 μM. The complexes were prepared in 300 mM NH_4OAc, 0.1 mM $Mg(OAc)_2$, 0.1 mM ATP, pH 7.6. Numbers on the tops of ions indicate the oligomeric forms of DnaB or mutants. Numbers in square brackets indicate the number of charges for the most abundant ion present for each oligomeric form. Hexamer of $(DnaB)_6$, $(F102W)_6$, $(D82N)_6$ with $(DnaC)_4$ ●; $(DnaC)_5$ △ or $(DnaC)_6$ ♦.

Alternatively, a complex of DnaC with $(D82N)_7$ may have formed in the mixture, but aggregated precluding nanoESI-MS analysis. The latter seems unlikely since $(D82N)_6(DnaC)_6$ was observed when the two proteins were present in equal amounts suggesting that all of the added DnaC was present in the solution sampled for nanoESI-MS. The majority of ions observed when titrating $(D82N)_6$ with six equivalents of DnaC corresponded to a $(D82N)_6(DnaC)_6$ complex with only a small amount of $(D82N)_6(DnaC)_5$ complex. Previous analytical ultracentrifugation studies suggested that DnaB/DnaC complex can exist *in vivo* as a mixture of complexes with a different number of bound DnaC molecules.[175] The results obtained from the nanoESI-MS study are consistent with this proposal, and with previous observations that in the presence of an ATP analogue or ADP, a maximum of six DnaC monomers bind cooperatively to the DnaB hexamer.[175] In the current work, no ions from complexes containing less than four DnaC molecules bound to $(DnaB)_6$, $(F102W)_6$ or $(D82N)_6$ were observed.

The ability of $(F102W)_6$ or $(D82N)_6$ to form complexes with six molecules of DnaC monomer, similar to that of the wild-type $(DnaB)_6$, indicates that mutations at the Phe102 dimer interface and at the Asp82 at the N-terminal domain of full-length DnaB do not prevent the formation of these complexes. Hence this supports that the interaction between the hexameric helicase and DnaC occurs at the C-terminal domain of DnaB in $(DnaB)_6$, which is consistent with previous studies.[470,518]

5.4.4 Formation of complexes of DnaB and mutants with ADP

Close inspection of the ions in the nanoESI mass spectra shown in Figure 5.6 reveal fine structure. This was particularly evident when 0.1 $Mg(OAc)_2$ was present (Figure

5.6 (e) to (h)) since the peaks were less broad. Figure 5.8 shows an expansion of the *m/z* range of ~8920-9120 from Figure 5.6 (g), representing the 34^+ ion from the spectrum of F102H. The average mass difference between the species corresponding to these ions is ~430 Da. The expected mass differences upon binding of ADP, ADP + Mg, ATP and ATP + Mg are 427.2, 451.5, 507.2 and 531.5 Da, respectively. The observed mass difference thus suggests that each adduct ion represents the binding of successive ADP molecules to $(F102H)_6$. Since ATP was used in preparation of samples and DnaB is known to be a potent ATPase,[520-522] the presence of ADP molecules bound to the protein suggests that during the time of experiments $(F102H)_6$ hydrolysed ATP that was present in the solution.

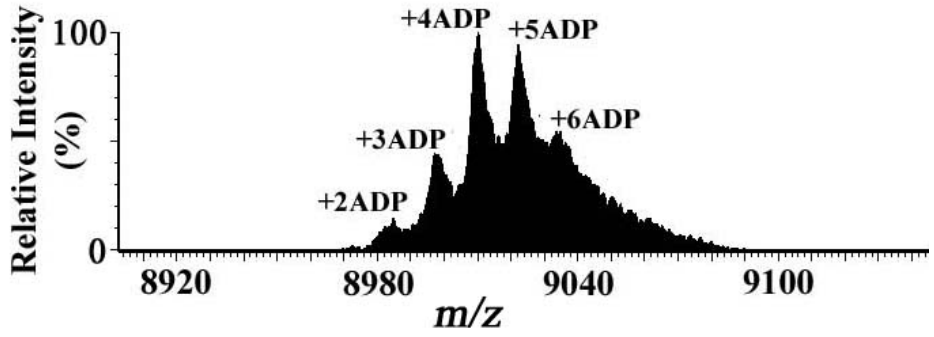


Figure 5.8 An expansion of the *m/z* range ~8920-9120 of the 34^+ ion from the nanoESI mass spectrum of F102H in 1000 mM NH_4OAc, 0.1 mM $Mg(OAc)_2$, 0.1 mM ATP, pH 7.6 (expansion of Figure 5.6 (g)).[519]

There were up to six molecules of ADP bound to $(F102H)_6$, with ions from the $(F102H)_6(ADP)_4$ and $(F102H)_6(ADP)_5$ complexes being the most abundant in the spectrum. However, this may not necessarily present a true picture of the relative abundances of these complexes with ADP in solution (or for $(F102H)_6$ *in vivo*), but may reflect the stability of $(F102H)_6(ADP)_n$ complexes in the mass spectrometer. Similar observations were also made for the wild-type DnaB, F102W, F102E and

D82N (data not shown). In separate experiments (Dixon NE; unpublished data) each of the mutants was shown to have ATPase activity. The observation of up to six ADP molecules binding to the hexameric helicase is consistent with fluorescence studies.[496] The ability to observe complexes of ADP with the hexameric form of DnaB and mutants is a first step for determining optimal experimental conditions for kinetics studies of the hydrolysis of ATP by various oligomeric forms of DnaB using nanoESI-MS analysis.

5.5 Conclusions

Oligomerisation of wild-type DnaB and a series of mutants of full-length DnaB (F102E, F102H, F102W, and D82N) were investigated using nanoESI-MS. It has been shown in earlier experiments that a heptameric form of DnaB was favoured in the presence of 30% (*v/v*) methanol.[497] In this work, mixtures of hexamer and heptamer as well as decamer and dodecamer (for (D82N) were observed in the absence of methanol. Optimal experimental conditions were established for the observation of different oligomeric forms of *E. coli* DnaB helicase and mutants using nanoESI-MS. It was found that at a higher concentration of Mg^{2+}, higher oligomeric forms were stabilised, except for the F102H protein where no higher oligomeric forms were observed. Stoichiometries of $(DnaB)_6$, $(F102W)_6$ and $(D82N)_6$ with DnaC monomers were also obtained. There were no less than four and up to a maximum of six DnaC molecules bound to $(DnaB)_6$, $(F102W)_6$ and $(D82N)_6$. These observations are consistent with previous studies, where DnaC was found to bind cooperatively to the hexameric helicase DnaB.[175,493]

Understanding the oligomerisation of wild-type DnaB and the mutants F102W, F102E, F102H and D82N will require further study. In future experiments, in-source CID and/or tandem MS/MS experiments will allow a comparison of the relative stabilities of oligomers formed by the different proteins. Additional information on the relative stabilities of the oligomeric forms of wild-type DnaB and mutants when complexed with loading partner DnaC as well as other replicative proteins may also provide insights into understanding how replication is initiated.

Chapter 6
Comparison of Unfolding Rates of Linear and Cyclised DnaB-N using Hydrogen/Deuterium Exchange

6.1 Introduction

Proteins are chains of amino acid residues that fold into a three-dimensional shape, providing distinct biological functions. The free termini of proteins are usually flexible, and consequently act as target points for the attack of proteolytic enzymes.[523] In recent years, naturally occurring circular proteins have been discovered in various microorganisms, plants and mammals.[524-529] Cyclisation of proteins confers stability against exoprotease digestion[527,530] and chemical denaturation,[531,532] and provides improved thermodynamic stability.[523,533,534] Researchers have previously attempted to increase enzyme stabilities by engineering disulfide bonds between termini. However, peptide bonds of the cyclised proteins have higher tolerance against exoprotease digestion and are inert in reducing environments. Therefore, the increased stability of cyclised proteins might enable their use *in vivo* as therapeutic agents.[523,533,535] Cyclisation of proteins may also have industrial applications as they can be used at higher temperatures. Other possible benefits of protein engineering and cyclisation might include positioning active site residues in a preferred conformation for catalysis, or reducing or diminishing the flexibility of peptide termini minimising entropic losses on receptor binding.[524]

In addition, it is believed that cyclised proteins could be engineered to have anti-HIV,[536] antibacterial and antifungal activity.[525,537]

In parallel with the discovery of increasing numbers of naturally occurring circular proteins, synthetic techniques have provided the opportunity to cyclise proteins where the N- and C-termini are in close proximity.[524] Goldenberg and Creighton produced the first circular protein by using a chemical cross-linking approach to prepare a cyclised version of bovine pancreatic trypsin inhibitor (BPTI).[538] However, they found no significant stabilization effects in the cyclised form of the protein BPTI. This was possibly because the direct cross-linking of the N- and C-termini introduced undesirable strain to the native structure.[538,539]

It was proposed that circular proteins could be synthesised by an intramolecular reaction between an N-terminal cysteine residue in one peptide and an α-thioester group of a second peptide.[540] In order to exploit this chemistry, recombinant proteins with N-terminal cysteine residues have been produced.[541,542] However, it was not until recently that these techniques have been well developed. This innovation was achieved by manipulation of a naturally occurring biological process known as protein splicing.[543-545]

6.1.1 Protein splicing

Over the past decade, many examples of self-catalysed peptide bond rearrangements have been discovered. These examples include autoprocessing of hedgehog protein,[546] formation of pyruvoyl enzymes,[547] autocleavage of glycosylasparaginase precursors[548] and protein splicing.[545,549,550] They have as a common feature the

observation that the self-catalysed reaction is prompted by the N→S or N→O acyl rearrangement of a peptide bond involving cysteine, serine or threonine. Protein splicing was first reported by Kane *et al*. in 1990.[551]

Inteins are proteins that can excise themselves post-translationally from nascent polypeptide chains, forming a new peptide bond between the new termini (N- and C-exteins).[552] This process is referred to as protein splicing, which is an intramolecular event[553] that does not require any coenzymes or sources of metabolic energy.[543] The detailed steps of protein splicing were not determined until the mid-1990s (1993-1996).[545,549,550,554-556] As currently understood, protein splicing involves four successive steps (see Figure 6.1), three of which require the catalytic properties of the inteins. Step 1 involves formation of a linear ester intermediate by an N→S or N→O acyl rearrangement of the peptide bond between the N-extein and the N-terminus of the intein at the upstream splice junction. This step involves the nucleophilic attack of the thiol or hydroxyl side chain of the intein N-terminal amino acid, cysteine (Cys) or serine (Ser) on the peptide carbonyl carbon of the adjacent N-extein. Step 2 involves the attack of the nucleophilic residue (Cys, Ser or Thr) at the downstream splice junction, forming a branched ester intermediate. Step 3 involves the cyclisation of the asparagine (Asn) residue at the intein C-terminal, coupled to cleavage of the branched ester intermediate, releasing the excised intein with a C-terminal aminosuccinimide residue and the ligated exteins joined by an ester bond. Step 4 involves spontaneous rearrangement of the ester linkage between the ligated exteins to the more stable peptide bond, resulting in a newly formed cyclised protein.

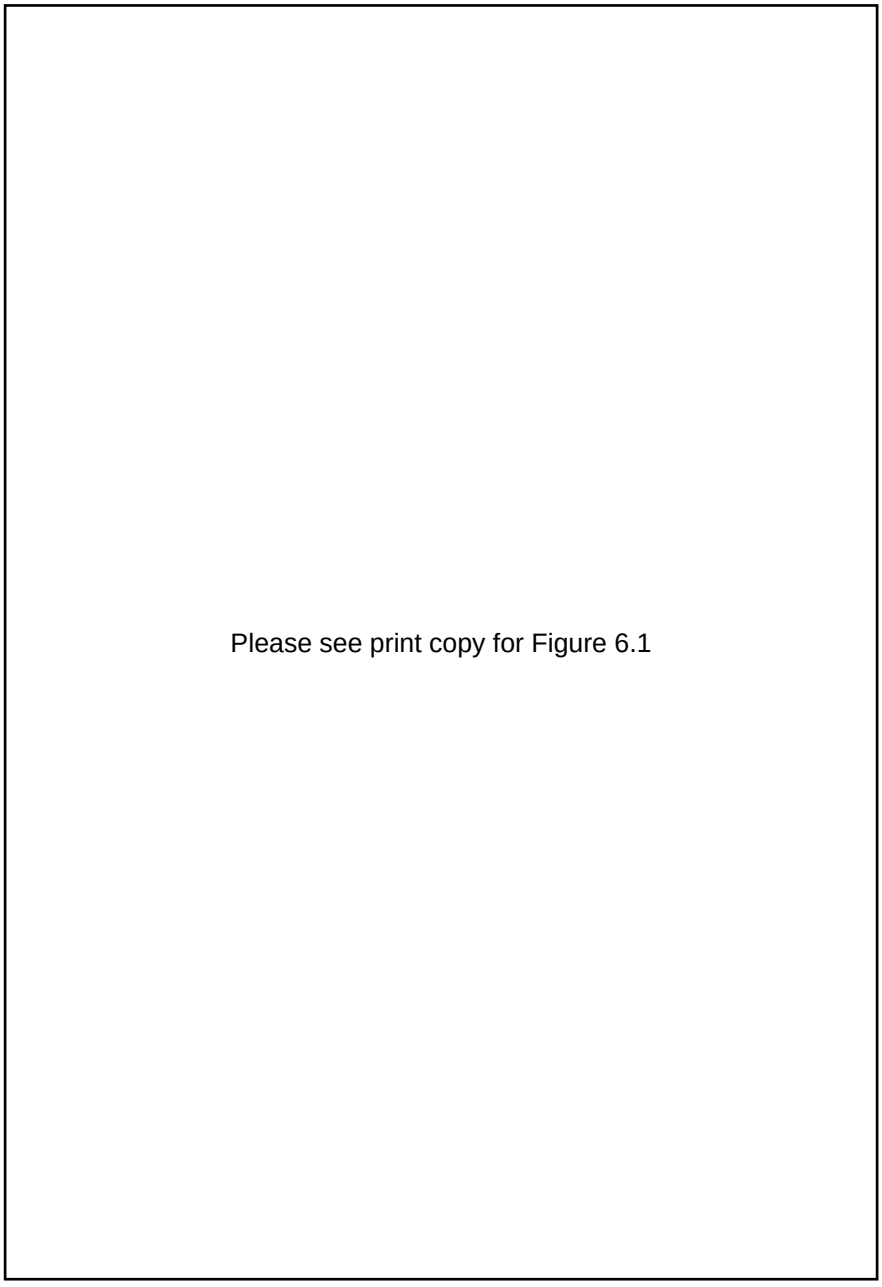


Figure 6.1 Proposed mechanism of protein splicing. X represents the sulphur (S) atom in Cys or the oxygen (O) atom in Ser or Thr. Step 1. A linear ester intermediate is formed as a result of N-X acyl rearrangement at the upstream splice junction. Step 2. An attack of the amino-terminal ester by the nucleophilic residues (Cys, Ser or Thr) leads to transesterification. Step 3. Cyclisation of asparagine results in the cleavage of the amide linkage at the intein C-terminus, yielding the free excised intein. A new cyclised protein is formed as a result of the second spontaneous N-X acyl rearrangement (step 4). Adapted from Perler *et al*..[557]

Over 100 inteins have been identified since the first discovery in 1990. Pietrovski developed a computer-based method for identifying new inteins,[558] while Perler constructed and updates a list of inteins for the intein database and registry (http://www.neb.com/neb/inteins.html). Inteins are most frequently found in proteins involved in nucleic acid processing such as DNA polymerase, but are also found in other types of proteins.[559,560] Several inteins have been used for protein splicing and trans-splicing *in vivo* and *in vitro*, including the *Mycobacterium tuberculosis* RecA intein,[561,562] the *Pyrococcus* Psp Pol-1 intein,[563] the DnaE split intein of the cyanobacterium *Synechocystis* sp. PCC6803,[533,564] and the DnaB split mini-intein of the cyanobacterium *Synechocystis* sp. PCC6803.[174] Examples of proteins which have been successfully cyclised by exploiting these reactions include β-lactamase,[523] dihydrofolate reductase (DHFR),[533] green fluorescent protein (GFP),[531] and the N-terminal domain of DnaB (DnaB-N).[174]

Many studies have shown that cyclisation of proteins enhances resistance against protease digestion, inertness to chemical denaturation and improves thermodynamic stability. In addition, protein cyclisation may be used to explore fundamental questions in protein folding. In particular, recent studies indicate that there is a statistically significant correlation between the rate of folding and the topology of a native fold. It has been suggested that cyclisation of proteins enhances their stability.[565,566] The backbone cyclisation is believed to lower the conformational entropy of a flexible peptide linker in both the folded[44] and unfolded state.[511]

6.2 Hydrogen/Deuterium Exchange (HDX)

Hydrogen/deuterium exchange (HDX) is a powerful technique for studying protein folding. It is a process that involves exchanging labile hydrogen atoms, namely those attached to nitrogen, oxygen or sulphur atoms in the protein molecules, with hydrogen or deuterium atoms of the solvent. The exchangeable hydrogen atoms located on amino acid side chains, as well as those on the N- or C-terminus exchange too rapidly to be measured by any technique so it is only those hydrogens in the peptide amide bonds that are measured.[123,567-571] For a labile hydrogen to exchange with solvent, it must be free from stable intramolecular hydrogen bonding and be accessible to the solvent.[123,572,573] Thus, those hydrogen atoms which are involved in hydrogen-bonding in various secondary structures, and those which are buried within the interior of a protein, will be prevented from hydrogen exchange, a phenomenon called "protection". During molecular motions ranging from local conformational fluctuations to global unfolding, these protected amide hydrogen atoms (NH) can be exchanged as the intramolecular hydrogen bonds are broken and the labile hydrogen atoms become accessible to the solvent. The rate of NH exchange can be affected by as much as 10 orders of magnitude as a result of interactions with neighbouring side chains.[574,575] This significant reduction in the exchange rates makes amide hydrogen exchange a sensitive probe for detecting and locating changes in protein conformations and dynamics.[567,570,572,576-580]

Analysis of hydrogen exchange rates is based on the well-established kinetic mechanism for slow amide hydrogen (NH) exchange in native proteins proposed by Linderstrøm-Lang and co-workers.[572,576,581] The exchange mechanism is given by (1):

$$NH_{(closed)} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} NH_{(open)} \xrightarrow[D_2O]{k_{ch}} ND \qquad (1)$$

$$k_{obs} = \frac{k_{op}k_{ch}}{k_{cl} + k_{ch}}$$

where k_{op} and k_{cl} are the rate constants for the opening and closing, respectively, of a particular exchangeable site and k_{ch} is the rate constant for chemical exchange (isotopic change) of the fully unprotected hydrogen. According to this model, a protected amide hydrogen (in the closed state) cannot exchange, however upon opening of the protein, the amide becomes accessible to the solvent allowing the exchange to occur.

There are two limiting situations for the above mechanism (see Figure 6.2).[576] The first occurs when interconversion between the closed and open state is much faster than the intrinsic chemical exchange rate, hence $k_{cl} >> k_{ch}$. The opening and closing will occur many times before the exchange takes place. This mechanism is called EX2. The observed rate constant for hydrogen exchange under EX2 conditions reflects the equilibrium constant between the closed and open states ($k_{obs} = k_{ch}(k_{op}/k_{cl})$). It has been shown that under EX2 conditions, the observed hydrogen exchange rate is related to the free energy of the protein unfolding process or exchange reaction (ΔG_{HX}) of a particular NH.[569] Thus, the hydrogen exchange rate in the EX2 limit can provide thermodynamic data.[569,582,583]

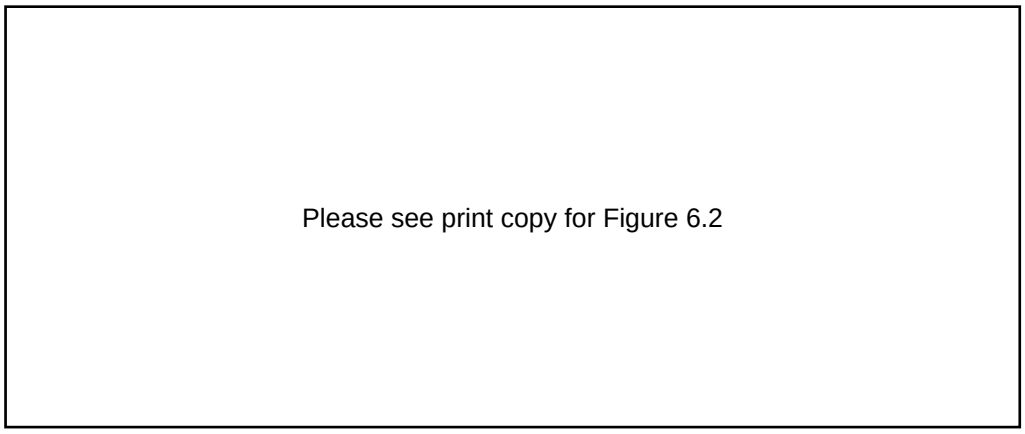


Figure 6.2 Kinetic mechanisms of amide hydrogen/deuterium exchange of native proteins. The EX2 mechanism occurs when $k_{cl} >> k_{ch}$, whereas the EX1 mechanism occurs when $k_{ch} >> k_{cl}$. Undeuterated sections of protein are shown in light grey and deuterated sections of the protein are shown in black. Adapted from Kaltashov and Eyles.[584]

The second situation occurs when interconversion between the closed and open structures is slow compared to the intrinsic (chemical) exchange rate, $k_{ch} >> k_{cl}$. As a result all of the amide hydrogens undergo isotopic exchange for each unfolding event (see Figure 6.2). Hence the experimentally observed rate constant for hydrogen exchange, k_{obs}, is directly related to the rate constant of the opening reaction ($k_{obs} = k_{op}$), which can be used to probe the kinetics rather than the thermodynamics of protein conformational changes.[569,583] This process is called EX1. Most naturally occurring amide exchange processes in proteins follow the EX2 mechanism. However, proteins can be induced to exhibit EX1 kinetics under some destabilising conditions such as by the addition of denaturants, extreme pH and/or high temperature.[582,583,585-587] Different amides of one protein may be able to undergo exchange by EX1 and EX2 mechanisms.[580]

6.3 Techniques for Probing Protein Conformational Dynamics and Interaction Sites of Protein Complexes

In the early 1950s, the use of HDX to study protein conformations relied on infrared spectroscopy[588] and density-gradient measurements.[589] Later on, radioactive tritium (^{3}H) was used instead of deuterium (^{2}H), allowing the HDX to be monitored using liquid scintillation counting.[590] NMR spectroscopy was introduced as a tool to measure HDX in 1958 by Saunders and Wishnia.[591] However, it was not until 20 years later that NMR was in widespread use for this purpose.[592,593] Although one-dimensional NMR facilitated observation of HDX at specific peptide bonds, its impact was limited because signals for most amide hydrogens occur over a narrow frequency band and therefore are not completely resolved even for small proteins.[572,593] Interest in using amide hydrogen exchange as a probe for protein structure and dynamics accelerated with the development of multi-dimensional high resolution NMR techniques, allowing more amide hydrogens to be resolved as a result of better resolution for NMR signals, thereby providing more residue-specific information on structural changes in proteins during folding/unfolding mechanisms.[594-597]

Mass spectrometry is ideal for monitoring hydrogen/deuterium exchange (HDX) because of its high resolution and accuracy, allowing detection of the mass difference of 1.0063 Da for each individual exchange between ^{1}H and ^{2}H. The application of HDX for probing protein conformational changes, coupled with ESI-MS, was first demonstrated by Katta and Chait in the early 1990s.[598] They used HDX MS to study acid- and alcohol-induced conformational changes in bovine ubiquitin.[598] Since then, the use of the HDX MS method to probe protein conformations and dynamics has

expanded rapidly.[599-603] Studies of amide hydrogen exchange have also been coupled with other soft ionisation mass spectrometric methods, including MALDI,[604,605] FAB[573,606] and FTICR.[607-609] In addition, studies of amide hydrogen exchange coupled with ESI-MS have also been used to characterise non-covalent protein-protein,[610,611] protein-substrate,[612] protein-ligand,[604,613,614] and protein-metal ion complexes.[385,407,600]

6.3.1 Hydrogen/deuterium exchange coupled with mass spectrometry (HDX MS)

Although NMR spectroscopy is a very powerful and non-destructive technique for studying hydrogen/deuterium exchange of proteins, it has several disadvantages. NMR spectroscopy requires a large amount of protein sample in order to obtain high quality NMR spectra. Furthermore, there are limitations to the size of proteins (< 30 kDa) that can be studied, as NMR signals can overlap extensively, resulting in poor resolution. The time taken to acquire NMR spectra may result in some denaturation, and the presence of paramagnetic cofactors such as ferric ions in proteins hinders NMR studies.

In recent years, the use of mass spectrometry (MS) to study HDX has become increasingly popular.[580,584,598,615-617] Owing to its sensitivity, only modest amounts of protein samples are required for analysis. MS is also suitable for analysis of larger proteins (> 30 kDa).[82] In contrast to NMR spectroscopy, MS data are not the average of all molecules in the sample. Instead co-existing protein species can be individually observed and the HDX can be monitored individually.[580] Mass spectrometry measures differences in individual populations within the bulk solution, while NMR

spectroscopy measures the average of the whole protein population. Therefore, these two methods are complementary techniques.

HDX observed using ESI-MS has been shown to be useful in probing protein conformational changes. For example, Gross and co-workers examined conformational properties of the apo- and holo-forms of the Ca^{2+}-binding protein calmodulin under near-native conditions.[600] The results demonstrated that the protein adopted a tight, less solvent-accessible conformation when the four Ca^{2+} ions were bound to the protein, as evidenced by a reduction in the number of amide hydrogens exchanged on addition of Ca^{2+}.

Mass spectrometry is a very useful tool for monitoring conformational dynamics of slow-exchanging proteins ($t_{1/2}$ > 1 minute).[602] However, it is also possible to use MS to study fast exchange process with a stopped-flow apparatus. For example, Miranker *et al*. used stopped-flow techniques to examine the kinetics of folding of hen egg-white lysozyme, and found that this occurred on a millisecond time scale.[580] Furthermore, this study highlighted the advantages offered by HDX MS for detecting intermediate folded states of the protein resulting from an EX1 mechanism.[580]

HDX MS alone cannot monitor exchange in a residue-specific manner, and is therefore unable to directly provide information on the location of hindered or exposed amide hydrogens. However, in combination with chemical cross-linking of side chains, proteolytic digestion and/or CID experiments it is possible to elucidate the interface regions of the complexes.[584,606,618-621] For example, Smith and co-workers coupled proteolytic digestion with HDX and MS to confirm the identity

of buried residues of horse heart cytochrome c.[606] Prior to HPLC-FAB-MS analysis, the protein was incubated in D_2O for different periods of time at different temperatures. After proton amide exchange was quenched by acidification to pH 2-3 at 0 °C, and rapid proteolytic digestion of the protein with pepsin,[606] the deuterium content of specific segments of the protein was then examined.[606] Other recent examples of application of HDX MS include mapping of protein interfaces in myoglobin,[622] and analysis of the conformations of bovine insulin,[623] hemoglobin[624] and oxidised and reduced *E. coli* thioredoxin.[625]

It is very desirable to be able to monitor HDX at individual amino acids. Attempts have been made to do this by using CID to fragment peptides into shorter pieces.[625-629] For example, in a study by Smith and co-workers, cytochrome c was labeled with deuterium at pD 7.0, acidified to quench HDX, and subsequently digested with pepsin.[629] The distribution of deuterium within the resulting peptide fragments was determined by MS and CID MS/MS experiments using a commercial ion trap instrument. The results showed that *b* ions from high-energy CID yielded the same deuterium distribution as those obtained from NMR experiments.[629] However, the deuterium content in most *y″* ions showed several discrepancies, suggesting that internal hydrogen exchange (scrambling) occurred during the gas phase fragmentation process.[625,629] The scrambling process seemed to depend on the amino acid sequence of the peptide, the nature of the charge carrier and other factors.[626,630]

A gas phase fragmentation method that shows promise for analysing sites of HDX and that may also offer a solution to the scrambling issue is electron-capture dissociation (ECD). ECD is an attractive method since it results in fragmentation

pathways where only *c*- and *z*-ions are produced, and exhibits minimal hydrogen scrambling in the gas phase.[631-634] ECD allows fragmentation of molecules via non-ergodic dissociation of ions with an odd number of electrons. This enables straightforward localisation of individual deuterium atoms after HDX in solution. ECD can be carefully controlled by varying the energy of an electron beam (5-7 eV), causing a large number of fragments even for larger proteins such as thiaminase (M_r 42 kDa).[608,634] This reduces the need for proteolytic digestion. Fragmentation can be further improved by directing the electron beam through the source region.[633]

6.4 Cyclisation of the N-terminal Domain of DnaB (DnaB-N)

The structure of the N-terminal domain of DnaB has been determined by both NMR spectroscopy[451] and X-ray crystallography.[452] The N- and C-termini of this domain are approximately 13 Å apart,[451] making it an attractive candidate for studying the effect of cyclisation using peptide linkers of different lengths. *E. coli* DnaB plays important roles during both the initiation and elongation stages of *E. coli* DNA replication (Chapter 5).[178,331,434] It is a hexameric protein composed of six identical 52 kDa subunits, each containing a small 12 kDa N-terminal domain and a larger 33 kDa C-terminal domain.[331,443,444,635] Both domains are required for helicase activity.[444,448] The N-terminal domain is essential for binding to primase and the C-terminal domain contains the binding sites for DNA, nucleotides and DnaC.[175,470,636] The functions of the N-terminal domain appear to include regulation of conformational changes which are important for helicase activity, regulation of ATP turnover, and coordination of helicase function with other replication proteins through specific protein-protein interactions.[178,331,440,453,464] Its many different roles in

DNA replication make DnaB-N an interesting candidate for studying protein conformational changes and unfolding.

Dixon and co-workers cyclised DnaB-N (N-terminal domain, residues 24-136 of DnaB helicase) *in vivo* using a synthetic *Synechocystis* sp. PCC6803 DnaB split mini-intein gene. The N- and C-termini were linked through a 9 amino acid flexible linker without introducing any conformational strain as judged by NMR experiments.[174] DnaB-N is an all-helical domain protein that consists of six α helices.[451,452] Its structure also contains two helical turns (residues 56-58 and 97-99). Charged amino acid side chains are evenly distributed over the protein surface and are easily accessible to the solvent. The hydrophobic core of DnaB-N is centred on the C-terminal end of the completely buried helix 1. An NMR structure of DnaB-N was determined for residues 30-134.[451]

Williams *et al.* found that the solution structure of cyclic DnaB-N was very similar to that of its linear version.[174] However, the unfolding of cyclic DnaB-N was found to occur at higher temperatures (~14 °C) than for linear DnaB-N.[174] Furthermore, differential scanning calorimetry studies showed that the cyclised DnaB-N, cz-DnaB-N, was more thermally stable (free energy, $\Delta\Delta G$, = 1.9 $kcalmol^{-1}$).[174] However, this study did not allow unambiguous evaluation of the thermostability of the cyclised protein since the extra nine amino acid residues of the linker were absent from the linear DnaB-N, lin-DnaB-N. Secondly, dimerisation of DnaB-N at high concentrations (millimolar range) interfered with the thermodynamic stabilization studies.[174]

In another study by Williams *et al*., these issues were addressed by using two new constructs: lin-DnaB-N(F102E) and cz-DnaB-N(F102E) (structures shown in Figures 6.3 (a) and (b), respectively), where Phe (F) 102 was replaced by a charged residue Glu (E) at the dimer interface so that only monomeric protein was present even at high concentrations.[511] Both new versions of linear and cyclised DnaB-N were created with the same amino acid sequences, resulting in the two proteins only differing by the mass of a water molecule.[511] Both proteins were found to fold and unfold reversibly as shown in previous studies.[174] These constructs and analogues with different linker lengths were studied in the current work.

6.5 Scope of This Chapter

The coupling of hydrogen/deuterium exchange with ESI-MS was used to probe the unfolding mechanism of DnaB-N. It was also used to probe the effect on cyclisation of the protein with a restriction at the N- and C-termini by inserting amino acid linkers with different lengths (3, 4, 5 and 9 amino acid residues). The amide proton exchange rates of the linear and cyclised proteins were also compared. Table 6.1 shows the amino acid sequences of the different linkers of DnaB-N used in this study.

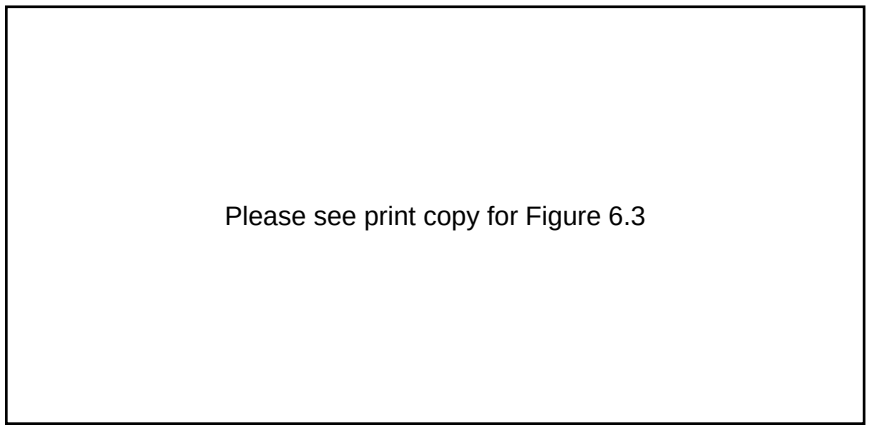


Figure 6.3 NMR structures of: (a) 9-linear-DnaB-N(F102E) and (b) 9-cyclised-DnaB-N(F102E), where the F102 residue is substituted by a charged residue, E, at the dimer interface preventing dimerisation at high protein concentrations. An arrow pointing between S and G indicates the fusion junction of the cyclised protein. Taken from Williams *et al.*.[511]

Table 6.1 Peptide sequences of the DnaB-N linkers used in this study.

Linker lengths (number of amino acids)	Amino acid sequences (N-terminus……DnaB-N…..C-terminus)*
3	**S**-DnaB-N-**TG**
4	**SF**-DnaB-N-**TG**
5	**SF**-DnaB-N-**TRG**
9	**SIEF**-DnaB-N-**TRESG**

* All of these proteins have the F102E mutation, and the amino acid sequences are shown as the linear version. In the cyclised versions, the Ser (S) of the linear version shown in the Table was joined through a peptide bond to the Gly (G) residue.

6.6 Results and Discussion

6.6.1 Hydrogen/deuterium exchange rates

Information about the mechanism of protein folding/unfolding can be directly obtained from amide proton exchange studies using ESI-MS. As a protein unfolds, its amide protons become exposed. In the presence of D_2O, the readily exchangeable

protons exchange for deuterons. Subsequently, the mass of the protein increases. In this work, a high concentration of D_2O (~99%) was used in order to ensure maximum hydrogen/deuterium exchange. The DnaB-N hydrogen/deuterium exchange was performed at pH 7.2 and 10 °C. Back-exchange of amide deuterons (protons replacing the deuterons) was minimised by using quenching solution (water:methanol:formic acid (90:9:1), pH 2.1) as the mobile phase for injection into the mass spectrometer. It has been shown that by decreasing the pH from 7 to 2-3 the amide proton exchange rate is reduced by ~10^4,[575] and by lowering the temperature from 20 to 0 °C the exchange rate is further reduced by an additional ten-fold.[575,577] The quenching solution travelled through a 200 μL sample loop that was immersed in ice, and a fast flow rate of 50 μL/min was used to transfer the deuterated sample into the mass spectrometer. In addition, the glass syringe used for injecting the exchanged proteins was also pre-cooled prior to sample injections.

Figure 6.4 shows the ESI mass spectra (transformed to a mass scale) obtained from HDX experiments for both linear and cyclised DnaB-N containing a three amino acid linker (M_r 12799 and 12781 Da, respectively). The top spectra in panels (a) and (b) (at time 0 min) were obtained from proteins that had not been exposed to D_2O solution. Subsequent spectra were obtained from proteins that had been exposed to D_2O solution at pH 7.2 for the indicated lengths of time. A striking feature is that two distinct populations (labelled A and B) were observed in the ESI mass spectra of both deuterated proteins. The mass increase of about 45 Da for 3-lin-DnaB-N (Figure 6.4 (a)) during the first two minutes was the result of exchange of the amide protons on the solvent-exposed surface of the folded protein, resulting in the observation of peak A (M_r 12844 ± 2 Da). The subsequent mass spectra obtained from the protein

after longer incubation times in the deuterated solvent showed a reduction in the relative intensity of peak A and the appearance of a new peak, B (M_r 12896 ± 0.5 Da). The mass difference between B and A (51 Da) indicates the number of deuterons incorporated as a result of protein unfolding. The bimodal isotopic pattern is characteristic of EX1 kinetics,[576] where $k_{ch} >> k_{cl}$ (see section 6.2). This pattern indicates that 51 amide protons were exposed in one unfolding event.

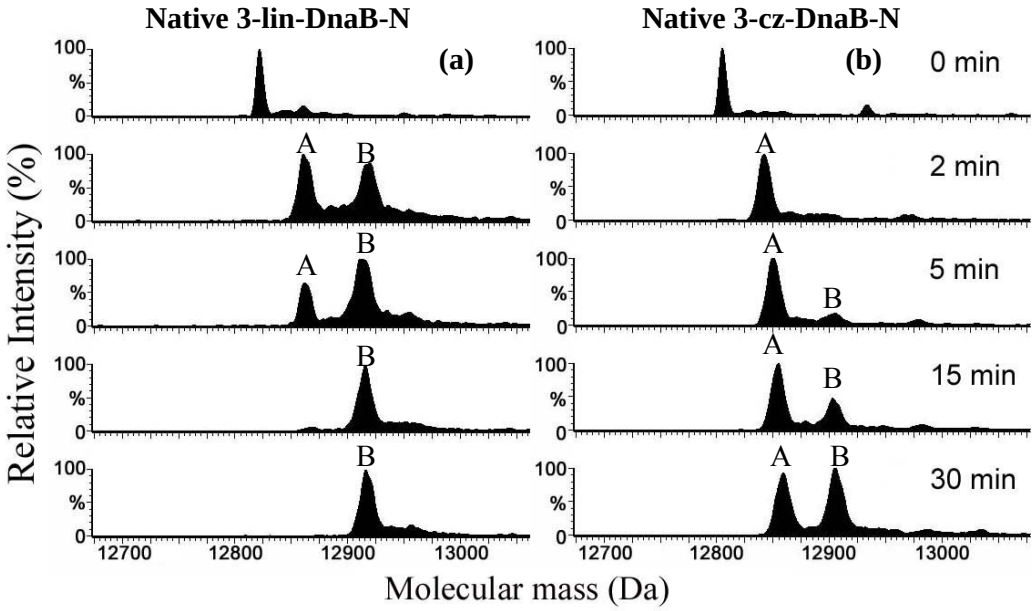


Figure 6.4 ESI-MS analysis of HDX as a function of time for: (a) 3-lin- and (b) 3-cz-DnaB-N in ~99% D_2O,10 mM NH_4OAc, pH 7.2, 10 °C. The ESI mass spectra were transformed to a mass scale using MassLynx software™. Peak A corresponds to the amide proton exchange of the solvent-exposed surface of the folded protein; peak B corresponds to the amide proton exchange of the unfolded protein.

Two populations were also observed in the ESI mass spectra obtained of the cyclised protein (Figure 6.4 (b)). The mass difference between type B and type A for 3-cz-DnaB-N was about 45 Da, hence 45 deuterons were incorporated as a result of one protein unfolding event. This is approximately the same as the difference observed for 3-lin-DnaB-N. However, for the cyclised protein a longer period of time was

taken for peak A to disappear and for peak B to form. In previous work in our laboratory, similar observations were made for 9-lin- and 9-cz-DnaB-N.[493]

It was surprising to observe EX1 kinetics for DnaB-N at near neutral pH. EX1 exchange, that is, where exchange of the slowly exchanging amide protons is governed by global unfolding of the protein (or part of the protein), is usually observed under denaturing conditions that potentiate global unfolding and allow uniform exposure of buried/protected amides to the solvent. These conditions occur in the presence of denaturants or under extreme conditions such as high or low pH or high temperature.[577,586,637,638] An EX1 mechanism has been observed by ESI-MS for proteins in denaturing conditions including the 62-residue IgG binding domain of protein L (pH 11, 60 °C),[586] oxidised and reduced *E. coli* thioredoxin and its cysteine alkylated derivatives in 2% acetic acid,[625,637] turkey ovomucoid third domain (pH 9.79)[570] and hen egg-white lysozyme (pH 3.8, 69 °C).[580]

EX1 and EX2 exchange behaviour can be differentiated through measurement of the observed exchange rate at varying pH. Since EX2 is strongly pH dependent, the rate would be expected to increase with increasing pH, whereas for EX1 the rate is independent of pH.[587] In our laboratory, it was shown that the unfolding rate for both 9-lin- and 9-cz-DnaB-N decreased slightly as the pH was increased from 6.8 to 7.8.[493] This is not characteristic of EX2 exchange, confirming the proposal that EX1 unfolding behaviour is observed for DnaB-N. The nine amino acid linker was originally chosen as it was calculated that this length allows the linker to be fully flexible.[174] Therefore, in comparisons of the conformational entropy of linear and cyclised proteins, linker properties and effects of restrictions of the distance between

the N- and C-termini on unfolding could be ignored.[565] Similar experiments were carried out to determine whether EX1 behaviour was also observed for the linear and cyclised DnaB-N with 4 and 5 amino acid linkers and also to compare the rates of unfolding for linear and cyclised proteins joined by different linkers.

Table 6.2 shows the average molecular masses of the A and B forms of the proteins. Table 6.3 shows the average number of amide protons exchanged (obtained from three separate experiments) for each of the proteins and includes the data for the 9 amino acid linker determined previously in this laboratory and in the current work.[493] The tables show results for when the exchange was carried out in both 10 mM and 100 mM NH_4OAc. First, it is clear that the number of slowly exchanging amide protons (those that are exposed on global unfolding, B-A), are always the same (approximately 50 in 10 mM NH_4OAc), and independent of the linker length used for cyclisation (Table 6.3). In 100 mM NH_4OAc, more amide protons rapidly exchanged (70 cf. 50 in 10 mM) suggesting that more amide protons were exposed to the surface under these conditions (A form). DnaB-N has previously been shown to be highly sensitive to pH and salt, possibly due to titration of a buried and uncharged (His64) side chain and of two other partially buried His residues.[451] The effect of salt concentration on the rate of HDX will be examined in more detail in section 6.6.2.

Table 6.2 Average molecular masses of peaks A and B from HDX of DnaB-N with different linker lengths obtained in 99% D_2O in 10 and 100 mM NH_4OAc, pH 7.2, 10 °C. These values were determined from 15-20 data points from three separate HDX experiments.

Please see print copy for Table 6.2

Table 6.3 Average numbers of amide protons exchanged, determined from three separate HDX experiments in 99% D_2O in 10 and 100 mM NH_4OAc, pH 7.2, 10 °C.

Please see print copy for Table 6.3

The HDX data for all the linear and cyclised proteins were then analysed and compared to the rate of loss of form A (same rate as appearance of form B). Figure 6.5 shows plots of relative abundances of peaks A and B as a function of exchange time in 10 mM NH_4OAc for each of the proteins examined. Each data point is the average obtained from three sets of experiments. The relative abundance of each peak was obtained by dividing the relative intensity of the peak by the sum of the relative intensities of peaks A and B. A comparison between linear DnaB-N proteins with varying linker lengths of 3, 4 and 5 amino acids (see Figures 6.5 (a), (b) and (c), respectively) indicated that the time required for complete conversion of type A protein to type B was approximately 20 minutes. However, the time required for the cyclised proteins to convert from type A to type B (see Figure 6.5 (d), (e) and (f)) varied between 90 and 180 minutes. Since DnaB-N exhibits EX1 behaviour, the observed rate reflects the protein unfolding rate ($k_{obs} = k_{op}$). The unfolding rate constant, k_{op}, was obtained by plotting the natural log of the relative abundance of peak A (folded state of the protein) against time. These plots are shown in Figure 6.6. Table 6.4 summarises the first order rate constants obtained from these plots.

DnaB-N is a small all-helical protein with a calculated contact order (average sequence separation between contacting residues in the native state) of 12.6.[511] Based on correlation studies between protein folding rates and contact order by Plaxco *et al*., the folding rate constant of DnaB-N was expected to be at least 1000 s^{-1}.[639,640] However, the folding rate constant was slower (see Table 6.4) which is also consistent with an EX1 process.

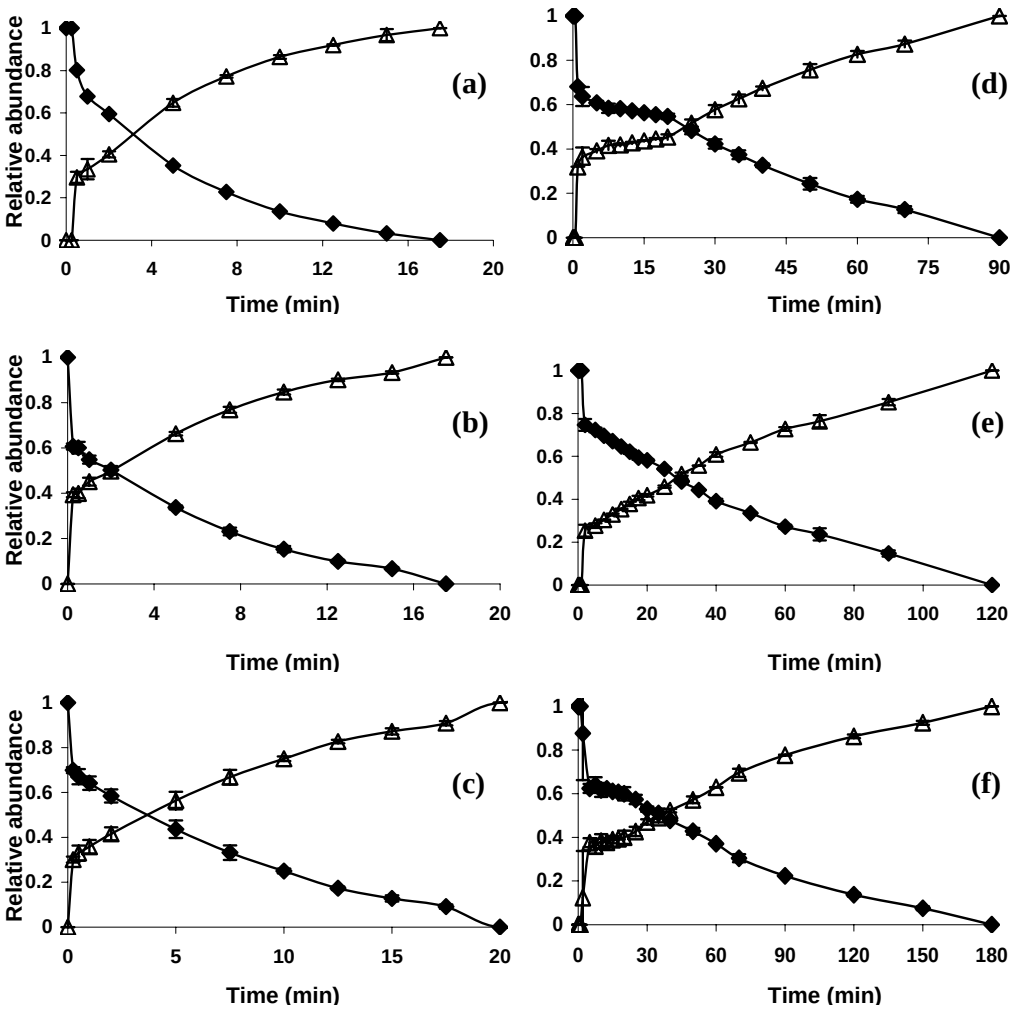


Figure 6.5 Relative abundances of peaks A and B in ESI mass spectra obtained during HDX experiments for linear DnaB-N with linkers containing 3, 4 and 5 amino acids ((a), (b) and (c), respectively), and for cyclised DnaB-N with linkers containing 3, 4 and 5 amino acids ((d), (e) and (f), respectively). All experiments were performed in 10 mM NH_4OAc, 99% D_2O, pH 7.2, 10 °C. Each data point is the average obtained from three sets of experiments. ◆ A form; △ B form.

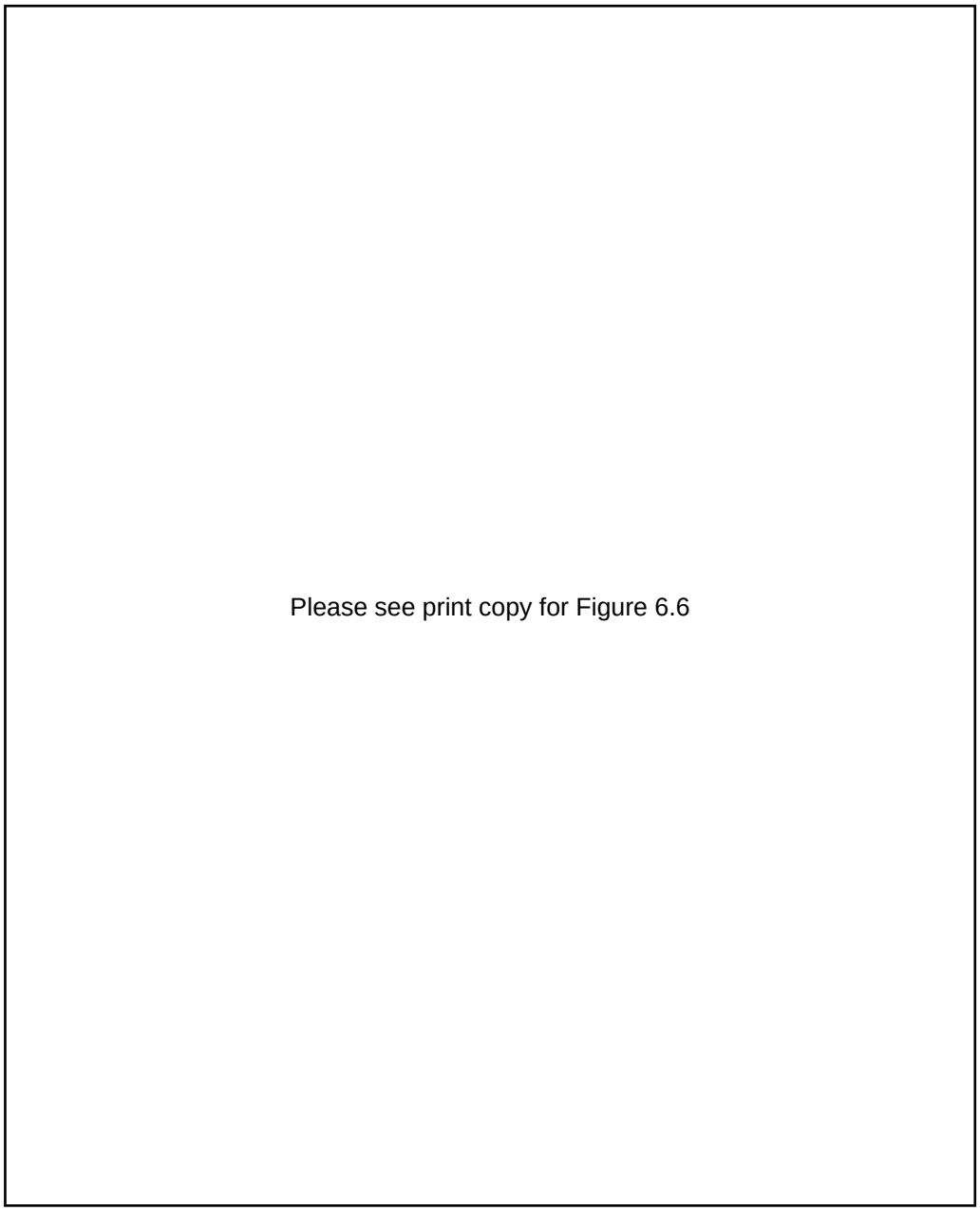


Figure 6.6 First order plots of HDX of linear and cyclised DnaB-N with amino acid linkers composed of: (a) 3; (b) 4; (c) 5; and (d) 9 amino acids. All experiments were performed in 10 mM NH_4OAc, ~99% D_2O, pH 7.2, 10 °C. Each data point is the average of three independent experiments. * Data obtained by Stephen J Watt. ◆ Linear protein; ■ cyclised protein.

Table 6.4 First order rate constants for unfolding of linear and cyclised DnaB-N with different linker lengths in 99% D_2O, 10 mM NH_4OAc, pH 7.2, 10 °C.

Please see print copy for Table 6.4

There are several interesting features evident in the data in Table 6.4. First, the unfolding rates for all the cyclised proteins are lower than for their linear counterparts. This observation is consistent with the previous NMR and MS studies of 9-lin- and 9-cz-DnaB-N, where the amide proton exchange rate of the cyclised protein was approximately ten-fold slower than that of its linear counterpart.[511] This indicates that cyclisation increases the stability of the protein, which is consistent with the theory that backbone cyclisation reduces the entropy of a flexible peptide linker.[565] Secondly, the linker length has no effect on the degree of stabilisation caused by cyclisation, i.e. all the ratios ($k_{op(lin)}/k_{op(cz)}$) are ~8. These data show that global unfolding of the cyclised proteins was about 8 times less frequent than for the linear proteins. This has implications for the mechanism of global unfolding in DnaB-N. Based on results for the proteins with the 9 amino acid linker, it has previously been proposed that global unfolding events of DnaB-N involve a large expansion of the N- and C-termini that exceeds the lengths of the peptide linker.[511] This proposal is also consistent with the data for the different linker lengths. The results here showed that backbone cyclisation did not stop DnaB-N from global

unfolding, but merely restricted the spatial separation between the N- and C-termini of the protein, leading to a slower unfolding rate. These data also suggest that initiation of the unfolding of DnaB-N can occur at different sites on the molecule.

As the linker length increased, the unfolding rate for both linear and cyclised DnaB-N decreased, suggesting that proteins with longer linker lengths have greater stabilities. This at first may be surprising, however, a previous study showed that introducing the nine amino acid residue linker to DnaB-N did not introduce any conformational strain to the protein.[174] It has previously been shown that protein stability is dependent on the length of a peptide linker attached to the native structure when the linker length is shorter than the distance separating the N- and C-termini.[566] On the other hand, when the linker is too long, steric effects could push the termini apart and as a result the structure may be distorted or unfolded.

6.6.2 Effect of salt concentration on H/D exchange rates

It has been shown previously by CD spectroscopy that both linear and cyclised DnaB-N were stabilised at high salt concentrations.[511] In the current work, the amide exchange rates of the linear and cyclised proteins were compared at two different NH_4OAc concentrations (10 and 100 mM). Figure 6.7 shows the ESI mass spectra (transformed to a mass scale) of the linear and cyclised proteins with a three amino acid linker subjected to HDX in ~99% D_2O, 100 mM NH_4OAc, pH 7.2, 10 °C. The top spectra in panels (a) and (b) at time 0 min were obtained from 3-lin- and 3-cz-DnaB-N, respectively, that had not been exposed to a D_2O solution. Subsequent spectra were obtained using proteins that had been exposed to D_2O (pH 7.2, 10 °C) for the indicated lengths of time. Similar to the ESI mass spectra obtained for 10 mM

NH_4OAc (Figure 6.4), two distinct populations (labelled A and B) were observed for both the linear and cyclised proteins. The difference in mass between the A and B form was about 39 Da for 3-lin-DnaB-N (Figure 6.7 (a) and Tables 6.2 and 6.3). A comparison of Figures 6.7 (a) and (b) reveals that a longer period of time was required to completely convert form A to form B for the cyclised protein.

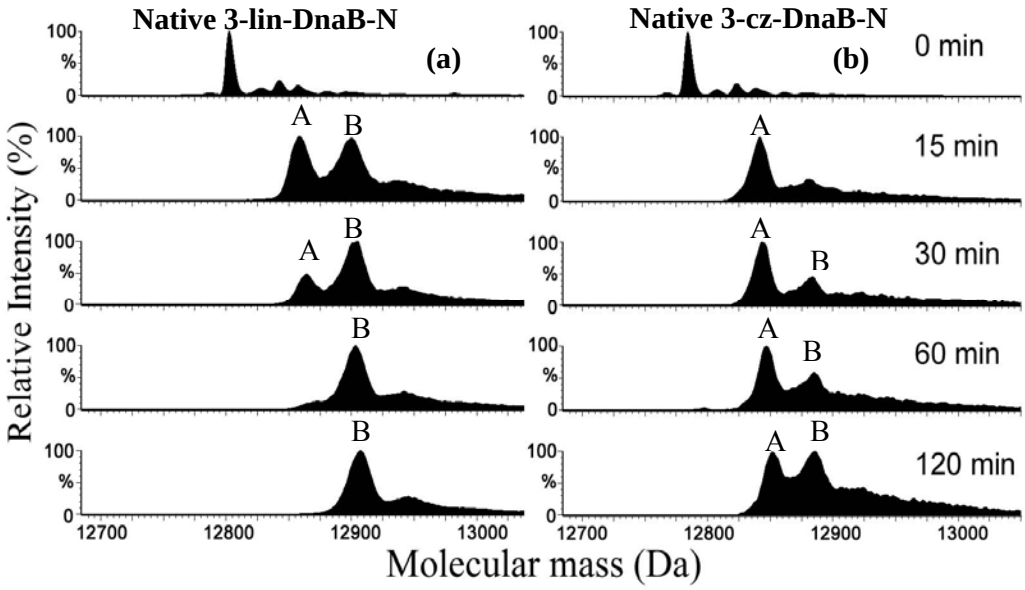


Figure 6.7 ESI-MS analysis of HDX as a function of time for: (a) 3-lin- and (b) 3-cz-DnaB-N in ~99% D_2O, 100 mM NH_4OAc, pH 7.2, 10 °C.

Figure 6.8 shows plots of relative abundances of peaks A and B observed for the different linear and cyclised DnaB-N molecules, as a function of exchange time in 99% D_2O, 100 mM NH_4OAc, pH 7.2, 10 °C. Each data point is the average obtained from three separate sets of experiments. The unfolding rate constants (k_{op}) for all the linear and cyclised proteins in the presence of 100 mM NH_4OAc were obtained from the slope of the plots of the logarithm of the relative abundance of peak A against time (see Figure 6.9). The rate constants for linear and cyclised DnaB-N with varying linker lengths obtained from HDX experiments in a solution containing 100 mM

NH_4OAc (99% D_2O) are summarised in Table 6.5. The results presented in Table 6.5 show that as the linker length increases, the unfolding rate decreases. This is similar to the results shown in Table 6.4. The exception, however, is in the rates obtained from the proteins containing the three and four amino acid linkers, which are the same within experimental error. The ratios between the unfolding rates for the linear and cyclised proteins ($k_{op(lin)}/k_{op(cz)}$) varied between 8 and 11. These values are approximately the same when taking into account experimental errors, and are also the same as those obtained when the proteins were studied in 10 mM NH_4OAc. These HDX results show that all the proteins (both linear and cyclised) are more stable in the presence of a higher NH_4OAc concentration, as evidenced by a five-fold reduction in the unfolding rate (compare Tables 6.4 and 6.5). Furthermore, reference to Table 6.3, suggests that at a higher salt concentration, a structure where ~40 amide protons (B-A) are buried is stabilised relative to the structure in 10 mM MH_4OAc, where ~50 amide protons are buried.

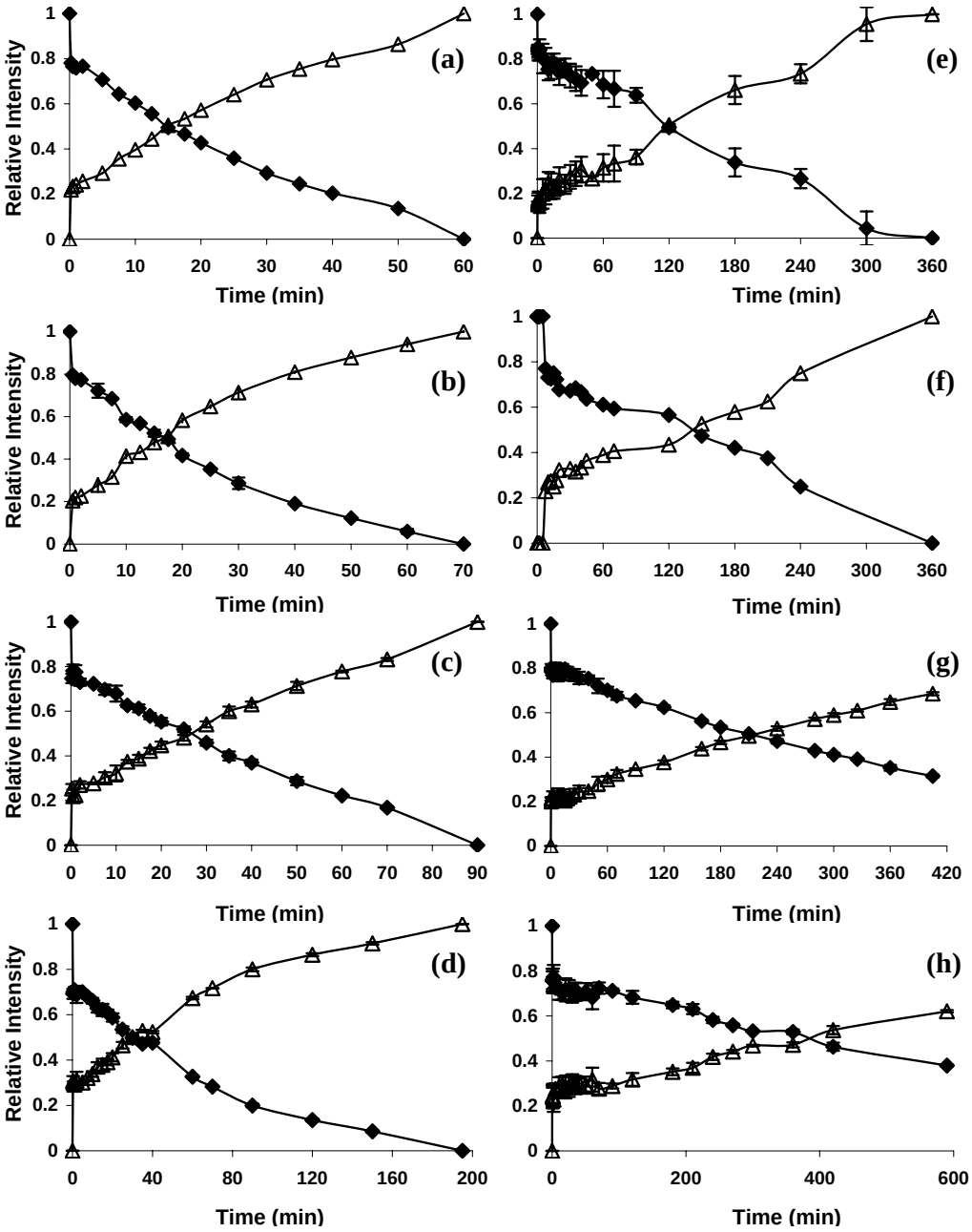


Figure 6.8 Relative abundances of peaks A and B in ESI mass spectra obtained during HDX experiments for linear DnaB-N with linkers containing 3, 4, 5 and 9 amino acids ((a), (b), (c) and (d), respectively), and for cyclised DnaB-N with linkers containing 3, 4, 5 and 9 amino acids ((e), (f), (g) and (h), respectively). All experiments were performed in 100 mM NH_4OAc, 99% D_2O, pH 7.2, 10 °C. ◆ A form; △ B form.

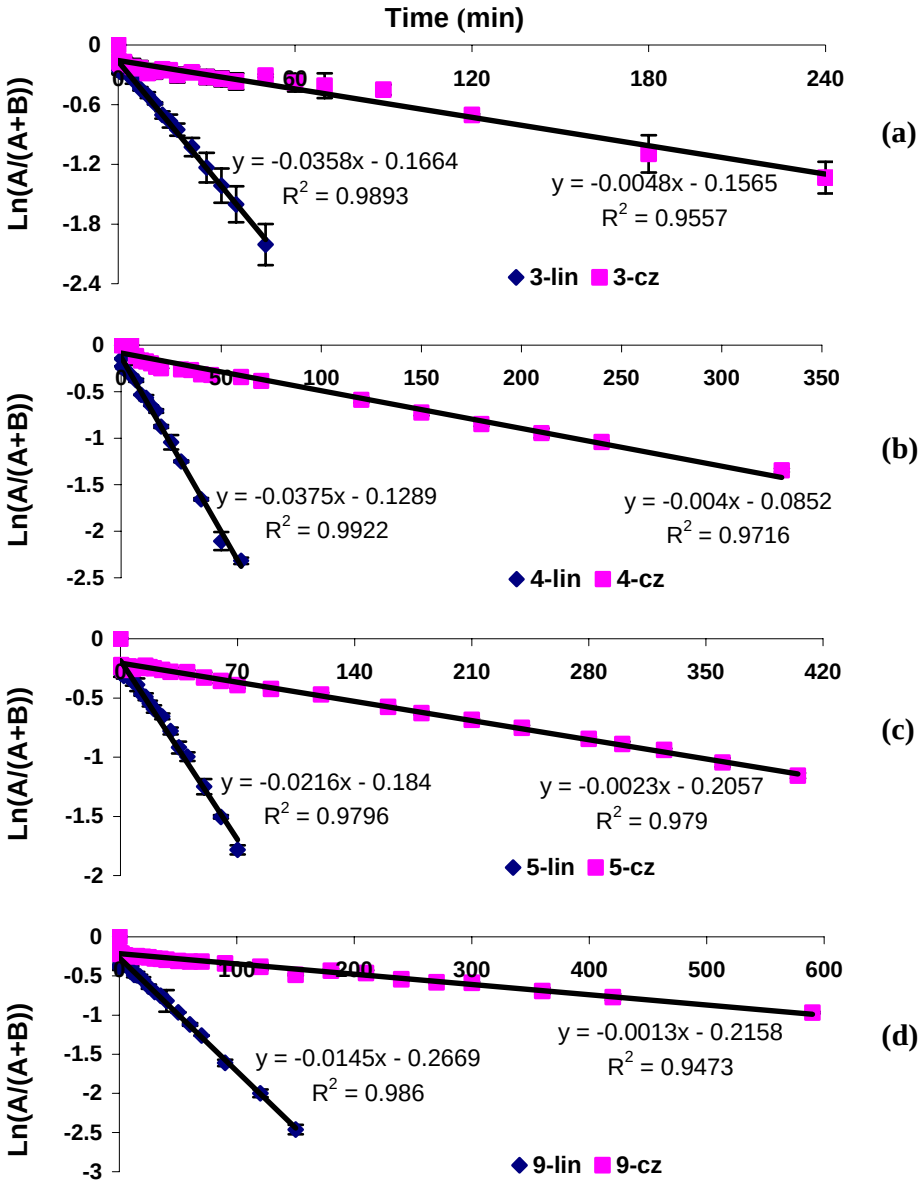


Figure 6.9 First order plots of HDX of linear and cyclised DnaB-N containing linkers comprised of: (a) 3; (b) 4; (c) 5; and (d) 9 amino acid linker in 100 mM NH_4OAc, 99% D_2O), pH 7.2, 10 °C. Each data point was determined from the average of three separate experiments. ◆ Linear protein; ■ cyclised protein.

Table 6.5 First order rate constants for unfolding of linear and cyclised DnaB-N with different linker lengths in 99% D_2O, 100 mM NH_4OAc, pH 7.2, 10 °C.

Linker length (number of amino acids)	k_{op} (min^{-1})[a]		$k_{op(lin)}/k_{op(cz)}$
	Linear	Cyclised	
3	0.036 ± 0.005	0.0048 ± 0.0009	8 ± 2
4	0.0375 ± 0.003	0.00405 ± 0.00005	9.3 ± 0.7
5	0.0216 ± 0.0005	0.0023 ± 0.0001	9.4 ± 0.5
9	0.0145 ± 0.0009	0.0013 ± 0.00008	11 ± 1

[a] Rate constants were determined from the average of three sets of experiments.

6.7 Conclusions

Hydrogen/deuterium exchange coupled with ESI-MS (HDX MS) proved to be a powerful tool for studying protein dynamics. In the current work, HDX MS was used to determine the unfolding rates for both linear and cyclised DnaB-N. It was interesting to observe that at near neutral pH the HDX process for all the linear and cyclised DnaB-N containing different linker lengths occurred via an EX1 process. Since the HDX of DnaB-N followed the EX1 regime, the exchange rate constants observed therefore reflect the unfolding rates of the proteins. HDX data showed that the rates for all the cyclised DnaB-N proteins were approximately eight to ten-fold slower than for the corresponding linear proteins. This observation is consistent with previous HDX data determined by NMR spectroscopy confirming that backbone cyclisation played an important role in protein stabilisation. It is also important to note that cyclisation did not stop the proteins from unfolding, suggesting that the unfolding of DnaB-N does not necessarily commence by unzipping from the N- and C-termini.

The effect of varying salt concentrations on the exchange rates was also investigated. DnaB-N (both linear and cyclised) proteins appeared to be more stable against global unfolding in the presence of a higher salt concentration. This was evident by a five-fold reduction in the unfolding rates as the salt concentration was increased from 10 mM to 100 mM.

REFERENCES

1. Dempster, A. J. *Phys. Rev.* **1921**, 18, 415-422.

2. Munson, M. S. B.; Field, F. H. *J. Am. Chem. Soc.* **1966**, 88, 2621-2630.

3. Beckey, H. D. *Intl J. Mass Spectrom. Ion Process* **1969**, 2, 500-503.

4. Hsu, C. S.; Green, M. *Rapid Commun. Mass Spectrom.* **2001**, 15, 236-239.

5. Torgerson, D. F.; Skowronski, R. P.; Macfarlane, R. D. *Biochem. Biophys. Res. Commun.* **1974**, 60, 616-621.

6. Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N.; Bycroft, B. W. *Biochem. Biophys. Res. Commun.* **1981**, 101, 632-638.

7. Akashi, S. *Med. Res. Rev.* **2006**, 26, 339-368.

8. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yohida, T. *Rapid Commun. Mass Spectrom.* **1988**, 2, 151-153.

9. Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, 60, 2299-2301.

10. Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, 88, 4451-4459.

11. Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, 88, 4671-4675.

12. Anderegg, R. J.; Wagner, D. S.; Blackburn, R. K.; Opiteck, G. J.; Jorgenson, J. W. *J. Protein Chem.* **1997**, 16, 523-526.

13. Huang, E. C.; Pramanik, B. N.; Tsarbopoulos, A.; Reichert, P.; Ganguly, A. K.; Trotta, P. P.; Nagabhushan, T. L.; Covey, T. R. *J. Am. Soc. Mass Spectrom.* **1993**, 4, 624-630.

14. Pandey, A.; Mann, M. *Nature* **2000**, 405, 837-846.

15. Roepstorff, P. *Curr. Opin. Biotechnol.* **1997**, 8, 6-13.

16. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, 246, 64-71.

17. Griffiths, W. J.; Jonsson, A. P.; Liu, S.; Rai, D. K.; Wang, Y. *Biochem. J.* **2001**, 355, 545-561.

18. Aebersold, R.; Mann, M. *Nature* **2003**, 422, 198-207.

19. Nelson, R. W.; Dogruel, D.; Williams, P. *Rapid Commun. Mass Spectrom.* **1994**, 8, 627-631.

20. Chernushevich, I. V.; Fell, L. M.; Bloomfield, N.; Metalnikov, P. S.; Loboda, A. V. *Rapid Commun. Mass Spectrom.* **2003**, 17, 1416-1424.

21. Krutchinsky, A. N.; Loboda, A. V.; Spicer, V. L.; Dworschak, R.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **1998**, 12, 508-518.

22. Krone, J. R.; Nelson, R. W.; Dogruel, D.; Williams, P.; Granzow, R. *Anal. Biochem.* **1997**, 244, 124-132.

23. Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev.* **2004**, 23, 368-389.

24. Musselman, B.; Jonscher, K. *Am. Genomic/Proteomic Technol.* **2002**, 2, 22, 24-27.

25. Stensballe, A.; Jensen, O. N. *Proteomics* **2001**, 1, 955-966.

26. Wang, Y.; Schneider Bradley, B.; Covey Thomas, R.; Pawliszyn, J. *Anal. Chem.* **2005**, 77, 8095-8101.

27. Hillenkamp, F. *NATO ASI Ser., Ser. C: Math. Phys. Sci* **1998**, 510, 181-191.

28. Strupat, K.; Rogniaux, H.; Van Dorsselaer, A.; Roth, J.; Vogl, T. *J. Am. Soc. Mass Spectrom.* **2000**, 11, 780-788.

29. Friess, S. D.; Daniel, J. M.; Hartmann, R.; Zenobi, R. *Intl J. Mass Spectrom.***2002**, 219, 269-281.

30. Zehl, M.; Allmaier, G. *Anal. Chem.* **2005**, 77, 103-110.

31. Tissot, B.; Gonnet, F.; Iborra, A.; Berthou, C.; Thielens, N.; Arlaud, G. J.; Daniel, R. *Biochemistry* **2005**, 44, 2602-2609.

32. Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Intl J. Mass Spectrom.* **2002**, 216, 1-27.

33. Kiselar, J. G.; Downard, K. M. *J. Am. Soc. Mass Spectrom.* **2000**, 11, 746-750.

34. Strupat, K.; Sagi, D.; Peter-Katalinic, J.; Bonisch, H.; Schafer, G. *Analyst* **2000**, 125, 563-567.

35. Dole, M.; Hines, R. L.; Mack, L. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *Macromolecules* **1968**, 1, 96-97.

36. Light-Wahl, K. J.; Schwartz, B. L.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, 116, 5271-5278.

37. Przybylski, M.; Glocker, M. O. *Angew. Chem., Int. Ed.* **1996**, 35, 806-826.

38. Loo, J. A. *Intl J. Mass Spectrom.***2000**, 200, 175-186.

39. Gaskell, S. J. *J. Mass Spectrom.* **1997**, 32, 677-688.

40. Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, 64, 2287-2294.

41. Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, 71, 4451-4463.

42. Dole, M.; Mack, L. L.; Hines, R. L. *J. Chem. Phys.* **1968**, 49, 2240-2249.

43. Kebarle, P.; Peschke, M. *Anal. Chim. Acta* **2000**, 406, 11-35.

44. Tang, L.; Kebarle, P. *Anal. Chem.* **1991**, 63, 2709-2715.

45. Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, 65, 3654-3668.

46. Kebarle, P.; Tang, L. *Anal. Chem.* **1993**, 65, 972A-986A.

47. Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, 11, 450-457.

48. Guo, M. Q.; Zhang, S. Q.; Song, F. R.; Wang, D. W.; Liu, Z. Q.; Liu, S. Y. *J. Mass Spectrom.* **2003**, 38, 723-731.

49. Wang, Y.; Schubert, M.; Ingendoh, A.; Franzen, J. *Rapid Commun. Mass Spectrom.* **2000**, 14, 12-17.

50. Stephenson, J. L., Jr.; McLuckey, S. A. *J. Mass Spectrom.* **1998**, 33, 664-672.

51. Chowdhury, S. K.; Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1990**, 4, 81-87.

52. Katta, V.; Chowdhury, S. K.; Chait, B. T. *Anal. Chem.* **1991**, 63, 174-178.

53. Jensen, O. N.; Kulkarni, S.; Aldrich, J. V.; Barofsky, D. F. *Nucleic Acids Res.* **1996**, 24, 3866-3872.

54. Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrom. Rev.* **2000**, 19, 65-107.

55. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. *J. Mass Spectrom.* **2001**, 36, 849-865.

56. Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. *Rapid Commun. Mass Spectrom.* **1996**, 10, 889-896.

57. Sobott, F.; Hernandez, H.; McCammon, M. G.; Tito, M. A.; Robinson, C. V. *Anal. Chem.* **2002**, 74, 1402-1407.

58. Fitzgerald, M. C.; Chernushevich, I.; Standing, K. G.; Whitman, C. P.; Kent, S. B. H. *Proc. Natl Acad. Sci. USA* **1996**, 93, 6851-6856.

59. Loo, J. A. *J. Mass Spectrom.* **1995**, 30, 180-183.

60. Light-Wahl, K. J.; Winger, B. E.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, 115, 5869-5870.

61. Pramanik, B. N.; Bartner, P. L.; Mirza, U. A.; Liu, Y.-H.; Ganguly, A. K. *J. Mass Spectrom.* **1998**, 33, 911-920.

62. Zhu, M. M.; Chitta, R.; Gross, M. L. *Intl J. Mass Spectrom.***2005**, 240, 213-220.

63. Tjernberg, A.; Carnoe, S.; Oliv, F.; Benkestock, K.; Edlund, P.-O.; Griffiths, W. J.; Hallen, D. *Anal. Chem.* **2004**, 76, 4325-4331.

64. Schwartz, B. L.; Gale, D. C.; Smith, R. D.; Chilkoti, A.; Stayton, P. S. *J. Mass Spectrom.* **1995**, 30, 1095-1102.

65. De Vriendt, K.; Sandra, K.; Desmet, T.; Nerinckx, W.; Van Beeumen, J.; Devreese, B. *Rapid Commun. Mass Spectrom.* **2004**, 18, 3061-3067.

66. Benkestock, K.; Edlund, P. O.; Roeraade, J. *Rapid Commun. Mass Spectrom.* **2002**, 16, 2054-2059.

67. Rostom, A. A.; Sunde, M.; Richardson, S. J.; Schreiber, G.; Jarvis, S.; Bateman, R.; Dobson, C. M.; Robinson, C. V. *Proteins-Struct. Func. Genet.* **1998**, 3-11.

68. Hu, P.; Loo, J. A. *J. Mass Spectrom.* **1995**, 30, 1076-1082.

69. Loo, J. A. *Intl J. Mass Spectrom.***2001**, 204, 113-123.

70. Cheng, X.; Morin, P. E.; Harms, A. C.; Bruce, J. E.; Ben-David, Y.; Smith, R. D. *Anal. Biochem.* **1996**, 239, 35-40.

71. Kapur, A.; Beck, J. L.; Brown, S. E.; Dixon, N. E.; Sheil, M. M. *Protein Sci.* **2002**, 11, 147-157.

72. Kamadurai, H. B.; Subramaniam, S.; Jones, R. B.; Green-Church, K. B.; Foster, M. P. *Protein Sci.* **2003**, 12, 620-626.

73. Light-Wahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G., II; Thrall, B. D.; Smith, R. D. *J. Am. Chem. Soc.* **1993**, 115, 803-804.

74. Ganem, B.; Li, Y. T.; Henion, J. D. *Tetrahedron Lett.* **1993**, 34, 1445-1448.

75. Goodlett, D. R.; Camp, D. G.; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, 22, 181-183.

76. Hsieh, Y. L.; Li, Y. T.; Henion, J. D.; Ganem, B. *Biol. Mass Spectrom.* **1994**, 23, 272-276.

77. David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, 74, 2029-2033.

78. Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J.-F.; Mergny, J.-L. *Biochemistry* **2003**, 42, 10361-10371.

79. Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; Pauw Edwin, D. *Rapid Commun. Mass Spectrom.* **2002**, 16, 1729-1736.

80. Baker, E. S.; Lee, J. T.; Sessler, J. L.; Bowers, M. T. *J. Am. Chem. Soc.* **2006**, 128, 2641-2648.

81. Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1999**, 13, 2489-2497.

82. Loo, J. A. *Mass Spectrom. Rev.* **1997**, 16, 1-23.

83. Schwartz, B. L.; Bruce, J. E.; Anderson, G. A.; Hofstadler, S. A.; Rockwood, A. L.; Smith, R. D.; Chilkoti, A.; Stayton, P. S. *J. Am. Soc. Mass Spectrom.* **1995**, 6, 459-465.

84. Hu, P.; Ye, Q.-Z.; Loo, J. A. *Anal. Chem.* **1994**, 66, 4190-4194.

85. Loo, J. A.; Hu, P. F.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. *J. Am. Soc. Mass Spectrom.* **1997**, 8, 234-243.

86. Rostom, A. A.; Robinson, C. V. *Curr. Opin. Struct. Biol.* **1999**, 9, 135-141.

87. Greig, M. J.; Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. *J. Am. Chem. Soc.* **1995**, 117, 10765-10766.

88. Cheng, X. H.; Harms, A. C.; Goudreau, P. N.; Terwilliger, T. C.; Smith, R. D. *Proc. Natl Acad. Sci. USA* **1996**, 93, 7022-7027.

89. Potier, N.; Donald, L. J.; Chernushevich, I.; Ayed, A.; Ens, W.; Arrowsmith, C. H.; Standing, K. G.; Duckworth, H. W. *Protein Sci.* **1998**, 7, 1388-1395.

90. SannesLowery, K. A.; Hu, P. F.; Mack, D. P.; Mei, H. Y.; Loo, J. A. *Anal. Chem.* **1997**, 69, 5130-5135.

91. Pramanik, B. N.; Bartner, P. L.; Mirza, U. A.; Liu, Y. H.; Ganguly, A. K. *J. Mass Spectrom.* **1998**, 33, 911-920.

92. McCammon, M. G.; Robinson, C. V. *Curr. Opin. Chem. Biol.* **2004**, 8, 60-65.

93. 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. *J. Biol. Chem.* **2002**, 277, 36425-36432.

94. Rogniaux, H.; Sanglier, S.; Strupat, K.; Azza, S.; Roitel, O.; Ball, V.; Tritsch, D.; Branlant, G.; Van Dorsselaer, A. *Anal. Biochem.* **2001**, 291, 48-61.

95. Drummond, J. T.; Loo, R. R. O.; Matthews, R. G. *Biochemistry* **1993**, 32, 9282-9289.

96. Xu, N.; Pasa-Tolic, L.; Smith, R. D.; Ni, S.; Thrall, B. D. *Anal. Biochem.* **1999**, 272, 26-33.

97. Hensley, P. *Structure* **1996**, 4, 367-373.

98. Otto-Bruc, A.; Antonny, B.; Minh Vuong, T.; Chardin, P.; Chabre, M. *Biochemistry* **1993**, 32, 8636-8645.

99. Chen, H.; Sanyal, G. *J. Pharm. Sci.* **1999**, 88, 1122-1126.

100. Yuan, T.; Walsh, M. P.; Sutherland, C.; Fabian, H.; Vogel, H. J. *Biochemistry* **1999**, 38, 1446-1455.

101. Lebowitz, J.; Lewis, M. S.; Schuck, P. *Protein Sci.* **2002**, 11, 2067-2079.

102. Winzor, D. J.; Harding, S. E. *Protein-Ligand Interactions: Hydrodynamics and Calorimetry* **2001**, 105-135.

103. Wills, P. R.; Jacobsen, M. P.; Winzor, D. J. *Prog. Colloid Polym. Sci.* **1997**, 107, 1-10.

104. Szabo, A.; Stolz, L.; Granzow, R. *Curr. Opin. Struct. Biol.* **1995**, 5, 699-705.

105. Traenkle, J. H.; Herrmann, C. *Protein-Protein Interactions (2nd Edition)* **2005**, 229-251.

106. Fanghaenel, J.; Wawra, S.; Luecke, C.; Wildemann, D.; Fischer, G. *Anal. Chem.* **2006**, In press.

107. Barcelo, F.; Capo, D.; Portugal, J. *Nucleic Acids Res.* **2002**, 30, 4567-4573.

108. Leavitt, S.; Freire, E. *Curr. Opin. Struct. Biol.* **2001**, 11, 560-566.

109. Jelesarov, I.; Bosshard, H. R. *J. Mol. Recognit.* **1999**, 12, 3-18.

110. Aivaliotis, M.; Karas, M.; Tsiotis, G. *Photosynth. Res.* **2006**, 88, 143-157.

111. Fethiere, J.; Venzke, D.; Madden, D. R.; Boettcher, B. *Biochemistry* **2005**, 44, 15906-15914.

112. Kitta, K.; Ohnishi-Kameyama, M.; Moriyama, T.; Ogawa, T.; Kawamoto, S. *Anal. Biochem.* **2006**, 351, 290-297.

113. Monie, T. P.; Hernandez, H.; Robinson, C. V.; Simpson, P.; Matthews, S.; Curry, S. *RNA* **2005**, 11, 1803-1808.

114. Moxley, R. A.; Jarrett, H. W.; Mitra, S. *J.Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2003**, 797, 269-288.

115. Peterson, G. L.; Hokin, L. E. *J. Biol. Chem.* **1981**, 256, 3751-3761.

116. Stanton, P. *Methods Mol. Biol* **2004**, 251, 55-73.

117. Wang, W.; Wei, L.; Wang, G. *J. Chromatogr. Sci.* **2003**, 41, 489-493.

118. Glusker, J. P.; Lewis, M.; Rossi, M. *Crystal Structure Analysis for Chemists and Biologists*, VCH Publishers, New York, **1994**.

119. Drenth, J. *Principles of Protein X-Ray Crystallography, 2nd Edition*, Springer-Verlag Inc., New York, **1999**.

120. Brunger, A. T. *Nat. Struct. Biol* **1997**, 4, 862-865.

121. Smyth, M. S.; Martin, J. H. J. *J. Clin. Pathol.-Mol Pathol.* **2000**, 53, 8-14.

122. Gadek, T. R. *BioTechniques* **2003**, Suppl, 21-24.

123. Smith, D. L.; Deng, Y. Z.; Zhang, Z. Q. *J. Mass Spectrom.* **1997**, 32, 135-146.

124. Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, 113, 7818-7819.

125. Ganem, B.; Li, Y. T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, 113, 6294-6296.

126. Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1991**, 113, 8534-8535.

127. Smith, R. D.; Bruce, J. E.; Wu, Q.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, 26, 191-202.

128. Beck, J. L.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, 20, 61-87.

129. Hernandez, H.; Robinson, C. V. *J. Biol. Chem.* **2001**, 276, 46685-46688.

130. Veenstra, T. D. *Biophys. Chem.* **1999**, 79, 63-79.

131. Sobott, F.; McCammon, M. G.; Hernandez, H.; Robinson, C. V. *Philos. Trans. R. Soc. London, Ser. A: Math., Phys. Eng. Sci.* **2005**, 363, 379-391.

132. Kansy, J. W.; Clack, B. A.; Gray, D. M. *J. Biomol. Struct. Dyn.* **1986**, 3, 1079-1110.

133. Veenstra, T. D. *Biochem. Biophys. Res. Commun.* **1999**, 257, 1-5.

134. Veenstra, T. D.; Johnson, K. L.; Tomlinson, A. J.; Naylor, S.; Kumar, R. *Eur. Mass Spectrom.* **1997**, 3, 453-459.

135. Lafitte, D.; Capony, J. P.; Grassy, G.; Haiech, J.; Calas, B. *Biochemistry* **1995**, 34, 13825-13832.

136. Lafitte, D.; Heck, A. J. R.; Hill, T. J.; Jumel, K.; Harding, S. E.; Derrick, P. J. *Eur. J. Biochem.* **1999**, 261, 337-344.

137. Chazin, W.; Veenstra, T. D. *Rapid Commun. Mass Spectrom.* **1999**, 13, 548-555.

138. Akerfeldt, K. S.; Coyne, A. N.; Wilk, R. R.; Thulin, E.; Linse, S. *Biochemistry* **1996**, 35, 3662-3669.

139. Veenstra, T. D.; Gross, M. D.; Hunziker, W.; Kumar, R. *J. Biol. Chem.* **1995**, 270, 30353-30358.

140. Gross, M. D.; Kumar, R.; Hunziker, W. *J. Biol. Chem.* **1988**, 263, 14426-14432.

141. Veenstra, T. D.; Johnson, K. L.; Tomlinson, A. J.; Naylor, S.; Kumar, R. *Biochemistry* **1997**, 36, 3535-3542.

142. Feng, R.; Castelhano, A. L.; Billedeau, R.; Yuan, Z. *J. Am. Soc. Mass Spectrom.* **1995**, 6, 1105-1111.

143. Veenstra, T. D.; Johnson, k. L.; Tomlinson, A. J.; Craig, T. A.; Kumar, R.; Naylor, S. *J. Am. Soc. Mass Spectrom.* **1998**, 9, 8-14.

144. Greig, M. J.; Gaus, H. J.; Griffey, R. H. *Rapid Commun. Mass Spectrom.* **1996**, 10, 47-50.

145. Smith, R. D.; Loo, J. A.; Loo, R. R. O.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, 10, 359-451.

146. Colgrave, M. L.; Beck, J. L.; Sheil, M. M.; Searle, M. S. *Chem. Commun.* **2002**, 556-557.

147. Mathews, C. K.; van Holde, K. E.; Ahern, K. G. *Biochemistry*, The Benjamin/Cummings Publising Company, Redwood City, CA, USA, **1990**.

148. Sinden, R. R. *DNA Structure and Function*, Academic Press, San Diego, **1994**.

149. Przybylski, M. *Natural Product Analysis: Chromatography, Spectroscopy, Biological Testing, [Symposium], Wuerzburg, Germany, Sept. 1997* **1998**, 153-166.

150. Doktycz, M. J.; Habibi-Goudarzi, S.; McLuckey, S. A. *Anal. Chem.* **1994**, 66, 3416-3422.

151. Bayer, E.; Bauer, T.; Schmeer, K.; Bleicher, K.; Maler, M.; Gaus, H. J. *Anal. Chem.* **1994**, 66, 3858-3863.

152. Ding, J. M.; Anderegg, R. J. *J. Am. Soc. Mass Spectrom.* **1995**, 6, 159-164.

153. Gabelica, V.; Rosu, F.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, 14, 464-467.

154. Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, 120, 9605-9613.

155. Gabelica, V.; De Pauw, E. *J. Mass Spectrom.* **2001**, 36, 397-402.

156. Gabelica, V.; De Pauw, E. *Intl J. Mass Spectrom.***2002**, 219, 151-159.

157. Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, 13, 91-98.

158. Roche, C. J.; Thomson, J. A.; Crothers, D. M. *Biochemistry* **1994**, 33, 926-935.

159. Johnson, R. K.; Zee-Cheng, R. K.; Lee, W. W.; Acton, E. M.; Henry, D. W.; Cheng, C. C. *Cancer Treatment Rep.* **1979**, 63, 425-439.

160. Quigley, G. J.; Wang, A. H.; Ughetto, G.; van der Marel, G.; van Boom, J. H.; Rich, A. *Proc. Natl Acad. Sci. USA* **1980**, 77, 7204-7208.

161. Frederick, C. A.; Williams, L. D.; Ughetto, G.; Van der Marel, G. A.; Van Boom, J. H.; Rich, A.; Wang, A. H. J. *Biochemistry* **1990**, 29, 2538-2549.

162. Cao, Y.; He, X.-W. *Spectrochim. Acta, Part A: Mol. Biomol. Spectrosc.* **1998**, 54A, 883-892.

163. Gale, D. C.; Goodlett, D. R.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, 116, 6027-6028.

164. Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, 6, 1154-1164.

165. Wan, K. X.; Shibue, T.; Gross, M. L. *J. Am. Chem. Soc.* **2000**, 122, 300-307.

166. Baca, M.; Kent, S. B. H. *J. Am. Chem. Soc.* **1992**, 114, 3992-3993.

167. Fabris, D.; Fenselau, C. *Anal. Chem.* **1999**, 71, 384-387.

168. Li, Y. T.; Hsieh, Y. L.; Henion, J. D.; Senko, M. W.; McLafferty, F. W.; Ganem, B. *J. Am. Chem. Soc.* **1993**, 115, 8409-8413.

169. Ayed, A.; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. *Rapid Commun. Mass Spectrom.* **1998**, 12.

170. Lei, Q. P.; Cui, X. Y.; Kurtz, D. M.; Amster, I. J.; Chernushevich, I. V.; Standing, K. G. *Anal. Chem.* **1998**, 70, 1838-1846.

171. Wickham, G.; Iannitti, P.; Boschenok, J.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **1995**, S197-203.

172. http://www.basic.northwestern.edu/biotools/oligocalc.html.

173. Hamdan, S.; Bulloch, E. M.; Thompson, P. R.; Beck, J. L.; Yang, J. Y.; Crowther, J. A.; Lilley, P. E.; Carr, P. D.; Ollis, D. L.; Brown, S. E.; Dixon, N. E. *Biochemistry* **2002**, 41, 5266-5275.

174. Williams, N. K.; Prosselkov, P.; Liepinsh, E.; Line, I.; Sharipo, A.; Littler, D. R.; Curmi, P. M. G.; Otting, G.; Dixon, N. E. *J. Biol. Chem.* **2002**, 277, 7790-7798.

175. Galletto, R.; Jezewska, M. J.; Bujalowski, W. *J. Mol. Biol.* **2003**, 329, 441-465.

176. Keniry, M. A.; Berthon, H. A.; Yang, J. Y.; Miles, C. S.; Dixon, N. E. *Protein Sci.* **2000**, 9, 721-733.

177. Hamdan, S.; Brown, S. E.; Thompson, P. R.; Yang, J. Y.; Carr, P. D.; Ollis, D. L.; Otting, G.; Dixon, N. E. *J. Struct. Biol.* **2000**, 131, 164-169.

178. Funnell, B. E.; Baker, T. A.; Kornberg, A. *J. Biol. Chem.* **1987**, 262, 10327-10334.

179. Jezewska, M. J.; Kim, U. S.; Bujalowski, W. *Biophys. J.* **1996**, 71, 2075-2086.

180. LeBowitz, J. H.; McMacken, R. *J. Biol. Chem.* **1986**, 261, 4738-4748.

181. Mathews, C. K.; van Holde, K. E.; Ahern, K. G. *Biochemistry*, The Benjamin/Cummings Publising Company, Redwood City, CA, USA, **2000**.

182. http://fig.cox.miami.edu/~cmallery/150/gene/16x5bc.jpg.

183. Watson, J. D.; Crick, F. H. *Nature* **1953**, 171, 737-738.

184. Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamental of Biochemistry*, John Wiley & Son, Inc., USA, **1999**.

185. Saenger, W.; Hunter, W. N.; Kennard, O. *Nature* **1986**, 324, 385-388.

186. Scarlett, G. P.; Elgar, S. J.; Cary, P. D.; Noble, A. M.; Orford, R. L.; Kneale, G. G.; Guille, M. J. *J. Biol. Chem.* **2004**, 279, 52447-52455.

187. http://lmb.uni-muenchen.de/groups/Biostruc/chap-08/dna_abz.gif.

188. Pohl, F. M. *Cold Spring Harbor Symposia on Quantitative Biology* **1983**, 47, 113-117.

189. Di Capua, E.; Stasiak, A.; Koller, T.; Brahms, S.; Thomae, R.; Pohl, F. M. *EMBO J.* **1983**, 2, 1531-1535.

190. Droege, P.; Pohl, F. M. *Nucleic Acids Res.* **1991**, 19, 5301-5306.

191. Delepierre, M.; Langlois D'Estaintot, B.; Igolen, J.; Roques, B. P. *Eur. J. Biochem.* **1986**, 161, 571-577.

192. Neidle, S.; Pearl, L. H.; Skelly, J. V. *Biochem. J.* **1987**, 243, 1-13, 11 plate.

193. Cohen, J. S. *Trends Biochem. Sci.* **1987**, 12, 133-135.

194. Long, E. C.; Barton, J. K. *Acc. Chem. Res.* **1990**, 23, 271-273.

195. Pandolfi, P. P. *Oncogene* **2001**, 20, 3116-3127.

196. Peterson, M. G.; Baichwal, V. R. *Trends Biotechnol.* **1993**, 11, 11-18.

197. Darnell, J. E., Jr. *Nat. Rev. Cancer* **2002**, 2, 740-749.

198. Odom, D. T.; Parker, C. S.; Barton, J. K. *Biochemistry* **1999**, 38, 5155-5163.

199. Pullman, A.; Pullman, B. *Q. Rev. Biophys.* **1981**, 14, 289-380.

200. Rosenberg, B.; VanCamp, L.; Krigas, T. *Nature* **1965**, 205, 698-699.

201. di Pietro, A.; de Vries, E. G. E.; Gietema, J. A.; Spierings, D. C. J.; de Jong, S. *Intl J. Biochem. Cell Biol.* **2005**, 37, 2437-2456.

202. Schweyer, S.; Soruri, A.; Heintze, A.; Radzun, H. J.; Fayyazi, A. *Intl J. Oncol.* **2004**, 25, 1671-1676.

203. Markman, M. *Oncologist* **1996**, 1, 18-21.

204. Rothenberg, M. L.; Ozols, R. F.; Glatstein, E.; Steinberg, S. M.; Reed, E.; Young, R. C. *J. Clin. Oncol.* **1992**, 10, 727-734.

205. Weiss, R. B.; Christian, M. C. *Drugs* **1993**, 46, 360-377.

206. Taylor, S. G. t.; Murthy, A. K.; Caldarelli, D. D.; Showel, J. L.; Kiel, K.; Griem, K. L.; Mittal, B. B.; Kies, M.; Hutchinson, J. C., Jr.; Holinger, L. D. *J. Clin. Oncol.* **1989**, 7, 846-856.

207. Wittes, R. E.; Cvitkovic, E.; Shah, J.; Gerold, F. P.; Strong, E. W. *Cancer Treatment Rep.* **1977**, 61, 359-366.

208. Rosenberg, B. *Platinum Complexes for the Treatment of Cancer: Why the Search Goes On?*, Research Signpost, India, **2005**.

209. Bloemink, M. J.; Reedijk, J. *Met. Ions Biol. Sys.* **1996**, 32, 641-685.

210. Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, 99, 2467-2498.

211. Poklar, N.; Pilch, D. S.; Lippard, S. J.; Redding, E. A.; Dunham, S. U.; Breslauer, K. J. *Proc. Natl Acad. Sci. USA* **1996**, 93, 7606-7611.

212. Rajski, S. R.; Williams, R. M. *Chem. Rev.* **1998**, 98, 2723-2795.

213. Lawley, P. D.; Brookes, P. *Nature* **1965**, 206, 480-483.

214. Lawley, P. D.; Brookes, P. *J. Mol. Biol.* **1967**, 25, 143-160.

215. Osborne, M. R.; Lawley, P. D. *Chem.-Biol. Interact.* **1993**, 89, 49-60.

216. Lawley, P. D.; Phillips, D. H. *Mutat. Res.* **1996**, 355, 13-40.

217. Rink, S. M.; Solomon, M. S.; Taylor, M. J.; Rajur, S. B.; McLaughlin, L. W.; Hopkins, P. B. *J. Am. Chem. Soc.* **1993**, 115, 2551-2557.

218. Wyatt, M. D.; Lee, M.; Garbiras, B. J.; Souhami, R. L.; Hartley, J. A. *Biochemistry* **1995**, 34, 13034-13041.

219. Hemminki, K. *Chem.-Biol. Interact.* **1987**, 61, 75-88.

220. Sequin, U. *Prog. Chem. Org. Nat. Prod.* **1986**, 50, 57-122.

221. Pavlopoulos, S.; Bicknell, W.; Craik, D. J.; Wickham, G. *Biochemistry* **1996**, 35, 9314-9324.

222. Hansen, M.; Yun, S.; Hurley, L. *Chem. Biol.* **1995**, 2, 229-240.

223. Owen, E. A.; Burley, G. A.; Carver, J. A.; Wickham, G.; Keniry, M. A. *Biochem. Biophys. Res. Commun.* **2002**, 290, 1602-1608.

224. Colgrave, M. L.; Iannitti-Tito, P.; Wickham, G.; Sheil, M. M. *Aust. J. Chem.* **2003**, 56, 401-413.

225. Iannitti, P.; Sheil, M. M.; Wickham, G. *J. Am. Chem. Soc.* **1997**, 119, 1490-1491.

226. Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chem. Biol.*; IRL Press, Oxford, UK, 1990, 297-336.

227. Uytterhoeven, K.; Sponer, J.; Van Meervelt, L. *Eur. J. Biochem.* **2002**, 269, 2868-2877.

228. Coll, M.; Frederick, C. A.; Wang, A. H.; Rich, A. *Proc. Natl Acad. Sci. USA* **1987**, 84, 8385-8389.

229. Geierstanger, B. H.; Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, 24, 463-493.

230. Ren, J.; Chaires, J. B. *Biochemistry* **1999**, 38, 16067-16075.

231. Pelton, J. G.; Wemmer, D. E. *Biochemistry* **1988**, 27, 8088-8096.

232. Pelton, J. G.; Wemmer, D. E. *Proc. Natl Acad. Sci. USA* **1989**, 86, 5723-5727.

233. Loontiens, F. G.; Regenfuss, P.; Zechel, A.; Dumortier, L.; Clegg, R. M. *Biochemistry* **1990**, 29, 9029-9039.

234. Patel, D. J. *Proc. Natl Acad. Sci. USA* **1982**, 79, 6424-6428.

235. Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl Acad. Sci. USA* **1985**, 82, 1376-1380.

236. Temperini, C.; Messori, L.; Orioli, P.; Di Bugno, C.; Animati, F.; Ughetto, G. *Nucleic Acids Res.* **2003**, 31, 1464-1469.

237. Sarma, R. H. *Nucleic Acid Geometry and Dynamics*, Pergamon Press, New York, **1980**.

238. Bailly, C.; Echepare, S.; Gago, F.; Waring, M. J. *Anti-Cancer Drug Design* **1999**, 14, 291-303.

239. Pullman, B. *Anti-Cancer Drug Des.* **1991**, 6, 95-105.

240. Lerman, L. S. *J. Mol. Biol.* **1961**, 3, 18-30.

241. Wang, A. H.; Nathans, J.; van der Marel, G.; van Boom, J. H.; Rich, A. *Nature* **1978**, 276, 471-474.

242. Wong, Y.-S.; Lippard, S. J. *J. Chem. Soc., Chem. Commun.* **1977**, 824-825.

243. Neidle, S.; Taylor, G. L. *FEBS Lett.* **1979**, 107, 348-354.

244. Neidle, S. *Prog. Med. Chem.* **1979**, 16, 151-221.

245. Roche, C. J.; Berkowitz, D.; Sulikowski, G. A.; Danishefsky, S. J.; Crothers, D. M. *Biochemistry* **1994**, 33, 936-942.

246. Chaires, J. B.; Dattagupta, N.; Crothers, D. M. *Biochemistry* **1982**, 21, 3933-3940.

247. Taatjes, D. J.; Gaudiano, G.; Resing, K.; Koch, T. H. *J. Med. Chem.* **1996**, 39, 4135-4138.

248. Chaires, J. B.; Herrera, J. E.; Waring, M. J. *Biochemistry* **1990**, 29, 6145-6153.

249. Clarke, M. J.; Zhu, F.; Frasca, D. R. *Chem. Rev.* **1999**, 99, 2511-2533.

250. Erkkila, K. E.; Odom, D. T.; Barton, J. K. *Chem. Rev.* **1999**, 99, 2777-2795.

251. Hanlon, S.; Chan, A.; Berman, S. *Biochim. Biophys. Acta, Nucleis Acids Protein Synth.* **1978**, 519, 526-536.

252. Barton, J. K. *J. Biomol. Struct. Dyn.* **1983**, 1, 621-632.

253. Barton, J. K.; Danishefsky, A.; Goldberg, J. *J. Am. Chem. Soc.* **1984**, 106, 2172-2176.

254. Jennette, K. W.; Lippard, S. J.; Vassiliades, G. A.; Bauer, W. R. *Proc. Natl Acad. Sci. USA* **1974**, 71, 3839-3843.

255. Yamagishi, A. *Chem. Commun.* **1983**, 572-573.

256. Kumar, C. V.; Barton, J. K.; Turro, N. J. *J. Am. Chem. Soc.* **1985**, 107, 5518-5523.

257. Barton, J. K.; Dannenberg, J. J.; Raphael, A. L. *J. Am. Chem. Soc.* **1982**, 104, 4967-4969.

258. Rehmann, J. P.; Barton, J. K. *Biochemistry* **1990**, 29, 1701-1709.

259. Sutin, N.; Creutz, C. *Pure Appl. Chem.* **1980**, 52, 2717-2738.

260. Balzani, V.; Moggi, L.; Manfrin, M. F.; Bolletta, F.; Laurence, G. S. *Coord. Chem. Rev.* **1975**, 15, 321-433.

261. Barton, J. K.; Goldberg, J. M.; Kumar, C. V.; Turro, N. J. *J. Am. Chem. Soc.* **1986**, 108, 2081-2088.

262. Satyanarayana, S.; Dabrowiak, J. C.; Chaires, J. B. *Biochemistry* **1992**, 31, 9319-9324.

263. Satyanarayana, S.; Dabrowiak, J. C.; Chaires, J. B. *Biochemistry* **1993**, 32, 2573-2584.

264. Eriksson, M.; Leijon, M.; Hiort, C.; Norden, B.; Graeslund, A. *Biochemistry* **1994**, 33, 5031-5040.

265. Norden, B.; Patel, N.; Hiort, C.; Graeslund, A.; Kim, S. K. *Nucleosides Nucleotides* **1991**, 10, 195-205.

266. Eriksson, M.; Leijon, M.; Hiort, C.; Norden, B.; Graeslund, A. *J. Am. Chem. Soc.* **1992**, 114, 4933-4934.

267. Rehmann, J. P.; Barton, J. K. *Biochemistry* **1990**, 29, 1710-1717.

268. Coggan, D. Z. M.; Haworth, I. S.; Bates, P. J.; Robinson, A.; Rodger, A. *Inorg. Chem.* **1999**, 38, 4486-4497.

269. Collins, J. G.; Sleeman, A. D.; Aldrich-Wright, J. R.; Greguric, I.; Hambley, T. W. *Inorg. Chem.* **1998**, 37, 3133-3141.

270. Greguric, I.; Aldrich-Wright, J. R.; Collins, J. G. *J. Am. Chem. Soc.* **1997**, 119, 3621-3622.

271. Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1994**, 116, 10286-10287.

272. Dupureur, C. M.; Barton, J. K. *Inorg. Chem.* **1997**, 36, 33-43.

273. Holmlin, R. E.; Stemp, E. D. A.; Barton, J. K. *Inorg. Chem.* **1998**, 37, 29-34.

274. Delaney, S.; Pascaly, M.; Bhattacharya, P. K.; Han, K.; Barton, J. K. *Inorg. Chem.* **2002**, 41, 1966-1974.

275. Rueba, E.; Hart, J. R.; Barton, J. K. *Inorg. Chem.* **2004**, 43, 4570-4578.

276. Murphy, C. J.; Arkin, M. R.; Ghatlia, N. D.; Bossmann, S.; Turro, N. J.; Barton, J. K. *Proc. Natl Acad. Sci. USA* **1994**, 91, 5315-5319.

277. Friedman, A. E.; Chambron, J. C.; Sauvage, J. P.; Turro, N. J.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, 112, 4960-4962.

278. Olson, E. J. C.; Hu, D.; Hoermann, A.; Jonkman, A. M.; Arkin, M. R.; Stemp, E. D. A.; Barton, J. K.; Barbara, P. F. *J. Am. Chem. Soc.* **1997**, 119, 11458-11467.

279. Turro, C.; Bossmann, S. H.; Jenkins, Y.; Barton, J. K.; Turro, N. J. *J. Am. Chem. Soc.* **1995**, 117, 9026-9032.

280. Arkin, M. R.; Stemp, E. D. A.; Turro, C.; Turro, N. J.; Barton, J. K. *J. Am. Chem. Soc.* **1996**, 118, 2267-2274.

281. Chow, C. S.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, 112, 2839-2841.

282. Sitlani, A.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1993**, 115, 12589-12590.

283. Fitzsimons, M. P.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, 119, 3379-3380.

284. Hall, D. B.; Kelley, S. O.; Barton, J. K. *Biochemistry* **1998**, 37, 15933-15940.

285. Dandliker, P. J.; Holmlin, R. E.; Barton, J. K. *Science* **1997**, 275, 1465-1468.

286. Dandliker, P. J.; Nunez, M. E.; Barton, J. K. *Biochemistry* **1998**, 37, 6491-6502.

287. Hudson, B. P.; Dupureur, C. M.; Barton, J. K. *J. Am. Chem. Soc.* **1995**, 117, 9379-9380.

288. Nordhoff, E.; Kirpekar, F.; Roepstorff, P. *Mass Spectrom. Rev.* **1997**, 15, 69-138.

289. Yates, J. R., III. *J. Mass Spectrom.* **1998**, 33, 1-19.

290. Gao, Q.; Cheng, X.; Smith, R. D.; Yang, C. F.; Goldberg, I. H. *J. Mass Spectrom.* **1996**, 31, 31-36.

291. Triolo, A.; Arcamone, F. M.; Raffaelli, A.; Salvadori, P. *J. Mass Spectrom.* **1997**, 32, 1186-1194.

292. Pyle, A. M.; Rehmann, J. P.; Meshoyrer, R.; Kumar, C. V.; Turro, N. J.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, 111, 3051-3058.

293. Aldrich-Wright, J.; Brodie, C.; Glazer, E. C.; Luedtke, N. W.; Elson-Schwab, L.; Tor, Y. *Chem. Commun.* **2004**, 1018-1019.

294. Beck, J. L.; Gupta, R.; Urathamakul, T.; Williamson, N. L.; Sheil, M. M.; Aldrich-Wright, J. R.; Ralph, S. F. *Chem. Commun.* **2003**, 626-627.

295. Haq, I.; Lincoln, P.; Suh, D.; Norden, B.; Chowdhry, B. Z.; Chaires, J. B. *J. Am. Chem. Soc.* **1995**, 117, 4788-4796.

296. Chaires, J. B.; Fox, K. R.; Herrera, J. E.; Britt, M.; Waring, M. J. *Biochemistry* **1987**, 26, 8227-8236.

297. Wang, A. H. J.; Ughetto, G.; Quigley, G. J.; Rich, A. *Biochemistry* **1987**, 26, 1152-1163.

298. Chen, K. X.; Gresh, N.; Pullman, B. *J. Biomol. Struct. Dyn.* **1985**, 3, 445-466.

299. Bailly, C.; Henichart, J. P.; Colson, P.; Houssier, C. *J. Mol. Recognit.***1992**, 5, 155-171.

300. Kielkopf, C. L.; Erkkila, K. E.; Hudson, B. P.; Barton, J. K.; Rees, D. C. *Nat. Struct. Biol.***2000**, 7, 117-121.

301. Gupta, R.; Beck, J. L.; Ralph, S. F.; Sheil, M. M.; Aldrich-Wright, J. R. *J. Am. Soc. Mass Spectrom.* **2004**, 15, 1382-1391.

302. Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, 34, 1328-1337.

303. Gupta, R.; Kapur, A.; Beck, J. L.; Sheil, M. M. *Rapid Commun. Mass Spectrom.* **2001**, 15, 2472-2480.

304. Rao, S. N.; Kollman, P. A. *Proc. Natl Acad. Sci. USA* **1987**, 84, 5735-5739.

305. Kapur, A., *PhD Thesis*, University of Wollongong, **2003**

306. Jones, C. J. *D-and F-Block Chemistry*, Royal Society of Chemistry, Cambridge UK, **2001**.

307. Nikolis, N.; Methenitis, C.; Pneumatikakis, G. *J. Inorg. Biochem.* **2003**, 95, 177-193.

308. Ishida, T.; Mihara, Y.; Hama, Y.; Hanatani, A.; Tarui, M.; Doi, M.; Nakaike, S.; Kitamura, K. *Chem. Pharm. Bull.* **1998**, 46, 739-743.

309. Brana, M. F.; Cacho, M.; Gradillas, A.; De Pascual-Teresa, B.; Ramos, A. *Curr. Pharm. Des.* **2001**, 7, 1745-1780.

310. Lah, J.; Vesnaver, G. *Biochemistry* **2000**, 39, 9317-9326.

311. Quigley, G. J.; Wang, A. H. J.; Ughetto, G.; Van der Marel, G.; Van Boom, J. H.; Rich, A. *Proc. Natl Acad. Sci. USA* **1980**, 77, 7204-7208.

312. Browne, K. A.; He, G. X.; Bruice, T. C. *J. Am. Chem. Soc.* **1993**, 115, 7072-7079.

313. Xodo, L. E.; Manzini, G.; Ruggiero, J.; Quadrifoglio, F. *Biopolymers* **1988**, 27, 1839-1857.

314. Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, 9, 513-538.

315. Schaeffer, P. M.; Headlam, M. J.; Dixon, N. E. *IUBMB Life* **2005**, 57, 5-12.

316. Jeruzalmi, D.; O'Donnell, M.; Kuriyan, J. *Cell* **2001**, 106, 429-441.

317. Kong, X. P.; Onrust, R.; O'Donnell, M.; Kuriyan, J. *Cell* **1992**, 69, 425-437.

318. Gulbis, J. M.; Kazmirski, S. L.; Finkelstein, J.; Kelman, Z.; O'Donnell, M.; Kuriyan, J. *Eur. J. Biochem.* **2004**, 271, 439-449.

319. McHenry, C. S. *Mol. Microbiol.* **2003**, 49, 1157-1165.

320. Kelman, Z.; O'Donnell, M. *Annu. Rev. Biochem.* **1995**, 64, 171-200.

321. Taft-Benz, S. A.; Schaaper, R. M. *Nucleic Acids Res.* **1998**, 26, 4005-4011.

322. Schaaper, R. M. *J. Biol. Chem.* **1993**, 268, 23762-23765.

323. Perrino, F. W.; Harvey, S.; McNeill, S. M. *Biochemistry* **1999**, 38, 16001-16009.

324. Kuriyan, J.; O'Donnell, M. *J. Mol. Biol.* **1993**, 234, 915-925.

325. Krishna, T. S.; Kong, X. P.; Gary, S.; Burgers, P. M.; Kuriyan, J. *Cell* **1994**, 79, 1233-1243.

326. Davey, M. J.; O'Donnell, M. *Curr. Opin. Chem. Biol.* **2000**, 4, 581-586.

327. 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. *Biochimie* **2001**, 83, 5-12.

328. Messer, W. *FEMS Microbiol. Rev.* **2002**, 26, 355-374.

329. Frick, D. N.; Richardson, C. C. *Annu. Rev. Biochem.* **2001**, 70, 39-80.

330. Benkovic, S. J.; Valentine, A. M.; Salinas, F. *Annu. Rev. Biochem.* **2001**, 70, 181-208.

331. Kornberg, A.; Baker, T. *DNA Replication*, *2nd Ed.* W.H. Freeman & Company, New York, **1992**.

332. Hill, T. M. *Annu. Rev. Microbiol.* **1992**, 46, 603-633.

333. Kornberg, A.; Lehman, I. R.; Bessman, M. J.; Simms, E. S. *Biochim. Biophys. Acta* **1956**, 21, 197.

334. De Lucia, P.; Cairns, J. *Nature* **1969**, 224, 1164-1166.

335. McHenry, C. S. *Annu. Rev. Biochem.* **1988**, 57, 519-550.

336. Bryant; R, F.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1983**, 22, 3537-3546.

337. Studwell-Vaughan, P. S.; O'Donnell, M. *J. Biol. Chem.* **1990**, 265, 1171-1178.

338. Patel, P. H.; Suzuki, M.; Adman, E.; Shinkai, A.; Loeb, L. A. *J. Mol. Biol.* **2001**, 308, 823-837.

339. Delarue, M.; Poch, O.; Tordo, N.; Moras, D.; Argos, P. *Protein Eng.* **1990**, 3, 461-467.

340. Steitz, T. A. *J. Biol. Chem.* **1999**, 274, 17395-17398.

341. Friedberg, E. C.; Wagner, R.; Radman, M. *Science* **2002**, 296, 1627-1630.

342. Maki, H.; Maki, S.; Lasken, R. S.; Kornberg, A. *UCLA Symposia on Molecular and Cellular Biology, New Series* **1987**, 47, 63-73.

343. Kornberg, T.; Gefter, M. L. *J. Biol. Chem.* **1972**, 247, 5369-5375.

344. Maki, H.; Kornberg, A. *Proc. Natl Acad. Sci. USA* **1987**, 84, 4389-4392.

345. Jeruzalmi, D.; O'Donnell, M.; Kuriyan, J. *Curr. Opin. Struct. Biol.* **2002**, 12, 217-224.

346. Onrust, R.; Finkelstein, J.; Turner, J.; Naktinis, V.; O'Donnell, M. *J. Biol. Chem.* **1995**, 270, 13366-13377.

347. Maki, S.; Kornberg, A. *J. Biol. Chem.* **1988**, 263, 6561-6569.

348. Scheuermann, R. H.; Echols, H. *Proc. Natl Acad. Sci. USA* **1984**, 81, 7747-7751.

349. Studwell-Vaughan, P. S.; O'Donnell, M. *J. Biol. Chem.* **1993**, 268, 11785-11791.

350. Gupta, R.; Hamdan, S. M.; Dixon, N. E.; Sheil, M. M.; Beck, J. L. *Protein Sci.* **2004**, 13, 2878-2887.

351. DeRose, E. F.; Li, D.; Darden, T.; Harvey, S.; Perrino, F. W.; Schaaper, R. M.; London, R. E. *Biochemistry* **2002**, 41, 94-110.

352. Drake, J. W. *Annu. Rev. Genet.* **1991**, 25, 125-146.

353. Fersht, A. R.; Knill-Jones, J. W.; Tsui, W. C. *J. Mol. Biol.* **1982**, 156, 37-51.

354. Taft-Benz, S. A.; Schaaper, R. M. *J. Bacteriol.* **1999**, 181, 2963-2965.

355. Viswanathan, M.; Lovett, S. T. *J. Biol. Chem.* **1999**, 274, 30094-30100.

356. Blanco, L.; Bernard, A.; Salas, M. *Gene* **1992**, 112, 139-144.

357. Barnes, M. H.; Spacciapoli, P.; Li, D. H.; Brown, N. C. *Gene* **1995**, 165, 45-50.

358. Strauss, B. S.; Sagher, D.; Acharya, S. *Nucleic Acids Res.* **1997**, 25.

359. Koonin, E. V. *Nucleic Acids Res.* **1992**, 20, 1997.

360. Derbyshire, V.; Grindley, N. D. F.; Joyce, C. M. *EMBO J.* **1991**, 10, 17-24.

361. Beese, L. S.; Steitz, T. A. *EMBO J.* **1991**, 10, 25-33.

362. Hamdan, S.; Carr, P. D.; Brown, S. E.; Ollis, D. L.; Dixon, N. E. *Structure* **2002**, 10, 535-546.

363. Reedijk, J. *Springer Ser. Mat. Sci.* **2004**, 78, 244-257.

364. Garcia, J. S.; Schmidt de Magalhaes, C.; Arruda, M. A. Z. *Talanta* **2006**, 69, 1-15.

365. Liu, C.; Xu, H. *J. Inorg. Biochem.* **2002**, 88, 77-86.

366. Kennedy, M. L.; Gibney, B. R. *Curr. Opin. Struct. Biol.* **2001**, 11, 485-490.

367. Beck, J. L.; McConachie, L. A.; Summors, A. C.; Arnold, W. N.; De Jersey, J.; Zerner, B. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1986**, 869, 61-68.

368. Curley, J. F.; Joyce, C. M.; Piccirilli, J. A. *J. Am. Chem. Soc.* **1997**, 119, 12691-12692.

369. Frausto da Silva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*, Clarendon Press, Oxford, UK, **1991**.

370. Sigel, H., Ed. Metal Ions in Biological Systems: *Reactivity of Coordination Compounds,* Marcel Dekker Inc., New York, **1976**,

371. Dempsey, B. R.; Wrona, M.; Moulin, J. M.; Gloor, G. B.; Jalilehvand, F.; Lajoie, G.; Shaw, G. S.; Shilton, B. H. *Biochemistry* **2004**, 43, 9361-9371.

372. Berry, S. M.; Bebout, D. C.; Butcher, R. J. *Abstracts of Papers, 230th ACS National Meeting, Washington, DC, United States, Aug. 28-Sept. 1, 2005* **2005**, INOR-341.

373. Tsunaka, Y.; Takano, K.; Matsumura, H.; Yamagata, Y.; Kanaya, S. *J. Mol. Biol.* **2005**, 345, 1171-1183.

374. Raushel, F. M.; Rawding, C. J.; Anderson, P. M.; Villafranca, J. J. *Biochemistry* **1979**, 18, 5562-5566.

375. Hannan, J. P.; Whittaker, S. B. M.; Hemmings, A. M.; James, R.; Kleanthous, C.; Moore, G. R. *J. Inorg. Biochem.* **2000**, 79, 365-370.

376. Blake, P. R.; Lee, B.; Summers, M. F.; Park, J. B.; Zhou, Z. H.; Adams, M. W. W. *New J. Chem.* **1994**, 18, 387-395.

377. Herberg, F. W.; Doyle, M. L.; Cox, S.; Taylor, S. S. *Biochemistry* **1999**, 38, 6352-6360.

378. Frey, C. M.; Stuehr, J. E. *J. Am. Chem. Soc.* **1978**, 100, 134-139.

379. Conlan, L. H.; Dupureur, C. M. *Biochemistry* **2002**, 41, 1335-1342.

380. Ono, T.-A.; Mino, H. *Biochemistry* **1999**, 38, 8778-8785.

381. Mullen, G. P.; Serpersu, E. H.; Ferrin, L. J.; Loeb, L. A.; Mildvan, A. S. *J. Biol. Chem.* **1990**, 265, 14327-14334.

382. Bogumil, R.; Kappl, R.; Huttermann, J. *Met. Ions Biol. Sys.* **2000**, 37, 365-405.

383. Reed, G. H.; Poyner, R. R. *Met. Ions Biol. Sys.* **2000**, 37, 183-207.

384. Freeman, A. D. J.; Declais, A.-C.; Lilley, D. M. J. *J. Mol. Biol.* **2003**, 333, 59-73.

385. van den Bremer, E. T. J.; Jiskoot, W.; James, R.; Moore, G. R.; Kleanthous, C.; Heck, A. J. R.; Maier, C. S. *Protein Sci.* **2002**, 11, 1738-1752.

386. Bisaillon, M.; Bougie, I. *J. Biol. Chem.* **2003**, 278, 33963-33971.

387. Frey, M. W.; Frey, S. T.; Horrocks, W. D., Jr.; Kaboord, B. F.; Benkovic, S. J. *Chem. Biol.* **1996**, 3, 393-403.

388. Darnall, D. W.; Birnbaum, E. R. *J. Biol. Chem.* **1970**, 245, 6484-6486.

389. Darnall, D. W.; Birnbaum, E. R. *Biochemistry* **1973**, 12, 3489-3491.

390. Smolka, G. E.; Birnbaum, E. R.; Darnall, D. W. *Biochemistry* **1971**, 10, 4556-4561.

391. Horrocks, W. D., Jr.; Sudnick, D. R. *J. Am. Chem. Soc.* **1979**, 101, 334-340.

392. Ye, Y.; Lee, H.-W.; Yang, W.; Yang, J. J. *J. Inorg. Biochem.* **2005**, 99, 1376-1383.

393. Ming, L.-J. *NATO ASI Ser., Ser. C* **1995**, 457, 245-264.

394. Atreya, H. S.; Mukherjee, S.; Chary, K. V. R.; Lee, Y.-M.; Luchinat, C. *Protein Sci.* **2003**, 12, 412-425.

395. Kemple, M. D.; Ray, B. D.; Lipkowitz, K. B.; Prendergast, F. G.; Rao, B. D. N. *J. Am. Chem. Soc.* **1988**, 110, 8275-8287.

396. Pintacuda, G.; Keniry, M. A.; Huber, T.; Park, A. Y.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2004**, 126, 2963-2970.

397. Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A. *Nature* **1985**, 313, 762-766.

398. Derbyshire, V.; Freemont, P. S.; Sanderson, M. R.; Beese, L.; Friedman, J. M.; Joyce, C. M.; Steitz, T. A. *Science* **1988**, 240, 199-201.

399. Steitz, T. A. *Curr. Opin. Struct. Biol.* **1993**, 3, 31-38.

400. Brautigam, C. A.; Steitz, T. A. *J. Mol. Biol.* **1998**, 277, 363-377.

401. Kim, E. E.; Wyckoff, H. W. *J. Mol. Biol.* **1991**, 218, 449-464.

402. Davies, J. F., II; Hostomska, Z.; Hostomsky, Z.; Jordan, S. R.; Matthews, D. A. *Science* **1991**, 252, 88-95.

403. Han, H.; Rifkind, J. M.; Mildvan, A. S. *Biochemistry* **1991**, 30, 11104-11108.

404. Potier, N.; Rogniaux, H.; Chevreux, G.; Van Dorsselaer, A. *Methods Enzymol.* **2005**, 402, 361-389.

405. De Vriendt, K.; Van Driessche, G.; Devreese, B.; Bebrone, C.; Anne, C.; Frere, J.-M.; Galleni, M.; Van Beeumen, J. *J. Am. Soc. Mass Spectrom.* **2006**, 17, 180-188.

406. Zhu, M. M.; Rempel, D. L.; Du, Z.; Gross, M. L. *J. Am. Chem. Soc.* **2003**, 125, 5252-5253.

407. Zhu, M. M.; Rempel, D. L.; Zhao, J.; Giblin, D. E.; Gross, M. L. *Biochemistry* **2003**, 42, 15388-15397.

408. Palumaa, P.; Tammiste, I.; Kruusel, K.; Kangur, L.; Joernvall, H.; Sillard, R. *Biochim. Biophys. Acta, Protein Proteomics* **2005**, 1747, 205-211.

409. Urvoas, A.; Amekraz, B.; Moulin, C.; Le Clainche, L.; Stoecklin, R.; Moutiez, M. *Rapid Commun. Mass Spectrom.* **2003**, 17, 1889-1896.

410. Hershfinkel, M. *Topics Curr. Genet.* **2006**, 14, 131-153.

411. Wu, Q.; Cheng, X.; Hofstadler, S. A.; Smith, R. D. *J. Mass Spectrom.* **1996**, 31, 669-675.

412. Li, Y. T.; Hsieh, Y. L.; Henion, J. D.; Ganem, B. *J. Am. Soc. Mass Spectrom.* **1993**, 4, 631-637.

413. Light-Wahl, K. J.; Loo, J. A.; Edmonds, C. G.; Smith, R. D.; Witkowska, H. E.; Shackleton, C. H. L.; Wu, C. S. *Biol. Mass Spectrom.* **1993**, 22, 112-120.

414. Smith, R. D.; Light-Wahl, K. J. *Biol. Mass Spectrom.* **1993**, 22, 493-501.

415. Lee, M. H.; Nowak, T. *Biochemistry* **1984**, 23, 6506-6513.

416. Hung, H. C.; Chang, G. G.; Yang, Z. R.; Tong, L. *Biochemistry* **2000**, 39, 14095-14102.

417. Miller, A.-F.; Brudvig, G. W. *Biochemistry* **1989**, 28, 8181-8190.

418. Jackson, C. M.; Brenckle, G. M.; Hogg, P. J.; Winzor, D. J. *J. Biol. Chem.* **1987**, 262, 13472-13475.

419. Jorgensen, T. J. D.; Roepstorff, P.; Heck, A. J. R. *Anal. Chem.* **1998**, 70, 4427-4432.

420. De Vriendt, K.; Sandra, K.; Desmet, T.; Nerinckx, W.; Van Beeumen, J.; Devreese, B. *Rapid Commun. Mass Spectrom.* **2004**, 18, 3061-3067.

421. Akashi, S.; Osawa, R.; Nishimura, Y. *J. Am. Soc. Mass Spectrom.* **2005**, 16, 116-125.

422. Biligh, S. W.; Haley, T.; Lowe, P. N. *J. Mol. Recognit.***2003**, 16, 139-148.

423. Housecroft, C. E.; Sharpe, A. G. *Inorganic Chemistry*, *2nd Ed.* Pearson Prentice Hall, England, **2005**.

424. Brautigam, C. A.; Aschheim, K.; Steitz, T. A. *Chem. Biol.* **1999**, 6, 901-908.

425. Banerjee, S.; Wei, B.; Bhattacharyya-Pakrasi, M.; Pakrasi, H. B.; Smith, T. J. *J. Mol. Biol.* **2003**, 333, 1061-1069.

426. Bernhard, S. A.; Gutfreund, H. *Prog. Biophys. Biophys. Chem.* **1960**, 10, 116-152.

427. Fersht, A. *Enzyme Structure and Mechanism*, W.H. Freeman & Company, New York, **1977**.

428. Lohman, T. M.; Bjornson, K. P. *Annu. Rev. Biochem.* **1996**, 65, 169-214.

429. Egelman, E. H. *J. Struct. Biol.* **1998**, 124, 123-128.

430. Patel, S. S.; Picha, K. M. *Annu. Rev. Biochem.* **2000**, 69, 651-697.

431. Lohman, T. M.; Bujalowski, W. *Biochemistry* **1994**, 33, 6167-6176.

432. Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *Biochemistry* **1998**, 37, 3116-3136.

433. Ellis, N. A. *Curr. Opin. Genet. Dev.* **1997**, 7, 354-363.

434. Matson, S. W. *Prog. Nucleic Acids Res. Mol. Biol.* **1991**.

435. Baker, T. A.; Funnell, B. E.; Kornberg, A. *J. Biol. Chem.* **1987**, 262, 6877-6885.

436. McMacken, R.; Wold, M. S.; LeBowitz, J. H.; Roberts, J. D.; Mallory, J. B.; Wilkinson, J. A. K.; Loehrlein, C. *UCLA Symposia on Molecular and Cellular Biology, New Series* **1983**, 10, 819-848.

437. Baker, T. A.; Sekimizu, K.; Funnell, B. E.; Kornberg, A. *Cell* **1986**, 45, 53-64.

438. Kim, S.; Dallmann, H. G.; McHenry, C. S.; Marians, K. J. *J. Biol. Chem.* **1996**, 271, 21406-21412.

439. Manna, A. C.; Pai, K. S.; Bussiere, D. E.; Davies, C.; White, S. W.; Bastia, D. *Cell* **1996**, 87, 881-891.

440. Arai, K.; Kornberg, A. *J. Biol. Chem.* **1981**, 256, 5260-5266.

441. Bujalowski, W.; Klonowska, M. M. *Biochemistry* **1993**, 32, 5888-5900.

442. Bujalowski, W.; Jezewska, M. J. *J. Mol. Biol.* **2000**, 295, 831-852.

443. Bujalowski, W.; Klonowska, M. M.; Jezewska, M. J. *J. Biol. Chem.* **1994**, 269, 31350-31358.

444. Nakayama, N.; Arai, N.; Kaziro, Y.; Arai, K. *J. Biol. Chem.* **1984**, 259, 88-96.

445. Miles, C. S.; Weigelt, J.; Stamford, N. P. J.; Dammerova, N.; Otting, G.; Dixon, N. E. *Biochem. Biophys. Res. Commun.* **1997**, 231, 126-130.

446. Chang, P.; Marians, K. J. *J. Biol. Chem.* **2000**, 275, 21687-21695.

447. Mitkova, A. V.; Khopde, S. M.; Biswas, S. B. *J. Biol. Chem.* **2003**, 278, 52253-52261.

448. Biswas, S. B.; Chen, P.-H.; Biswas, E. E. *Biochemistry* **1994**, 33, 11307-11314.

449. Sawaya, M. R.; Guo, S.; Tabor, D.; Richardson, C. C.; Ellenberger, T. *Cell* **1999**, 99, 167-177.

450. Singleton, M. R.; Sawaya, M. R.; Ellenberger, T.; Wigley, D. B. *Cell* **2000**, 101, 589-600.

451. Weigelt, J.; Brown, S. E.; Miles, C. S.; Dixon, N. E.; Otting, G. *Structure* **1999**, 7, 681-690.

452. Fass, D.; Bogden, C. E.; Berger, J. M. *Structure* **1999**, 7, 691-698.

453. Biswas, E. E.; Biswas, S. B.; Bishop, J. E. *Biochemistry* **1986**, 25, 7368-7374.

454. Jezewska, M. J.; Kim, U.-S.; Bujalowski, W. *Biophys. J.* **1996**, 71, 2075-2086.

455. San Martin, C.; Radermacher, M.; Wolpensinger, B.; Engel, A.; Miles, C. S.; Dixon, N. E.; Carazo, J. M. *Structure* **1998**, 6, 501-509.

456. Yu, X.; Jezewska, M. J.; Bujalowski, W.; Egelman, E. H. *J. Mol. Biol.* **1996**, 259, 7-14.

457. San Martin, M. C.; Stamford, N. P. J.; Dammerova, N.; Dixon, N. E.; Carazo, J. M. *J. Struct. Biol.* **1995**, 114, 167-176.

458. Donate, L.-E.; Llorca, O.; Barcena, M.; Brown, S. E.; Dixon, N. E.; Carazo, J. M. *J. Mol. Biol.* **2000**, 303, 383-393.

459. Barcena, M.; San Martin, C.; Weise, F.; Ayora, S.; Alonso, J. C.; Carazo, J. M. *J. Mol. Biol.* **1998**, 283, 809-819.

460. Fouts, E. T.; Yu, X.; Egelman, E. H.; Botchan, M. R. *J. Biol. Chem.* **1999**, 274, 4447-4458.

461. Yu, X.; Egelman, E. H. *Nat. Struct. Biol.***1997**, 4, 101-104.

462. Scherzinger, E.; Ziegelin, G.; Barcena, M.; Carazo, J. M.; Lurz, R.; Lanka, E. *J. Biol. Chem.* **1997**, 272, 30228-30236.

463. Kobori, J. A.; Kornberg, A. *J. Biol. Chem.* **1982**, 257, 13770-13775.

464. Wickner, S.; Hurwitz, J. *Proc. Natl Acad. Sci. USA* **1975**, 72, 921-925.

465. Jezewska, M. J.; Kim, U. S.; Bujalowski, W. *Biochemistry* **1996**, 35, 2129-2145.

466. Jezewska, M. J.; Rajendran, S.; Bujalowski, W. *J. Biol. Chem.* **1998**, 273, 9058-9069.

467. Bujalowski, W.; Jezewska, M. J. *Biochemistry* **1995**, 34, 8513-8519.

468. Wahle, E.; Lasken, R. S.; Kornberg, A. *J. Biol. Chem.* **1989**, 264, 2463-2468.

469. Wahle, E.; Lasken, R. S.; Kornberg, A. *J. Biol. Chem.* **1989**, 264, 2469-2475.

470. Barcena, M.; Ruiz, T.; Donate, L. E.; Brown, S. E.; Dixon, N. E.; Radermacher, M.; Carazo, J. M. *EMBO J.* **2001**, 20, 1462-1468.

471. Kobori, J. A.; Kornberg, A. *J. Biol. Chem.* **1982**, 257, 13763-13769.

472. Galletto, R.; Rajendran, S.; Bujalowski, W. *Biochemistry* **2000**, 39, 12959-12969.

473. Galletto, R.; Bujalowski, W. *Biochemistry* **2002**, 41, 8921-8934.

474. Galletto, R.; Bujalowski, W. *Biochemistry* **2002**, 41, 8907-8920.

475. Wilm, M. S.; Mann, M. *Intl J. Mass Spectrom.and Ion Process* **1994**, 136, 167-180.

476. Chung, E. W.; Henriques, D. A.; Renzoni, D.; Morton, C. J.; Mulhern, T. D.; Pitkeathly, M. C.; Ladbury, J. E.; Robinson, C. V. *Protein Sci.* **1999**, 8, 1962-1970.

477. Nettleton, E. J.; Sunde, M.; Lai, Z. H.; Kelly, J. W.; Dobson, C. M.; Robinson, C. V. *J. Mol. Biol.* **1998**, 281, 553-564.

478. Verentchikov, A. N.; Ens, W.; Standing, K. G. *Anal. Chem.* **1994**, 66, 126-133.

479. Rostom, A. A.; Robinson, C. V. *J. Am. Chem. Soc.* **1999**, 121, 4718-4719.

480. Tahallah, N.; Pinkse, M.; Maier, C. S.; Heck, A. J. R. *Rapid Commun. Mass Spectrom.* **2001**, 15, 596-601.

481. Krutchinsky, A. N.; Chernushevich, I. V.; Spicer, V. L.; Ens, W.; Standing, K. G. *J. Am. Soc. Mass Spectrom.* **1998**, 9, 569-579.

482. Van Berkel, W. J. H.; Van Den Heuvel, R. H. H.; Versluis, C.; Heck, A. J. R. *Protein Sci.* **2000**, 9, 435-439.

483. Fandrich, M.; Tito, M. A.; Leroux, M. R.; Rostom, A. A.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. *Proc. Natl Acad. Sci. USA* **2000**, 97, 14151-14155.

484. Smith, V. F.; Schwartz, B. L.; Randall, L. L.; Smith, R. D. *Protein Sci.* **1996**, 5, 488-494.

485. Bruce, J. E.; Smith, V. F.; Liu, C.; Randall, L. L.; Smith, R. D. *Protein Sci.* **1998**, 7, 1180-1185.

486. Garnier, C.; Lafitte, D.; Jorgensen, T. J.; Jensen, O. N.; Briand, C.; Peyrot, V. *Eur. J. Biochem.* **2001**, 268, 2402-2407.

487. Green, B. N.; Gotoh, T.; Suzuki, T.; Zal, F.; Lallier, F. H.; Toulmond, A.; Vinogradov, S. N. *J. Mol. Biol.* **2001**, 309, 553-560.

488. Fuerstenau, S. D.; Benner, W. H.; Thomas, J. J.; Brugidou, C.; Bothner, B.; Siuzdak, G. *Angew. Chem., Int. Ed.* **2001**, 40, 541-544.

489. Sobott, F.; McCammon, M. G.; Robinson, C. V. *Intl J. Mass Spectrom.***2003**, 230, 193-200.

490. Sobott, F.; Benesch, J. L. P.; Vierling, E.; Robinson, C. V. *J. Biol. Chem.* **2002**, 277, 38921-38929.

491. Bird, L. E.; Brannigan, J. A.; Subramanya, H. S.; Wigley, D. B. *Nucleic Acids Res.* **1998**, 26, 2686-2693.

492. Cheng, W.; Hsieh, J.; Brendza, K. M.; Lohman, T. M. *J. Mol. Biol.* **2001**, 310, 327-350.

493. Watt, S. J., *PhD Thesis*, University of Wollongong, **2005**

494. Sobott, F.; Robinson, C. V. *Intl J. Mass Spectrom.***2004**, 236, 25-32.

495. Rostom, A. A.; Fucini, P.; Benjamin, D. R.; Juenemann, R.; Nierhaus, K. H.; Hartl, F. U.; Dobson, C. M.; Robinson, C. V. *Proc. Natl Acad. Sci. USA* **2000**, 97, 5185-5190.

496. Bujalowski, W.; Klonowska, M. M. *Biochemistry* **1993**, 32, 5888-5900.

497. Beck, J. L.; Urathamakul, T.; Watt, S. J.; Sheil, M. M.; Schaeffer, P. M.; Dixon, N. E. *Expert Rev. Proteomic* **2006**, 3, 197-211.

498. Nettleton, E. J.; Tito, P.; Sunde, M.; Bouchard, M.; Dobson, C. M.; Robinson, C. V. *Biophys. J.* **2000**, 79, 1053-1065.

499. Miyata, T.; Yamada, K.; Iwasaki, H.; Shinagawa, H.; Morikawa, K.; Mayanagi, K. *J. Struct. Biol.* **2000**, 131, 83-89.

500. Yu, X.; VanLoock, M. S.; Poplawski, A.; Kelman, Z.; Xiang, T.; Tye, B. K.; Egelman, E. H. *EMBO Rep.* **2002**, 3, 792-797.

501. Stasiak, A. Z.; Larquet, E.; Stasiak, A.; Muller, S.; Engel, A.; Van Dyck, E.; West, S. C.; Egelman, E. H. *Curr. Biol.* **2000**, 10, 337-340.

502. Toth, E. A.; Li, Y.; Sawaya, M. R.; Cheng, Y. F.; Ellenberger, T. *Mol. Cell* **2003**, 12, 1113-1123.

503. Egelman, E. H.; Yu, X.; Wild, R.; Hingorani, M. M.; Patel, S. S. *Proc. Natl Acad. Sci. USA* **1995**, 92, 3869-3873.

504. Patel, S. S.; Hingorani, M. M. *J. Biol. Chem.* **1993**, 268, 10668-10675.

505. Yu, X.; Hingoran, M. M.; Patel, S. S.; Egelman, E. H. *Nat. Struct. Biol.***1996**, 3, 740-743.

506. VanLoock, M. S.; Chen, Y.-J.; Yu, X.; Patel, S. S.; Egelman, E. H. *J. Mol. Biol.* **2001**, 311, 951-956.

507. Crampton, D. J.; Ohi, M.; Qimron, U.; Walz, T.; Richardson, C. C. *J. Mol. Biol.* **2006**, 360, 667-677.

508. Guevremont, R.; Siu, K. W. M.; Leblanc, J. C. Y.; Berman, S. S. *J. Am. Soc. Mass Spectrom.* **1992**, 3, 216-224.

509. Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**, 310, 1658-1661.

510. Li, Y.; Cole, R. B. *Anal. Chem.* **2003**, 75, 5739-5746.

511. Williams, N. K.; Liepinsh, E.; Watt, S. J.; Prosselkov, P.; Matthews, J. M.; Attard, P.; Beck, J. L.; Dixon, N. E.; Otting, G. *J. Mol. Biol.* **2005**, 346, 1095-1108.

512. Yang, S.; Yu, X.; VanLoock, M. S.; Jezewska, M. J.; Bujalowski, W.; Egelman, E. H. *J. Mol. Biol.* **2002**, 321, 839-849.

513. Stasiak, A.; Tsaneva, I. R.; West, S. C.; Benson, C. J. B.; Yu, X.; Egelman, E. H. *Proc. Natl Acad. Sci. USA* **1994**, 91, 7618-7622.

514. Mastrangelo, I. A.; Hough, P. V.; Wall, J. S.; Dodson, M.; Dean, F. B.; Hurwitz, J. *Nature* **1989**, 338, 658-662.

515. Gomez-Llorente, Y.; Fletcher Ryan, J.; Chen Xiaojiang, S.; Carazo Jose, M.; San Martin, C. *J. Biol. Chem.* **2005**, 280, 40909-40915.

516. Chong, J. P. J.; Hayashi, M. K.; Simon, M. N.; Xu, R.-M.; Stillman, B. *Proc. Natl Acad. Sci. USA* **2000**, 97, 1530-1535.

517. Lanka, E.; Schuster, H. *Nucleic Acids Res.* **1983**, 11, 987-997.

518. Ludlam, A. V.; McNatt, M. W.; Carr, K. M.; Kaguni, J. M. *J. Biol. Chem.* **2001**, 276, 27345-27353.

519. Watt, S. J.; Urathamakul, T.; Schaeffer, P. M.; Williams, N. K.; Sheil, M. M.; Dixon, N. E.; Beck, J. L. *Rapid Commun. Mass Spectrom.* **2006**, 21, 132-140.

520. Biswas, E. E.; Biswas, S. B. *Biochemistry* **1999**, 38, 10919-10928.

521. Rajendran, S.; Jezewska, M. J.; Bujalowski, W. *J. Mol. Biol.* **2000**, 303, 773-795.

522. von Hippel, P. H.; Delagoutte, E. *BioEssays* **2003**, 25, 1168-1177.

523. Iwai, H.; Pluckthun, A. *FEBS Lett.* **1999**, 459, 166-172.

524. Trabi, M.; Craik, D. J. *Trends Biochem. Sci.* **2002**, 27, 132-138.

525. Blond, A.; Peduzzi, J.; Goulard, C.; Chiuchiolo, M. J.; Barthelemy, M.; Prigent, Y.; Salomon, R. A.; Farias, R. N.; Moreno, F.; Rebuffat, S. *Eur. J. Biochem.* **1999**, 259, 747-755.

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

527. Luckett, S.; Garcia, R. S.; Barker, J. J.; Konarev, A. V.; Shewry, P. R.; Clarke, A. R.; Brady, R. L. *J. Mol. Biol.* **1999**, 290, 525-533.

528. 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. *Biochemistry* **2000**, 39, 5722-5730.

529. Craik, D. J.; Daly, N. L.; Bond, T.; Waine, C. *J. Mol. Biol.* **1999**, 294, 1327-1336.

530. Felizmenio-Quimio, M. E.; Daly, N. L.; Craik, D. J. *J. Biol. Chem.* **2001**, 276, 22875-22882.

531. Iwai, H.; Lingel, A.; Pluckthun, A. *J. Biol. Chem.* **2001**, 276, 16548-16554.

532. Camarero, J. A.; Fushman, D.; Sato, S.; Giriat, I.; Cowburn, D.; Raleigh, D. P.; Muir, T. W. *J. Mol. Biol.* **2001**, 308, 1045-1062.

533. Scott, C. P.; Abel-Santos, E.; Wall, M.; Wahnon, D. C.; Benkovic, S. J. *Proc. Natl Acad. Sci. USA* **1999**, 96, 13638-13643.

534. Siebold, C.; Erni, B. *Biophys. Chem.* **2002**, 96, 163-171.

535. Satoh, T.; Li, S.; Friedman, T. M.; Wiaderkiewicz, R.; Korngold, R.; Huang, Z. W. *Biochem. Biophys. Res. Commun.* **1996**, 224, 438-443.

536. Gustafson, K. R.; Sowder, R. C.; Henderson, L. E.; Parsons, I. C.; Kashman, Y.; Cardellina, J. H.; McMahon, J. B.; Buckheit, R. W.; Pannell, L. K.; Boyd, M. R. *J. Am. Chem. Soc.* **1994**, 116, 9337-9338.

537. Tam, J. P.; Lu, Y.-A.; Yang, J.-L.; Chiu, K.-W. *Proc. Natl Acad. Sci. USA* **1999**, 96, 8913-8918.

538. Goldenberg, D. P.; Creighton, T. E. *J. Mol. Biol.* **1983**, 165, 407-413.

539. Goldenberg, D. P.; Creighton, T. E. *J. Mol. Biol.* **1984**, 179, 527-545.

540. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, 266, 776-779.

541. Camarero, J. A.; Pavel, J.; Muir, T. W. *Angew. Chemie, Int. Ed.* **1998**, 37, 347-349.

542. Erlanson, D. A.; Chytil, M.; Verdine, G. L. *Chem. Biol.* **1996**, 3, 981-991.

543. Xu, M.-Q.; Southworth, M. W.; Mersha, F. B.; Hornstra, L. J.; Perler, F. B. *Cell* **1993**, 75, 1371-1377.

544. Xu, M.-Q.; Comb, D. G.; Paulus, H.; Noren, C. J.; Shao, Y.; Perler, F. B. *EMBO J.* **1994**, 13, 5517-5522.

545. Shao, Y.; Xu, M. Q.; Paulus, H. *Biochemistry* **1995**, 34, 10844-10850.

546. Porter, J. A.; Von Kessler, D. P.; Ekker, S. C.; Young, K. E.; Lee, J. J.; Moses, K.; Beachy, P. A. *Nature* **1995**, 374, 363-366.

547. Recsei, P. A.; Huynh, Q. K.; Snell, E. E. *Proc. Natl Acad. Sci. USA* **1983**, 80, 973-977.

548. Brannigan, J. A.; Dodson, G.; Duggleby, H. J.; Moody, P. C. E.; Smith, J. L.; Tomchick, D. R.; Murzin, A. G. *Nature* **1995**, 278, 416-419.

549. Shao, Y.; Xu, M. Q.; Paulus, H. *Biochemistry* **1996**, 35, 3810-3815.

550. Shao, Y.; Paulus, H. *J. Pept. Res.* **1997**, 50, 193-198.

551. Kane, P. M.; Yamashiro, C. T.; Wolczyk, D. F.; Neff, N.; Goebl, M.; Stevens, T. H. *Science* **1990**, 250, 651-657.

552. Paulus, H. *Annu. Rev. Biochem.* **2000**, 69, 447-496.

553. Kawasaki, M.; Satow, Y.; Ohya, Y.; Anraku, Y. *FEBS Lett.* **1997**, 412, 518-520.

554. Xu, M. Q.; Perler, F. B. *EMBO J.l* **1996**, 15, 5146-5153.

555. Evans, T. C.; Benner, J.; Xu, M. Q. *J. Biol. Chem.* **1999**, 274, 18359-18363.

556. Camarero, J. A.; Muir, T. W. *J. Am. Chem. Soc.* **1999**, 121, 5597-5598.

557. Perler, F. B.; Xu, M.-Q.; Paulus, H. *Curr. Opin. Chem. Biol.* **1997**, 1, 292-299.

558. Pietrokovski, S. *Protein Sci.* **1994**, 3, 2340-2350.

559. Perler, F. B. *Cell* **1998**, 92, 1-4.

560. Liu, X.-Q. *Annu. Rev. Genet.* **2000**, 34, 61-76.

561. Mills, K. V.; Lew, B. M.; Jiang, S.-Q.; Paulus, H. *Proc. Natl Acad. Sci. USA* **1998**, 95, 3543-3548.

562. Shingledecker, K.; Jiang, S. Q.; Paulus, H. *Gene* **1998**, 207, 187-195.

563. Southworth, M. W.; Adam, E.; Panne, D.; Byer, R.; Kautz, R.; Perler, F. B. *EMBO J.* **1998**, 17, 918-926.

564. Wu, H.; Xu, M. Q.; Liu, X. Q. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1998**, 1387, 422-432.

565. Zhou, H. X. *J. Mol. Biol.* **2003**, 332, 257-264.

566. Deechongkit, S.; Kelly, J. W. *J. Am. Chem. Soc.* **2002**, 124, 4980-4986.

567. Ferraro, D. M.; Noel, D. L.; Robertson, A. D. *Biochemistry* **2004**, 43, 587-594.

568. Wand, A. J.; Roder, H.; Englander, S. W. *Biochemistry* **1986**, 25, 1107-1114.

569. Bai, Y.; Sosnick, T. R.; Mayne, L.; Englander, S. W. *Science* **1995**, 269, 192-197.

570. Arrington, C. B.; Teesch, L. M.; Robertson, A. D. *J. Mol. Biol.* **1999**, 285, 1265-1275.

571. Perrett, S.; Clarke, J.; Hounslow, A. M.; Fersht, A. R. *Biochemistry* **1995**, 34, 9288-9298.

572. Woodward, C.; Simon, I.; Tuchsen, E. *Mol. Cell. Biochem.* **1982**, 48, 135-160.

573. Zhang, Z.; Post, C. B.; Smith, D. L. *Biochemistry* **1996**, 35, 779-791.

574. Englander, S. W.; Kallenbach, N. R. *Q. Rev. Biophys.* **1983**, 16, 521-655.

575. Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins* **1993**, 17, 75-86.

576. Hvidt, A.; Nielsen, S. O. *Adv. Protein Chem.* **1966**, 21, 287-386.

577. Smith, D. L.; Deng, Y.; Zhang, Z. *J. Mass Spectrom.* **1997**, 32, 135-146.

578. Powell, K. D.; Wales, T. E.; Fitzgerald, M. C. *Protein Sci.* **2002**, 11, 841-851.

579. Liu, K.; Cho, H. S.; Hoyt, D. W.; Nguyen, T. N.; Olds, P.; Kelly, J. W.; Wemmer, D. E. *J. Mol. Biol.* **2000**, 303, 555-565.

580. Miranker, A.; Robinson, C. V.; Radford, S. E.; Aplin, R. T.; Dobson, C. M. *Science* **1993**, 262, 896-900.

581. Englander, S. W.; Downer, N. W.; Teitelbaum, H. *Annu. Rev. Biochem.* **1972**, 41, 903-924.

582. Mayo, S. L.; Baldwin, R. L. *Science* **1993**, 262, 873-876.

583. Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins: Struct., Funct., Genet.* **1994**, 20, 4-14.

584. Kaltashov, I. A.; Eyles, S. J. *J. Mass Spectrom.* **2002**, 37, 557-565.

585. Chamberlain, A. K.; Handel, T. M.; Marqusee, S. *Nat. Struct. Biol.***1996**, 3, 782-787.

586. Yi, Q.; Baker, D. *Protein Sci.* **1996**, 5, 1060-1066.

587. Roder, H.; Wagner, G.; Wuthrich, K. *Biochemistry* **1985**, 24, 7396-7407.

588. Lenormant, H.; Blout, E. R. *Nature* **1953**, 172, 770-771.

589. Hvidt, A.; Linderstrom-Lang, K. *Biochim. Biophys. Acta* **1954**, 14, 574-575.

590. Leach, S. J.; Hill, J. *Biochemistry* **1963**, 128, 807-813.

591. Saunders, M.; Wishnia, A. *Annu. New York Acad. Sci.* **1958**, 70, 870-874.

592. Wagner, G.; Wuthrich, K. *J. Mol. Biol.* **1979**, 130, 31-37.

593. Wagner, G.; Wuthrich, K. *J. Mol. Biol.* **1982**, 160, 343-361.

594. Wagner, G.; Stassinopoulou, C. I.; Wuthrich, K. *Eur. J. Biochem.* **1984**, 145, 431-436.

595. Englander, S. W.; Milne, J. *Protein Structure and Function, Proceedings International Symposium*, **1990**, 117-133.

596. Zhang, Y. Z.; Paterson, Y.; Roder, H. *Protein Sci.* **1995**, 4, 804-814.

597. Englander, S. W.; Mayne, L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, 21, 243-265.

598. Katta, V.; Chait, B. T. *Rapid Commun. Mass Spectrom.* **1991**, 5, 214-217.

599. Thevenon-Emeric, G.; Kozlowski, J.; Zhang, Z.; Smith, D. L. *Anal. Chem.* **1992**, 64, 2456-2458.

600. Nemirovskiy, O.; Giblin, D. E.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1999**, 10, 711-718.

601. Engen, J. R.; Smith, D. L. *Methods Mol. Biol.* **2000**, 146, 95-112.

602. Kaltashov, I. A.; Eyles, S. J. *Mass Spectrom. Rev.* **2002**, 21, 37-71.

603. Kaltashov, I. A. *Intl J. Mass Spectrom.* **2005**, 240, 249-259.

604. Mandell, J. G.; Falick, A. M.; Komives, E. A. *Anal. Chem.* **1998**, 70, 3987-3995.

605. Nazabal, A.; Dos Reis, S.; Bonneu, M.; Saupe, S. J.; Schmitter, J. M. *Biochemistry* **2003**, 42, 8852-8861.

606. Zhang, Z.; Smith, D. L. *Protein Sci.* **1993**, 2, 522-531.

607. Eyles, S. J.; Speir, J. P.; Kruppa, G. H.; Gierasch, L. M.; Kaltashov, I. A. *J. Am. Chem. Soc.* **2000**, 122, 495-500.

608. Horn, D. M.; Breuker, K.; Frank, A. J.; McLafferty, F. W. *J. Am. Chem. Soc.* **2001**, 123, 9792-9799.

609. Akashi, S.; Takio, K. *Protein Sci.* **2000**, 9, 2497-2505.

610. Mandell, J. G.; Falick, A. M.; Komives, E. A. *Proc. Natl Acad. Sci. USA* **1998**, 95, 14705-14710.

611. Robinson, C. V.; Gross, M.; Eyles, S. J.; Ewbank, J.; Mayhew, M.; Hartl, F. U.; Dobson, C. M.; Radford, S. E. *Nature* **1994**, 372, 646-651.

612. Wang, F.; Blanchard, J. S.; Tang, X. J. *Biochemistry* **1997**, 36, 3755-3759.

613. Zhu, M. M.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2004**, 15, 388-397.

614. Anderegg, R. J.; Wagner, D. S. *J. Am. Chem. Soc.* **1995**, 117, 1374-1377.

615. Smith, D. L.; Zhang, Z.; Liu, Y. *Pure Appl. Chem.* **1994**, 66, 89-94.

616. Arrington, C. B.; Robertson, A. D. *J. Mol. Biol.* **2000**, 300, 221-232.

617. Wagner, D. S.; Melton, L. G.; Yan, Y.; Erickson, B. W.; Anderegg, R. J. *Protein Sci.* **1994**, 3, 1305-1314.

618. Charlebois, J. P.; Patrie, S. M.; Kelleher, N. L. *Anal. Chem.* **2003**, 75, 3263-3266.

619. Kim, M. Y.; Maier, C. S.; Reed, D. J.; Deinzer, M. L. *Protein Sci.* **2002**, 11, 1320-1329.

620. Yamamoto, T.; Izumi, S.; Gekko, K. *J. Biochem.* **2004**, 135, 17-24.

621. Lanman, J.; Lam, T. T.; Barnes, S.; Sakalian, M.; Emmett, M. R.; Marshall, A. G.; Prevelige, P. E. *J. Mol. Biol.* **2003**, 325, 759-772.

622. Johnson, R. S.; Walsh, K. A. *Protein Sci.* **1994**, 3, 2411-2418.

623. Tito, P.; Nettleton, E. J.; Robinson, C. V. *J. Mol. Biol.* **2000**, 303, 267-278.

624. Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L., Jr. *Proc. Natl Acad. Sci. USA* **2003**, 100, 7057-7062.

625. Kim, M. Y.; Maier, C. S.; Reed, D. J.; Deinzer, M. L. *J. Am. Chem. Soc.* **2001**, 123, 9860-9866.

626. Hoerner, J. K.; Xiao, H.; Dobo, A.; Kaltashov, I. A. *J. Am. Chem. Soc.* **2004**, 126, 7709-7717.

627. Demmers, J. A. A.; Van Duijn, E.; Haverkamp, J.; Greathouse, D. V.; Koeppe, R. E., II; Heck, A. J. R.; Killian, J. A. *J. Biol. Chem.* **2001**, 276, 34501-34508.

628. Demmers, J. A. A.; Haverkamp, J.; Heck, A. J. R.; Koeppe, R. E., II; Killian, J. A. *Proc. Natl Acad. Sci. USA* **2000**, 97, 3189-3194.

629. Deng, Y.; Pan, H.; Smith, D. L. *J. Am. Chem. Soc.* **1999**, 121, 1966-1967.

630. Demmers, J. A. A.; Rijkers, D. T. S.; Haverkamp, J.; Killian, J. A.; Heck, A. J. R. *J. Am. Chem. Soc.* **2002**, 124, 11191-11198.

631. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, 120, 3265-3266.

632. Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, 72, 563-573.

633. Haselmann, K. F.; Budnik, B. A.; Olsen, J. V.; Nielsen, M. L.; Reis, C. A.; Clausen, H.; Johnsen, A. H.; Zubarev, R. A. *Anal. Chem.* **2001**, 73, 2998-3005.

634. McLafferty, F. W.; Horn, D. M.; Breuker, K.; Ge, Y.; Lewis, M. A.; Cerda, B.; Zubarev, R. A.; Carpenter, B. K. *J. Am. Soc. Mass Spectrom.* **2001**, 12, 245-249.

635. Arai, K.; Kornberg, A. *J. Biol. Chem.* **1981**, 256, 5253-5259.

636. Nakayama, N.; Arai, N.; Bond, M. W.; Kaziro, Y.; Arai, K. *J. Biol. Chem.* **1984**, 259, 97-101.

637. Maier, C. S.; Schimerlik, M. I.; Deinzer, M. L. *Biochemistry* **1999**, 38, 1136-1143.

638. Sivaraman, T.; Arrington, C. B.; Robertson, A. D. *Nat. Struct. Biol.***2001**, 8, 331-333.

639. Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.* **1998**, 277, 985-994.

640. Plaxco, K. W.; Simons, K. T.; Ruczinski, I.; David, B. *Biochemistry* **2000**, 39, 11177-11183.

APPENDICES

Appendix 1 *Assignments for ions observed in ESI mass spectra of reaction mixtures containing ruthenium compounds and dsDNA.* N.O. = ion not observed.

Assignment	Observed *m/z* for D1	Observed *m/z* for D2	Observed *m/z* for D3
$[Ru(phen)_3+oligo-8H]^{6-}$	1733.9	1733.5	1732.6
$[Ru(phen)_3+oligo-7H]^{5-}$	2081.1	2080.1	2079.3
$[2Ru(phen)_3+oligo-10H]^{6-}$	1840.6	1840.1	1839.6
$[2Ru(phen)_3+oligo-9H]^{5-}$	2208.8	2208.1	2207.4
$[3Ru(phen)_3+oligo-11H]^{5-}$	N.O.	2336.2	2335.1
$[4Ru(phen)_3+oligo-13H]^{5-}$	N.O.	2464.4	2463.6
$[Ru(phen)_2(pda)+oligo-8H]^{6-}$	1738.5	1737.7	1737.1
$[Ru(phen)_2(pda)+oligo-7H]^{5-}$	2086.3	2085.3	2084.6
$[2Ru(phen)_2(pda)+oligo-10H]^{6-}$	1849.9	1848.9	N.O.
$[2Ru(phen)_2(pda)+oligo-9H]^{5-}$	2219.8	2219.4	N.O.
$[Ru(phen)_2(dpq)+oligo-8H]^{6-}$	1742.7	1742.0	1741.3
$[Ru(phen)_2(dpq)+oligo-7H]^{5-}$	2091.4	2090.8	2089.7
$[2Ru(phen)_2(dpq)+oligo-10H]^{6-}$	1857.7	1857.4	1856.3
$[2Ru(phen)_2(dpq)+oligo-9H]^{5-}$	2229.6	2228.9	2228.4
$[3Ru(phen)_2(dpq)+oligo-12H]^{6-}$	1972.9	1972.9	1971.7
$[3Ru(phen)_2(dpq)+oligo-11H]^{5-}$	2368.2	2367.2	2366.5
$[4Ru(phen)_2(dpq)+oligo-13H]^{5-}$	N.O.	2505.9	N.O.
$[Ru(phen)_2(dpqMe_2)+oligo-8H]^{6-}$	1747.2	N.O.	1745.6
$[Ru(phen)_2(dpqMe_2)+oligo-7H]^{5-}$	N.O.	N.O.	2095.5
$[2Ru(phen)_2(dpqMe_2)+oligo-10H]^{6-}$	1866.8	1866.7	1865.9
$[2Ru(phen)_2(dpqMe_2)+oligo-9H]^{5-}$	2239.7	2243.3	2238.9
$[3Ru(phen)_2(dpqMe_2)+oligo-13H]^{7-}$	N.O.	1702.5	N.O.
$[3Ru(phen)_2(dpqMe_2)+oligo-12H]^{6-}$	1987.1	1986.2	1985.7
$[3Ru(phen)_2(dpqMe_2)+oligo-11H]^{5-}$	2384.1	2383.9	2383.2
$[4Ru(phen)_2(dpqMe_2)+oligo-15H]^{7-}$	N.O.	1805.2	N.O.
$[4Ru(phen)_2(dpqMe_2)+oligo-14H]^{6-}$	2107.4	2106.3	2105.8
$[4Ru(phen)_2(dpqMe_2)+oligo-13H]^{5-}$	2528.3	2528.1	2527.3
$[5Ru(phen)_2(dpqMe_2)+oligo-17H]^{7-}$	N.O.	1908.0	1907.6
$[5Ru(phen)_2(dpqMe_2)+oligo-16H]^{6-}$	N.O.	2226.3	N.O.
$[5Ru(phen)_2(dpqMe_2)+oligo-15H]^{5-}$	N.O.	2687.5	N.O.

$[Ru(phen)_2(dpqC)+oligo-8H]^{6-}$	1751.7	1751.0	1750.4
$[Ru(phen)_2(dpqC)+oligo-7H]^{5-}$	2102.2	2101.4	2100.5
$[2Ru(phen)_2(dpqC)+oligo-10H]^{6-}$	1876.0	1875.4	1874.9
$[2Ru(phen)_2(dpqC)+oligo-9H]^{5-}$	2251.8	2250.8	2249.6
$[3Ru(phen)_2(dpqC)+oligo-12H]^{6-}$	2000.2	1999.7	2002.1
$[3Ru(phen)_2(dpqC)+oligo-11H]^{5-}$	2400.1	2400.0	2398.8
$[4Ru(phen)_2(dpqC)+oligo-13H]^{5-}$	N.O.	2549.8	N.O.
$[Ru(phen)_2(dppz)+oligo-8H]^{6-}$	1750.9	1750.3	1749.7
$[Ru(phen)_2(dppz)+oligo-7H]^{5-}$	2101.5	2100.8	2100.0
$[2Ru(phen)_2(dppz)+oligo-11H]^{7-}$	1606.6	N.O.	N.O.
$[2Ru(phen)_2(dppz)+oligo-10H]^{6-}$	1874.5	1874.2	1873.2
$[2Ru(phen)_2(dppz)+oligo-9H]^{5-}$	2249.8	2249.1	2248.6
$[3Ru(phen)_2(dppz)+oligo-13H]^{7-}$	1712.7	1712.0	1711.5
$[3Ru(phen)_2(dppz)+oligo-12H]^{6-}$	1998.3	1997.7	1997.1
$[3Ru(phen)_2(dppz)+oligo-11H]^{5-}$	2398.3	2397.4	2396.6
$[4Ru(phen)_2(dppz)+oligo-15H]^{7-}$	1818.7	1818.4	1817.8
$[4Ru(phen)_2(dppz)+oligo-14H]^{6-}$	2122.0	2121.5	2120.8
$[4Ru(phen)_2(dppz)+oligo-13H]^{5-}$	2546.4	2546.1	2544.7
$[5Ru(phen)_2(dppz)+oligo-17H]^{7-}$	N.O.	1924.1	1923.6
$[5Ru(phen)_2(dppz)+oligo-16H]^{6-}$	N.O.	2244.8	2244.4
$[5Ru(phen)_2(dppz)+oligo-15H]^{5-}$	N.O.	2694.5	2693.2

Appendix 2 *Assignments for ions observed in ESI mass spectra of reaction mixtures containing either daunomycin or distamycin, DNA and ruthenium compounds*. N.O. = ion not observed.

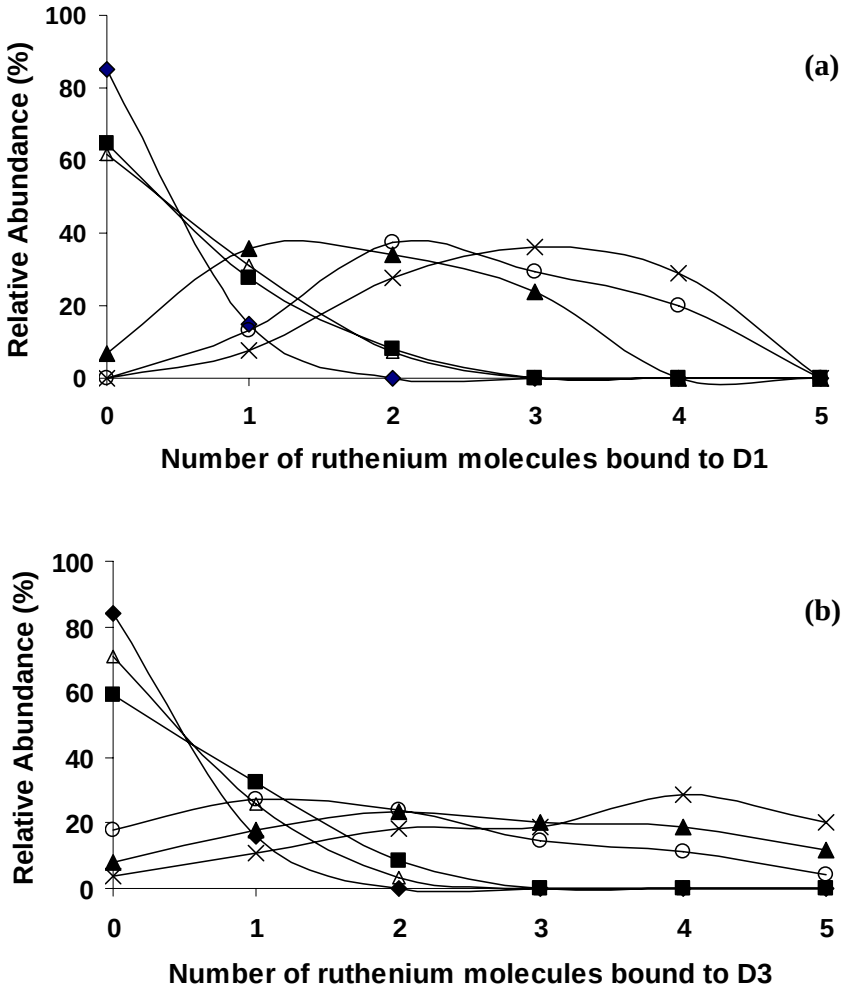
Assignment	Observed *m/z* for D2	Observed *m/z* for D3
[5daunomycin+oligo-7H]$^{7-}$	1770.6	N.O.
[6daunomycin+oligo-7H]$^{7-}$	1846.0	N.O.
[7daunomycin+oligo-7H]$^{7-}$	1921.4	N.O.
[8daunomycin+oligo-7H]$^{7-}$	1996.9	N.O.
[daunomycin+oligo-6H]$^{6-}$	1714.3	1713.7
[2daunomycin+oligo-6H]$^{6-}$	1802.3	1801.6
[3daunomycin+oligo-6H]$^{6-}$	1890.3	1889.6
[4daunomycin+oligo-6H]$^{6-}$	1978.2	N.O.
[5daunomycin+oligo-6H]$^{6-}$	2066.0	N.O.
[6daunomycin+oligo-6H]$^{6-}$	2153.9	N.O.
[7daunomycin+oligo-6H]$^{6-}$	2241.9	N.O.
[8daunomycin+oligo-6H]$^{6-}$	2329.9	N.O.
[daunomycin+oligo-5H]$^{5-}$	2057.6	2056.5
[2daunomycin+oligo-5H]$^{5-}$	2162.8	2162.4
[4daunomycin+Ru(phen)$_3$+oligo-8H]$^{6-}$	2085.0	N.O.
[4daunomycin+2Ru(phen)$_3$+oligo-10H]$^{6-}$	2194.1	N.O.
[4daunomycin+3Ru(phen)$_3$+oligo-12H]$^{6-}$	2298.6	N.O.
[5daunomycin+Ru(phen)$_3$+oligo-7H]$^{5-}$	2607.1	N.O.
[5daunomycin+Ru(phen)$_3$+oligo-8H]$^{6-}$	2172.6	N.O.
[5daunomycin+Ru(phen)$_3$+oligo-9H]$^{7-}$	1862.4	N.O.
[5daunomycin+2Ru(phen)$_3$+oligo-9H]$^{5-}$	2736.1	N.O.
[5daunomycin+2Ru(phen)$_3$+oligo-10H]$^{6-}$	2279.3	N.O.
[5daunomycin+2Ru(phen)$_3$+oligo-11H]$^{7-}$	1953.6	N.O.
[5daunomycin+3Ru(phen)$_3$+oligo-12H]$^{6-}$	2385.9	N.O.
[6daunomycin+Ru(phen)$_3$+oligo-8H]$^{6-}$	2260.6	N.O.
[6daunomycin+Ru(phen)$_3$+oligo-9H]$^{7-}$	1937.7	N.O.
[6daunomycin+2Ru(phen)$_3$+oligo-10H]$^{6-}$	2367.4	N.O.
[6daunomycin+2Ru(phen)$_3$+oligo-11H]$^{7-}$	2028.8	N.O.
[daunomycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	1829.7	1829.0
[2daunomycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	1917.3	N.O.
[4daunomycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	2093.4	N.O.
[4daunomycin+Ru(phen)$_2$(dpq)+oligo-9H]$^{7-}$	1792.5	N.O.
[5daunomycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	2181.5	N.O.
[5daunomycin+Ru(phen)$_2$(dpq)+oligo-9H]$^{7-}$	1869.5	N.O.
[6daunomycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	2269.2	N.O.

[6daunomycin+Ru(phen)$_2$(dpq)+oligo-9H]$^{7-}$	1944.6	N.O.
[daunomycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	1838.7	1837.8
[2daunomycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	1926.7	1925.5
[daunomycin+Ru(phen)$_2$(dpqC)+oligo-7H]$^{5-}$	2206.6	N.O
[3daunomycin+2Ru(phen)$_2$(dpqC)+oligo-10H]$^{6-}$	2138.4	N.O.
[3daunomycin+2Ru(phen)$_2$(dpqC)+oligo-11H]$^{7-}$	1833.0	N.O.
[3daunomycin+1Ru(phen)$_2$(dpqC)+oligo-12H]$^{6-}$	2263.1	N.O.
[4daunomycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	2102.5	N.O.
[4daunomycin+Ru(phen)$_2$(dpqC)+oligo-9H]$^{7-}$	1802.3	N.O.
[4daunomycin+2Ru(phen)$_2$(dpqC)+oligo-9H]$^{5-}$	2672.7	N.O.
[4daunomycin+2Ru(phen)$_2$(dpqC)+oligo-10H]$^{6-}$	2226.7	N.O.
[4daunomycin+2Ru(phen)$_2$(dpqC)+oligo-11H]$^{7-}$	1908.6	N.O.
[4daunomycin+3Ru(phen)$_2$(dpqC)+oligo-12H]$^{6-}$	2350.7	N.O.
[4daunomycin+3Ru(phen)$_2$(dpqC)+oligo-13H]$^{7-}$	2014.4	N.O.
[5daunomycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	2190.3	N.O.
[5daunomycin+Ru(phen)$_2$(dpqC)+oligo-9H]$^{7-}$	1877.4	N.O.
[5daunomycin+2Ru(phen)$_2$(dpqC)+oligo-10H]$^{6-}$	2314.7	N.O.
[5daunomycin+2Ru(phen)$_2$(dpqC)+oligo-11H]$^{7-}$	1983.7	N.O.
[6daunomycin+Ru(phen)$_2$(dpqC)+oligo-9H]$^{7-}$	1952.8	N.O.
[6daunomycin+2Ru(phen)$_2$(dpqC)+oligo-10H]$^{6-}$	2402.7	N.O.
[daunomycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	1838.2	1837.4
[2daunomycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	1926.2	1925.3
[3daunomycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	2013.7	N.O.
[daunomycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	1961.9	1961.0
[2daunomycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	2049.6	N.O.
[daunomycin+Ru(phen)$_2$(dppz)+oligo-7H]$^{5-}$	2205.9	2204.9
[3daunomycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	2137.0	N.O.
[3daunomycin+2Ru(phen)$_2$(dppz)+oligo-11H]$^{7-}$	1831.8	N.O.
[3daunomycin+3Ru(phen)$_2$(dppz)+oligo-11H]$^{5-}$	2713.4	N.O.
[3daunomycin+3Ru(phen)$_2$(dppz)+oligo-12H]$^{6-}$	2261.0	N.O.
[3daunomycin+3Ru(phen)$_2$(dppz)+oligo-13H]$^{7-}$	1938.5	N.O.
[4daunomycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	2101.7	N.O.
[4daunomycin+2Ru(phen)$_2$(dppz)+oligo-9H]$^{5-}$	2670.8	N.O.
[4daunomycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	2224.9	N.O.
[4daunomycin+2Ru(phen)$_2$(dppz)+oligo-11H]$^{7-}$	1907.6	N.O.
[4daunomycin+3Ru(phen)$_2$(dppz)+oligo-12H]$^{6-}$	2349.1	N.O.
[4daunomycin+3Ru(phen)$_2$(dppz)+oligo-13H]$^{7-}$	2013.8	N.O.
[5daunomycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	2189.7	N.O.
[5daunomycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	2312.7	N.O.
[5daunomycin+3Ru(phen)$_2$(dppz)+oligo-12H]$^{6-}$	2437.5	N.O.
[6daunomycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	2401.2	N.O.

[3daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-10H]$^{6-}$	2129.7	N.O.
[3daunomycin+3Ru(phen)$_2$(dpqMe$_2$)+oligo-12H]$^{6-}$	2250.0	N.O.
[4daunomycin+Ru(phen)$_2$(dpqMe$_2$)+oligo-7H]$^{5-}$	2518.0	N.O.
[4daunomycin+Ru(phen)$_2$(dpqMe$_2$)+oligo-8H]$^{6-}$	2098.0	N.O.
[4daunomycin+Ru(phen)$_2$(dpqMe$_2$)+oligo-9H]$^{7-}$	1798.1	N.O.
[4daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-9H]$^{5-}$	2661.6	N.O.
[4daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-10H]$^{6-}$	2218.0	N.O.
[4daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-11H]$^{7-}$	1900.8	N.O.
[4daunomycin+3Ru(phen)$_2$(dpqMe$_2$)+oligo-11H]$^{5-}$	2806.1	N.O.
[4daunomycin+3Ru(phen)$_2$(dpqMe$_2$)+oligo-12H]$^{6-}$	2338.4	N.O.
[5daunomycin+Ru(phen)$_2$(dpqMe$_2$)+oligo-8H]$^{6-}$	2186.1	N.O.
[5daunomycin+Ru(phen)$_2$(dpqMe$_2$)+oligo-9H]$^{7-}$	1873.1	N.O.
[5daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-9H]$^{5-}$	2767.6	N.O.
[5daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-10H]$^{6-}$	2305.9	N.O.
[5daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-11H]$^{7-}$	1976.5	N.O.
[6daunomycin+2Ru(phen)$_2$(dpqMe$_2$)+oligo-10H]$^{6-}$	2393.5	N.O.
[distamycin+oligo-6H]$^{6-}$	N.O.	1706.1
[2distamycin+oligo-6H]$^{6-}$	1787.0	1786.2
[4distamycin+oligo-6H]$^{6-}$	1947.7	N.O.
[5distamycin+oligo-6H]$^{6-}$	2027.9	N.O.
[6distamycin+oligo-6H]$^{6-}$	2108.2	N.O.
[distamycin+oligo-5H]$^{5-}$	N.O.	2047.5
[2distamycin+oligo-5H]$^{5-}$	2144.5	2143.8
[distamycin+2Ru(phen)$_3$+oligo-10H]$^{6-}$	1920.9	N.O.
[4distamycin+Ru(phen)$_3$+oligo-7H]$^{5-}$	2465.4	N.O.
[4distamycin+Ru(phen)$_3$+oligo-8H]$^{6-}$	2054.5	N.O.
[4distamycin+Ru(phen)$_3$+oligo-9H]$^{7-}$	1760.9	N.O.
[4distamycin+2Ru(phen)$_3$+oligo-9H]$^{5-}$	2593.3	N.O.
[4distamycin+2Ru(phen)$_3$+oligo-10H]$^{6-}$	2161.2	N.O.
[distamycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	N.O.	1821.5
[2distamycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	1902.3	1901.4
[2distamycin+2Ru(phen)$_2$(dpq)+oligo-9H]$^{5-}$	2422.0	N.O.
[2distamycin+2Ru(phen)$_2$(dpq)+oligo-10H]$^{6-}$	2017.8	N.O.
[2distamycin+3Ru(phen)$_2$(dpq)+oligo-11H]$^{5-}$	2559.9	N.O.
[4distamycin+Ru(phen)$_2$(dpq)+oligo-7H]$^{5-}$	2475.9	N.O.
[4distamycin+Ru(phen)$_2$(dpq)+oligo-8H]$^{6-}$	2062.9	N.O.
[4distamycin+Ru(phen)$_2$(dpq)+oligo-9H]$^{7-}$	1768.3	N.O.
[4distamycin+2Ru(phen)$_2$(dpq)+oligo-9H]$^{5-}$	2614.3	N.O.
[4distamycin+2Ru(phen)$_2$(dpq)+oligo-10H]$^{6-}$	2178.5	N.O.
[distamycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	N.O.	1830.4
[2distamycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	1911.3	1910.6

[2distamycin+2Ru(phen)$_2$(dpqC)+oligo-9H]$^{5-}$	2443.7	N.O.
[2distamycin+2Ru(phen)$_2$(dpqC)+oligo-10H]$^{6-}$	2035.9	N.O.
[2distamycin+2Ru(phen)$_2$(dpqC)+oligo-11H]$^{7-}$	1745.3	N.O.
[2distamycin+3Ru(phen)$_2$(dpqC)+oligo-11H]$^{5-}$	2592.6	N.O.
[2distamycin+3Ru(phen)$_2$(dpqC)+oligo-12H]$^{6-}$	2160.4	N.O.
[2distamycin+3Ru(phen)$_2$(dpqC)+oligo-13H]$^{7-}$	1851.3	N.O.
[2distamycin+4Ru(phen)$_2$(dpqC)+oligo-13H]$^{5-}$	2794.0	N.O.
[2distamycin+4Ru(phen)$_2$(dpqC)+oligo-14H]$^{6-}$	2284.4	N.O.
[4distamycin+Ru(phen)$_2$(dpqC)+oligo-8H]$^{6-}$	2071.9	N.O.
[distamycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	N.O.	1830.1
[2distamycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	1910.7	1910.1
[2distamycin+2Ru(phen)$_2$(dppz)+oligo-9H]$^{5-}$	2441.7	N.O.
[2distamycin+2Ru(phen)$_2$(dppz)+oligo-10H]$^{6-}$	2034.7	2033.9
[2distamycin+2Ru(phen)$_2$(dppz)+oligo-11H]$^{7-}$	1743.7	2033.9
[distamycin+Ru(phen)$_2$(dppz)+oligo-7H]$^{5-}$	N.O.	2195.9
[2distamycin+3Ru(phen)$_2$(dppz)+oligo-11H]$^{5-}$	2590.0	N.O.
[2distamycin+3Ru(phen)$_2$(dppz)+oligo-12H]$^{6-}$	2158.2	N.O.
[2distamycin+3Ru(phen)$_2$(dppz)+oligo-13H]$^{7-}$	1850.0	N.O.
[2distamycin+4Ru(phen)$_2$(dppz)+oligo-13H]$^{5-}$	2738.3	N.O.
[2distamycin+4Ru(phen)$_2$(dppz)+oligo-14H]$^{6-}$	2281.7	N.O.
[3distamycin+3Ru(phen)$_2$(dppz)+oligo-12H]$^{6-}$	2237.7	N.O.
[4distamycin+Ru(phen)$_2$(dppz)+oligo-8H]$^{6-}$	2071.2	N.O.

Appendix 3 *Relative abundances of non-covalent complexes obtained from reaction mixtures containing* a 6:1 ratio of *ruthenium compound and duplex: (a) D1 and (b) D2.*◆ *$[Ru(phen)_3)]^{2+}$;* ■ *$[Ru(phen)_2(dpq)]^{2+}$;* ▲ *$[Ru(phen)_2(dpqC)]^{2+}$;* × *$[Ru(phen)_2(dppz)]^{2+}$;* ○ *$[Ru(phen)_2(dpqMe_2)]^{2+}$;* △ *$[Ru(phen)_2(pda)]^{2+}$.*



Appendix 4 *Calculation for K_d of the first metal binding to ε186.*

$\varepsilon186 \cdot Me_A \rightleftharpoons \varepsilon186 + Me_A$ Me_A = first metal ion bound to the protein

For $[Mn^{2+}]_{initial}$ = 12 μM; all $[\varepsilon186]_{initial}$ = 2 μM

	Species		
	ε186·Me_A complex	**ε186**	**Me_A**
Initial concentrations (M)	0	2×10^{-6}	12×10^{-6}
Equilibrium concentrations (M)	$0.2040^{a} \times 2 \times 10^{-6} = 4.08 \times 10^{-7}$	$0.7960^{a} \times 2 \times 10^{-6} = 1.59 \times 10^{-6}$	$12 \times 10^{-6} - 4.08 \times 10^{-7} = 11.60 \times 10^{-6}$

[a] Relative intensity calculated from the ESI mass spectrum.

$$K_d = \frac{[\varepsilon186] \times [Me_A]}{[\varepsilon186 \cdot Me_A]}$$

$$= \frac{1.59 \times 10^{-6} \times 11.60 \times 10^{-6}}{4.08 \times 10^{-7}}$$

$$= 4.52 \times 10^{-5} \text{ M}$$

Appendix 5 *Calculation for the specific enzymatic activity of ε186*

For 0.5 mM Mn^{2+}, [*p*NP-TMP] 3 mM; A_{420} slope = 0.1391 min^{-1}

A = εcl

ε = extinction coefficient at 420 nm of *p*-nitrophenol at pH 8.0 (12,950 $M^{-1}cm^{-1}$)

c = 0.1391 min^{-1}/ 12950 $M^{-1}cm^{-1}$ x 1 cm)

= 1.074 x 10^{-5} M min^{-1} in 1 mL

= 1.074 x 10^{-8} moles min^{-1}

= 1.074 x 10^{-2} μmoles min^{-1}…… (1)

ε186 in the cuvette = 2.1 x 10^{-3} mg…… (2)

Thus ε186 activity = (1) / (2)

= 1.074 x 10^{-2} μmoles min^{-1} / 2.1 x 10^{-3} mg

= 5.11 μmoles min^{-1} mg^{-1} ; 1 unit (U) is μmoles min^{-1}

= 5.11 Umg^{-1}