University of Alberta

Development and Application of ESI-MS Based Techniques to Study Non-Covalent Protein-Ligand Complexes in Solution and the Gas Phase

by

Lu Deng

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctoral of Philosophy

Department of Chemistry

Spring 2013
Edmonton, Alberta

Abstract

This thesis describes the development and application of electrospray ionization mass spectrometry (ESI-MS) based techniques to study noncovalent protein-ligand complexes in solution and the gas phase.

The application of the direct ESI-MS assay for quantifying the stoichiometry and absolute affinity of protein-metal ion binding *in vitro* was described. The direct ESI-MS assay was also used to determine the dissociation rate constants (k_{off}) for the model high affinity interaction between biotin (B) and the homotetramer of natural core streptavidin (S_4) at pH 7 and temperatures ranging from 15 to 45 °C. Importantly, the dissociation activation energies determined by ESI-MS agree, within 1 kcal mol^{-1}, with the reported value using a radiolabeled biotin assay. In addition to providing a quantitative measure of k_{off}, the results of the ESI-MS measurements revealed that sequential binding of B to S_4 occurs in a non-cooperative fashion with the four ligand binding sites being kinetically and thermodynamically equivalent and independent.

The structure and stability of the gaseous protonated ions of the (S_4 + 4B) complexes were investigated using various experimental and computational methods. Rate constants were determined for the dissociation of $(S_4 + 4B)^{n+}$ ions using time-resolved blackbody infrared radiative dissociation (BIRD). These kinetic data, together with results of ion mobility measurements and molecular dynamics simulations suggest that significant structural changes do not occur upon transfer of the complexes from solution to the gas phase by ESI and at least

some of the specific intermolecular interactions are preserved in the gas phase. Comparison of the dissociation kinetics for the gaseous $(S_4 + 4B)^{13+}$ ions with those determined for the $(S_4 + 4B)$ complexes in solution provides evidence for a late dissociative transition state and the rehydration of B and protein binding cavity in solution during dissociation. Intermolecular interactions in $(S_4 + 4B)^{13+}$ and $(\mathrm{scFv+L1})^{10+}$ ions were investigated using a collision-induced dissociation (CID)-functional group replacement (FGR) strategy. Comparison of the results obtained by the CID-FGR approach with those determined by the BIRD-FGR approach suggests that the CID-FGR method does not represent a reliable approach for identifying interactions in the gaseous protein-ligand complexes.

Acknowledgement

First and foremost, I would like to give my genuine gratitude and appreciation to my advisor, Professor John S. Klassen for his constant support and valuable guidance in the last five years of my doctoral study. I also would like to thank my supervisor committee members, Prof. Christopher Cairo, Prof. Julianne M. Gibbs-Davis, Prof. Mariusz Klobukowski, Prof. Xinfang Li and Prof. Don J. Douglas for their thoughtful and helpful comments into my research.

I am heartily thankful to Dr. Elena N. Kitova for her helpful advice and continued encouragements. I thank Dr. Michele Richards, Dr. Pavel I. Kitov and Aron Broom for their help with molecular dynamics simulations. I thank Ruixiang Blake Zheng for making the plasmids of streptavidin mutants. I am very grateful to all my fellow labmates in the Klassen group, past and present, for being so nice and helpful at all times. Special thanks go to Dr. Rambod Daneshfar, Dr. Glen Shoemaker, Dr. Lan Liu, Dr. Amr El-hawiet, Dr. Naoto Soya, Nian Sun, Yixuan Zhang, Hong Lin and Ling Han.

I give sincere thanks to Lu Zou, Yao Zou, Li Xia, Wenmin Ye, Cao Cong, Feifei Fu, and my other friends for all their help and company.

Finally, I would like to thank my parents, my sister, and my fiancé, Jianzhao Huang, for always being by my side.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A	Pre-exponential factor
Ab	Abundance
B	Biotin
B	Magnetic field
BIRD	Blackbody infrared radiative dissociation
CID	Collision-induced dissociation
CRM	Charge residue model
DC	Direct current
E_a	Activation energy
ECD	Electron capture dissociation
EHSS	Exact hard sphere scattering
ELISA	Enzyme linked immunosorbent assay
ESI	Electrospray ionization
FA	Fatty acid
FGR	Functional group replacement
FT	Fourier transform
FTICR	Fourier transform ion cyclotron resonance
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IEM	Ion evaporation model
IMS	Ion mobility separation
IRMPD	Infrared multiphoton dissociation

ITC	Isothermal titration calorimetry
K_a	Binding constant
k_{on}	Association rate constant
k_{off}	Dissociation rate constant
L	Ligand
La	α-lactalbumin
Lg	β-lactoglobulin
L_{ref}	Reference ligand
Lyz	Lysozyme
L1	α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-α-Man*p*-OCH_3
MD	Molecular dynamics
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
m/z	mass to charge
NOESY	Nuclear overhauser effect spectroscopy
NMR	Nuclear magnetic resonance
P	Protein
PA	Hard sphere projection approximation
PBS	Phosphate buffered saline
P_{ref}	Reference protein
PSA	Projection superposition approximation
R	The ratio of total abundance

RF	Radio frequency
RF	Response factor
RMSD	Root mean square deviation
scFv	Single chain antibody of the monoclonal antibody Se155-4
SPR	Surface plasmon resonance
S_4	Homotetrameric streptavidin
t_D	Drift time
TM	Trajectory method
TOF	Time of flight
TRIS	2-Amino-2-hydroxymethyl-propane-1,3-diol
TS	Transition state
T-Wave	Travelling wave
Ubq	Ubiquitin
WT	Wild type
ω_c	Cyclotron frequency
Ω	Collision cross section

Chapter 1

Study of Non-Covalent Protein-Ligand Complexes by Electrospray Ionization Mass Spectrometry

1.1 Introduction

Protein-ligand interactions play essential roles in many important biological processes, including the immune response, cell-cell communications, inflammation and bacterial and viral infections.[1-5] In order to fully understand the mechanisms by which proteins carry out their diverse biological functions, detailed insights into the structures of the protein-ligand complexes and the corresponding kinetic and thermodynamic parameters of the interactions are required.

There are a number of established analytical techniques used to characterize interactions of non-covalent biological complexes, each with particular strengths and weaknesses. Isothermal titration calorimetry (ITC) is one of the most widely used and most reliable methods for quantifying binding thermochemistry in solution. It is also the only technique that provides a direct measure of the enthalpy of association.[6-8] Conventional ITC method usually requires large amounts (~mg) of pure protein and ligand for each analysis and suffers from low throughput (5-10 measurements/day). Recently, new ITC technologies, such as the nanoITC, have improved sensitivity and require lower sample amounts.[9]

Surface plasmon resonance (SPR) spectroscopy is another commonly used quantitative method for measuring both kinetic[10,11] and thermodynamic parameters[12-14]. SPR affords high sensitivity, and requires a very small amount of sample(~ng). A potential limitation of SPR is the need to immobilize one of the binding partners (usually the ligand) on a sensor chip, which may affect the nature of binding interaction. Indeed, there are examples where the ITC and SPR spectroscopy yield divergent binding data for the same ligand-protein interaction.[15]

Enzyme linked immunosorbent assay (ELISA) is another widely used method for quantifying protein-ligand interactions.[16] While there are many ways of implementing ELISA, the assay typically requires the immobilization of ligand to a surface which is incubated with solutions containing the protein. The protein is often linked to an enzyme, and in the final step a substance containing the enzyme's substrate is added. The subsequent reaction produces a detectable signal. The ELISA method, once setup, is fast and relatively sensitive. However, the large immobilization surface area can lead to nonspecific binding and increased background. Also, most ELISAs rely upon enzyme-mediated amplification of signal to achieve reasonable sensitivity.

Spectroscopic approaches (e.g. dichroism or fluorescence-based approaches) [17-21] are very sensitive and are also commonly used to characterize the interactions of macromolecular complexes. However, the spectroscopic methods require spectroscopic activity of the analyte, and thus often require the introduction of a chromophore or fluorophore which can complicate the interpretation of the data and, in some instances, influence the rates of the reactions being investigated.[22]

X-ray crystallography represents the most widely used method to establish the higher-order structure of biological complexes and to identify interactions between binding partners.[23,24] Unfortunately, not all biomolecules and their complexes are readily crystallized. Furthermore, there is the possibility that the interactions present in the crystalline state are different from those in solution.[25] Additionally, X-ray crystallography analysis does not provide a direct measurement of the strength of the non-covalent interactions.

Nuclear magnetic resonance (NMR) spectroscopy is also widely used to characterize the structures of biological molecules and their complexes in solution and to quantify the strength (usually reported as the dissociation constant, K_d) of the interactions.[25-28] Recently, transferred NOESY NMR spectroscopy has been used to estimate the dissociation kinetics for streptavidin-small ligand.[29] However, NMR measurements are usually limited to relatively small proteins, with molecular weight (MW) up to ~40 kDa.[28] In addition, NMR measurements generally require large amounts of sample (typically mg quantities).

Electrospray ionization mass spectrometry (ESI-MS) has emerged as a promising tool for quantifying noncovalent protein interactions in solution.[30-35] The ESI-MS measurements can be categorized as either “direct” or “indirect” in nature. The direct ESI-MS assay, which is the focus of this thesis, is based on the detection of free and ligand-bound protein ions by ESI-MS leading to the determination of the association constant (K_a) from the abundance ratio of the bound and unbound protein ions measured for a solution of known initial concentrations of protein and ligand. The advantages of the direct ESI-MS assay include simplicity (no labeling or immobilization), speed (measurements normally can be completed in less than ~1 min), direct insight into binding stoichiometry and the ability to measure multiple binding equilibria simultaneously. Additionally, when performed using nanoflow ESI (nanoESI), which operates at solution flow rates in the nL/min range, the assay typically consumes picomoles or less of analyte per analysis. Besides its wide applications on quantifying noncovalent interactions, ESI-MS has emerged as an important addition to the arsenal of techniques available for measuring the kinetics of chemical and biochemical reactions.[36-47] Furthermore, gas-phase studies of desolvated noncovalent protein complexes by ESI-MS afford an opportunity to investigate their intrinsic properties, i.e., their structure and kinetic stability, free of solvent effects.[33,48-56] An overview of ESI-MS techniques, starting with instrumentation followed by three application areas (solution binding, solution kinetics as well as the gas phase structure and stability), is given below.

1.2 Electrospray ionization mass spectrometry

Before describing the different ESI-MS approaches for studying non-covalent biological complexes, it is useful to first review the basic principles of ESI-MS.

1.2.1 Electrospray ionization

ESI is a “gentle” ionization method that can be used to introduce biological molecules into the gas phase without fragmentation. In 1991, Henion and coworkers detected a non-covalent biological complex, a receptor-ligand complex consisting of a small hydrophilic immunosuppressive binding protein FKBP, with its specific ligand FK506, by ESI-MS for the first time.[57] Since then, ESI-MS has been used to detect a wide variety of non-covalent biological complexes, including antibody-antigen,[57] lectin-carbohydrate,[58] enzyme-substrate/inhibitor,[59] multiprotein[38,39,60] and DNA-ligand complexes[61,62].

Shown in Figure 1.1 is a diagram describing the ESI process involved in the formation of gas phase ions.

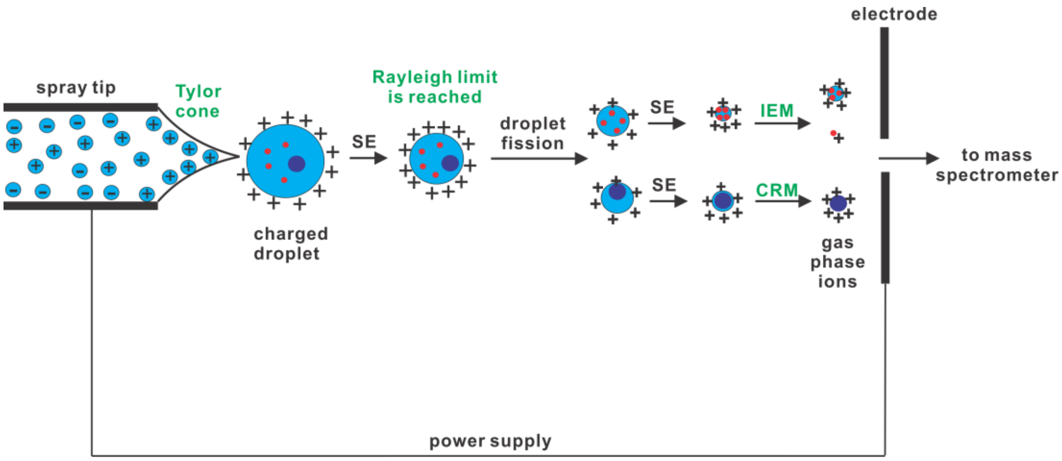


Figure 1.1. Schematic representation of ESI carried out in positive ion mode and the processes that lead to the formation of gas phase ions. SE, IEM and CRM here stand for solvent evaporation, ion evaporation model and charge residue model, respectively.

In ESI-MS, analyte solution is loaded into an ESI emitter, typically a stainless steel capillary. The mechanism of the ESI process[63] involves three steps:

(a) The production of charged droplets.

By applying a high voltage to the emitter, the electric field induces charge separation in the solution, whereby the charged electrolytes drift towards the liquid surface leading to the formation of a liquid cone referred to as a Taylor cone.[64] The increase of surface area due to the cone formation is resisted by the surface tension of the liquid. Under a sufficiently high field, the liquid cone becomes unstable and a fine jet emerges from the cone tip. In the positive mode, the surface of the jet, which is charged by an excess of positive ions, will break up into small charged droplets due to the repulsion between the charges.

(b) The shrinkage of the droplets by solvent evaporation accompanied with droplet fissions.

The charged droplets produced at the spray needle will shrink due to solvent evaporation causing an increase in the electric field normal to the surface of the droplet while the charge remains constant. As the charge density on the droplet surface increases to near the Rayleigh limit, the point at which the Columbic repulsion of the surface charges is equal to the surface tension of the droplets, these droplets undergo Rayleigh fission, eventually forming small highly charged offspring droplets.

(c) The formation of gas phase ions.

Two mechanisms have been proposed to account for the formation of gas-phase ions from the very small and highly charged droplets: the ion evaporation model (IEM) and the charge residue model (CRM). The ion evaporation model, proposed by Iribarne and Thomson,[65,66] assumes ion emission directly from very small and highly charged droplets and accounts for the production of gas phase ions of small molecules. The charge residue model, proposed by Dole,[67] assumes that the droplets undergo many fission events and finally produces very small droplets containing a single analyte molecule. These droplets undergo further solvent evaporation leading to the production of gas phase ions. It is generally believed that CRM most accurately describes the formation of gas phase macromolecule ions.

Recently Michael Gross and co-workers[68] proposes a modification of CRM in which CRM is preceded by IEM. This mechanism is expected to operate when salt additives (buffers) such as ammonium acetate or triethylacetate are present in millimolar concentrations in the solution that is electrosprayed.

In the present work, nanoESI was used in all cases. The mechanism of nanoESI is the same as that of ESI, except that by using a narrow glass tip, nanoESI operates at lower solution flow rates, (10-50 nL/min) and correspondingly emits smaller droplets than conventional ESI (1-10 μL/min).[69] Thus, only picomoles or less of analyte are required per analysis, a very important feature in the analysis of biological molecules where there is often only a limited amount of sample. In addition to being more sensitive than ESI, nanoESI readily allows the transfer of non-covalent complexes from buffered aqueous solutions to the gas phase and, therefore, can be directly performed on complex solutions that more closely resemble physiological conditions. The short lifetimes of the nanoESI droplets, which are estimated to be tens of microseconds, are also likely advantageous for preserving the original solution composition through the formation of gaseous ions. Furthermore, nanoESI can minimize nonspecific aggregation that can occur during the ESI process as there are fewer analyte molecules per droplet.[70,71] These features of nanoESI make it the method of choice for investigating non-covalent complexes directly by MS.

1.2.2 MS instrumentation

There are a number of different types of mass analyzers including magnetic sector, quadrupole, ion trap, fourier transform ion cyclotron resonance (FTICR) and time of flight (TOF). In the present study, nanoESI combined with FTICR and hybrid quadrupole-ion mobility separation (IMS)-TOF mass spectrometers were used.

1.2.2.1 Fourier transform ion cyclotron resonance mass spectrometer

Shown in Figure 1.2a is a schematic diagram of the Bruker Apex-II nanoESI-FTICR mass spectrometer used in the present work (Chapters 2 and 4). Briefly, fine droplets produced by nanoESI are introduced into the mass spectrometer through a heated metal capillary, which facilitates the desolvation of the droplets. The resulting gaseous ions are transmitted through a skimmer and accumulated in the hexapole for a certain time period to enhance the signal-to-noise ratio. After accumulation, ions are ejected from the hexapole, accelerated by a high voltage through the fringing field of a 9.4 T superconducting magnet, decelerated, and trapped by a combination of electric and magnetic field in FTICR cell for detection. The typical gas pressure for the instrument is ~5×10^{-10} mbar, maintained by the differential pumping system.

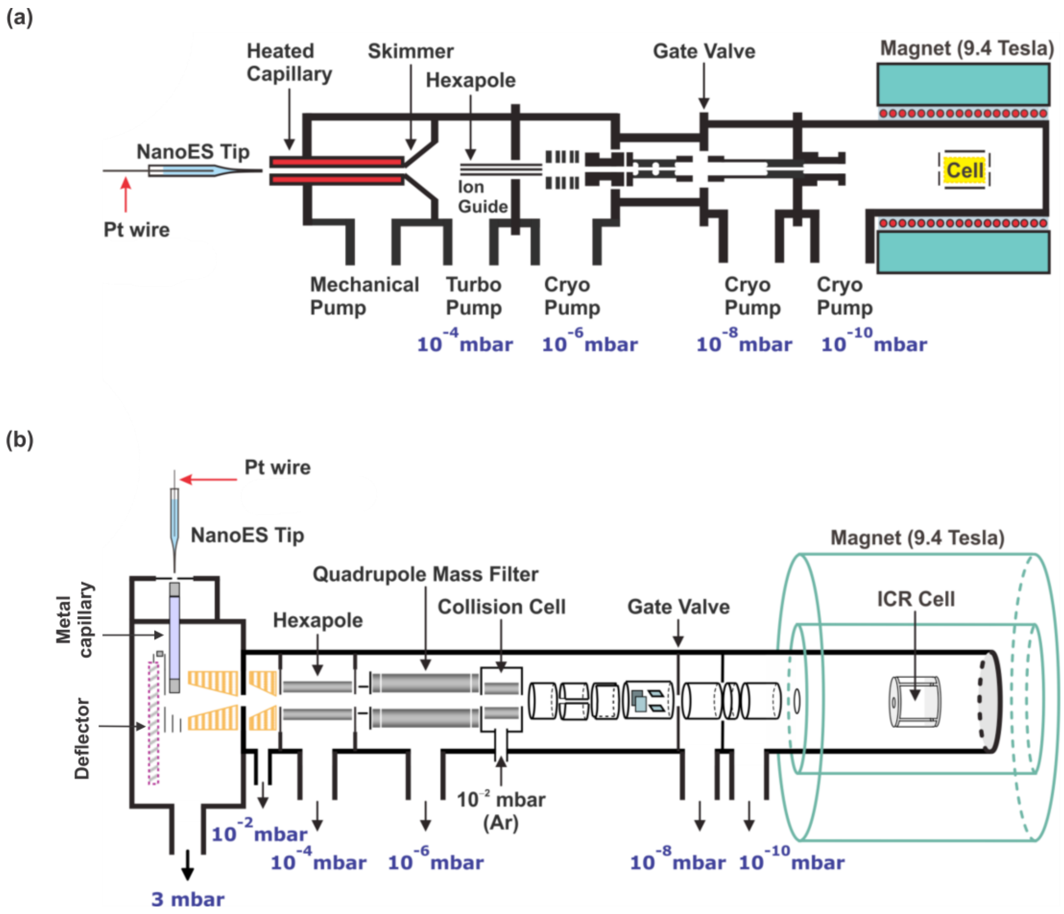


Figure 1.2. Schematic diagrams of (a) the Bruker Apex-II nanoESI-FTICR mass spectrometer and (b) the Bruker Apex-Qe nanoESI-FTICR mass spectrometer used in this study. Figures were reproduced from the Bruker user's manual.

The other FTICR mass spectrometer used in this thesis (Chapter 3) is a Bruker Apex-Qe nanoESI-FTICR mass spectrometer (Figure 1.2b). The Apex-Qe is a hybrid quadrupole-FTICR mass spectrometer, in which the quadrupole can act as a mass filter to isolate targeted ions for tandem MS (MS/MS) analysis. In the present work, however, the quadrupole was operated in radio frequency (RF)-only mode as it served as a wide band-pass filter to transmit ions. The operation scheme of the Apex-Qe is very similar to Apex-II, however the ion source represents the main difference. The ions generated in the ESI process, with the assistance of a neublizer and counter-drying gas, enter the vacuum system of the Apex Qe through a metal capillary. From the capillary exit, ions enter the transmitting funnels and skimmers orthogonally with the assistance of a deflector. The ions are then stored electrodynamically in the hexapole followed by further accumulation in the quadrupole. After accumulation, the ions are transferred through a series of ion optics into the ICR cell for detection.

FTICR mass spectrometers were used in this study for their high resolving power and mass accuracy. The general operating principles of FTICR are described in many reviews[72-74] and, therefore, only a brief overview is given here. The ICR cell consists of three pairs of plates (trapping, excitation and detection) and is located inside a spatially uniform static superconducting high field magnet cooled by liquid helium and liquid nitrogen. When the ions pass into the magnetic field they are bent into a circular motion in a plane perpendicular to the field (see Fig. 1.3) by the Lorentz force. The cyclotron frequency, ω_c is expressed in eq 1.1:

$$\omega_c = \frac{qB}{m} = \frac{zeB}{m} \tag{1.1}$$

where ω_c is the cyclotron frequency, q is the charge of the ion (q = ze, where z and e are the charge and the elementary charge, respectively), B is the magnetic strength and m is the mass of the ion. A notable feature of eq. 1.1 is that all ions of a given m/z rotate at the same frequency, independent of their velocities.

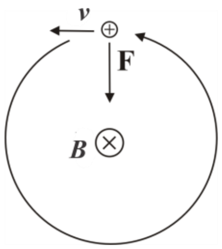


Figure 1.3. Illustration of the cyclotron motion of a positive ion of charge q moving with a velocity (*v*) in the presence of constant magnetic field (*B*) directed orthogonal to the motion of the ion. The ion experiences a Lorentz force F = q ($v \times B$), which causes the ion to move in a counterclockwise orbit.

At this stage, no signal is observed because the radius of the motion is very small. In order to produce a signal for the ions trapped in the ICR cell, excitation of each individual *m/z* is achieved by a swept RF pulse across the excitation plates of the cell. If the frequency of the applied field matches their ω_c, the ions will absorb energy and thus move circularly with a bigger orbital radius but the same ω_c. Shown in Figure 1.4 is the spiral trajectory of excited ions with the same *m/z* and ω_c. As they coherently pass the detection plates of the cell which are parallel to the magnetic axis, the orbiting ions induce an alternating current (image current) on the plates (Figure 1.4). The amplitude of this image current is proportional to the number of ions in the ICR cell while the frequency of the alternating current matches the ω_c of the ions. FT transforms the detected image current from the time domain signal into the frequency domain and a mass spectrum can be generated as ω_c is related to *m/z*.

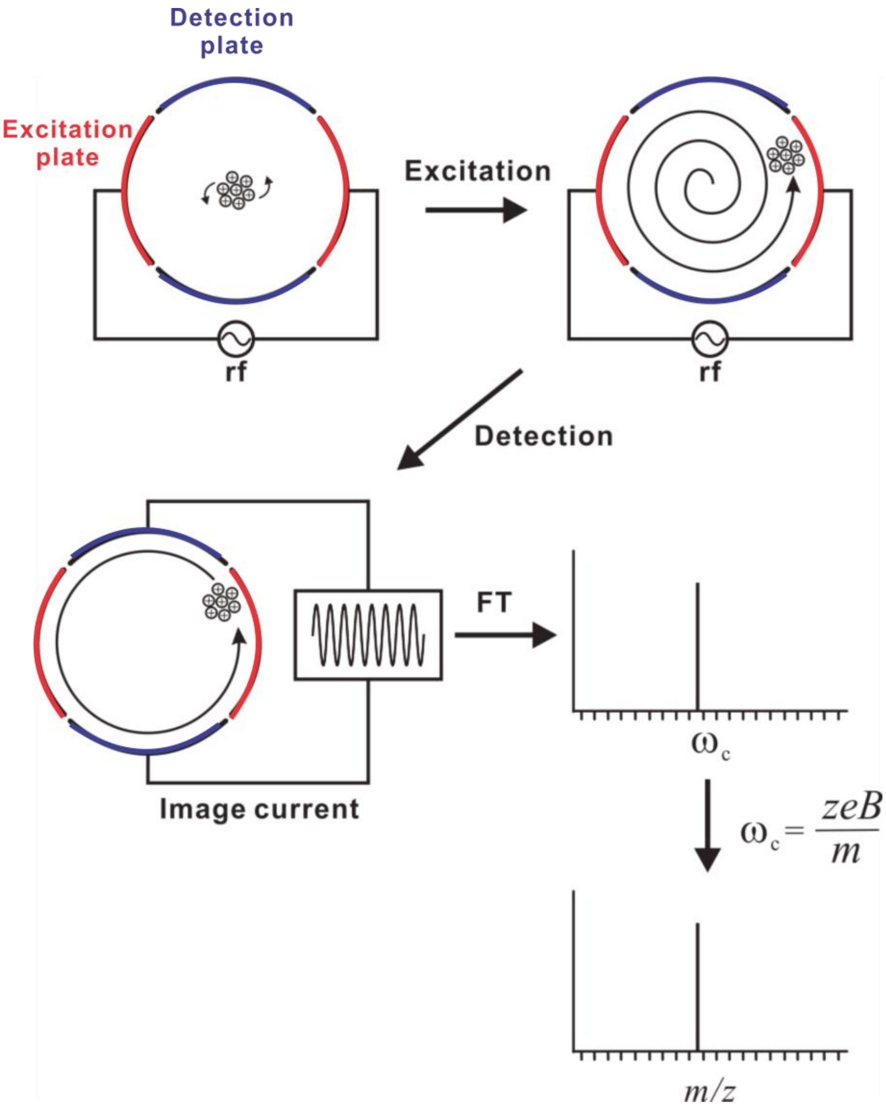


Figure 1.4. Illustration of excitation, image current detection and the production of mass spectrum by FTICR.

A notable feature of FTICR is its ultrahigh resolution which comes from the fact that all ions of a given m/z rotate at the same ω_c, independent of their kinetic energy. The resolution of FTICR mass spectrometer can routinely reach hundreds of thousands. The resolving power is proportional to the magnetic field strength (with higher magnetic field having higher resolution), and the acquisition time.

The acquisition time is the duration of the detection phase, determined by the dataset size and the frequency of sampling. Longer acquisition time (larger dataset size) results in higher resolution. Therefore, high vacuum (10^{-10} mbar) is necessary in the cell region of FTICR mass spectrometer, to avoid collisions with gas particles and deactivation of the ions.

1.2.2.2 Quadrupole-IMS-TOF mass spectrometer

Shown in Figure 1.5a is a schematic diagram of the Synapt G2 quadrupole-IMS-TOF mass spectrometer (Waters UK Ltd., Manchester, UK) equipped with a nanoESI source used in this work (Chapter 4). Briefly, fine droplets produced by nanoESI are introduced into the mass spectrometer through a “Z-spray source”, which minimizes neutral contamination and enhances the signal-to-noise. The resulting gaseous ions are transmitted through a quadrupole mass filter to the ion mobility section of the instrument (Triwave). The mobility separated ions are then detected by an orthogonal acceleration TOF mass analyzer (QuanTOF) equipped with a high field pusher and a dual-stage reflection.

The other Waters Synapt G2-S nanoESI-quadrupole-IMS-TOF mass spectrometer used in this study (Chapter 5) is shown in Figure 1.5b. The operation scheme of the Synapt G2-S is very similar to Synapt G2, but a newer generation with improved “step-wave” design in the ion source for better transmission. It is suitable for proteomics experiments where the sample complexity can be reduced by the ion mobility separation.

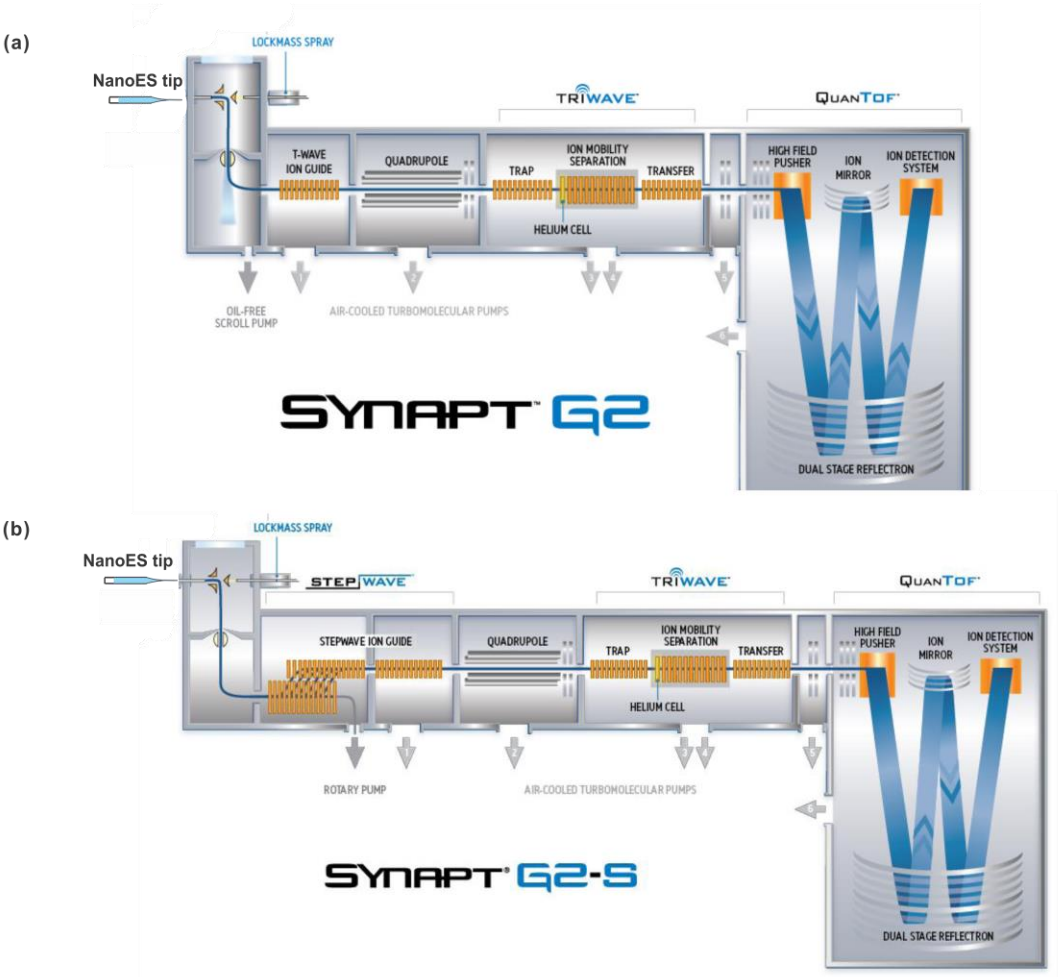


Figure 1.5. Schematic diagrams of (a) Waters Synapt G2 nanoESI-quadrupole -IMS-TOF mass spectrometer and (b) Waters Synapt G2-S nanoESI-quadrupole-IMS-TOF mass spectrometer used in this study. Figures were reproduced from the Waters user's manual.

The Waters Synapt G2 and G2-S nanoESI-quadrupole-IMS-TOF mass spectrometer were used in this study for their wide mass range, high sensitivity and high IMS efficiency which enables differentiating samples based on size, shape, and charge, as well as mass. Here a brief overview of three main parts of the Synapt mass spectrometer is given.

1.2.2.2.1 Quadrupole

The quadrupole consists of four cylindrical metal rods that are accurately positioned in a radial array and the diametrically opposed rods are paired. A direct current (DC) potential and a RF potential, 180 degrees out of phase, is applied to each pair of rods.[75] In the Synapt mass spectrometer, quadrupole contains two parts, a quadrupole prefilter followed by a quadrupole mass filter. (Figure 1.6). The use of prefilter increases the absolute sensitivity by minimizing the effects of fringing fields at the entrance to the quadrupole.[76]

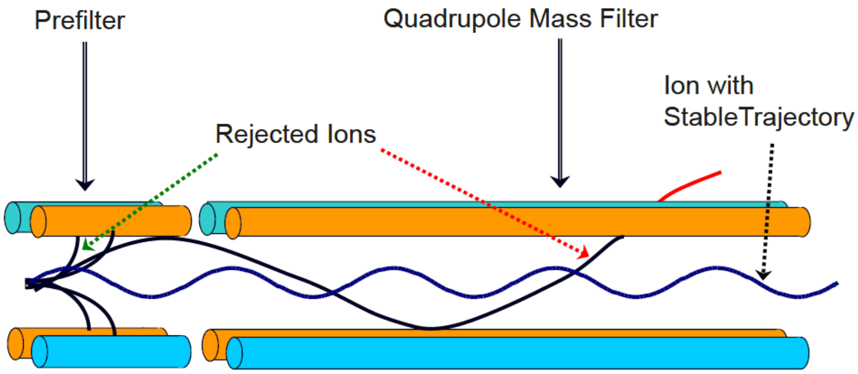


Figure 1.6. A schematic diagram of the quadrupole in the Waters Synapt mass spectrometer.

In the MS-TOF mode, the quadrupole operates in RF only mode that transmits and guides ions over a wide *m/z* range to later components of the apparatus. In the MS/MS-TOF mode, the quadrupole operates with both RF and DC. Depending on the specific voltage and frequency applied, ions of a particular *m/z* can be selected and transit down the entire length of the rods; other ions outside the *m/z* range hit the rods and are expelled.

1.2.2.2.2 Triwave

The Triwave section of the Waters Synapt G2 or G2-S mass spectrometer consists of three T-wave ion guides (Trap, IMS and Transfer) as shown in more detail in Figure 1.7a. Each ion guide comprises a series of planar electrodes arranged orthogonally to the ion transmission axis[77], as shown in Figure 1.7b. Opposite phases of a RF voltage are applied to adjacent electrodes and provide a radially confining effective potential barrier. In the presence of a background gas, ion axial motion through the ion guide can be significantly slowed or stopped due to the presence of axial traps generated by the ring geometry. To propel ions through the gas, a transient DC voltage is superimposed on the RF applied to a pair of adjacent electrodes in a repeating sequence along the length of the device. The series of potential hills generated are subsequently applied to the next pair of electrodes downstream at regular time intervals providing a continuous sequence of T-waves. The ions within the device are driven away from the potential hills and consequently are carried through the device with the waves, minimizing their transit time. Ions with high mobility are rapidly pushed through the device by the T-Waves. Ions with low mobility experience greater “friction” from the bath gas causing them to be retarded.

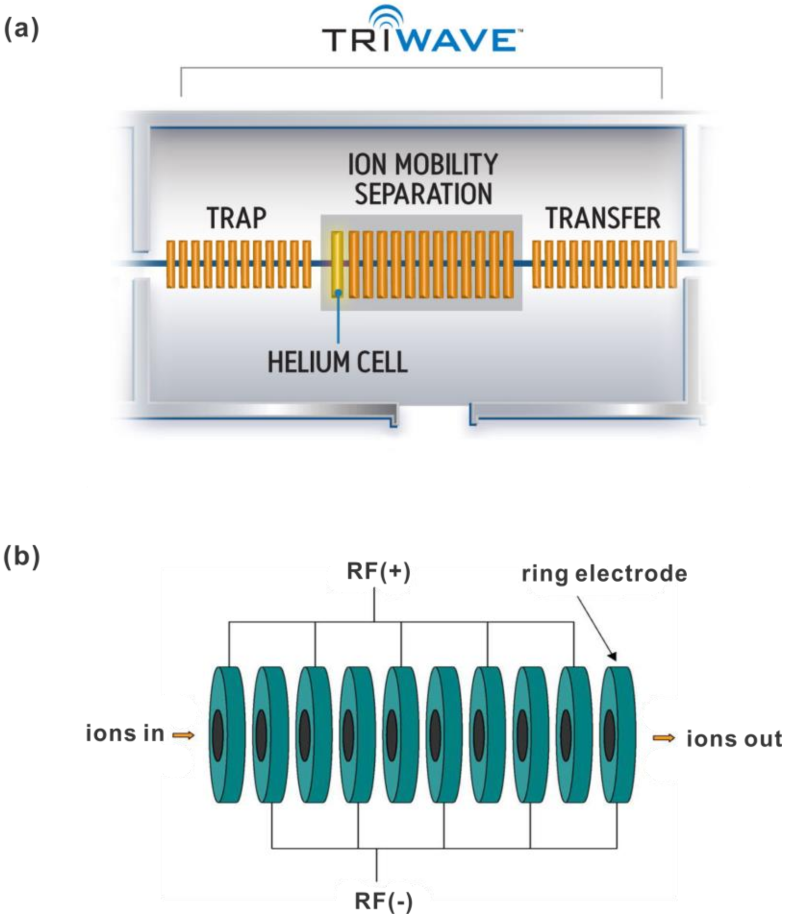


Figure 1.7. (a) A schematic diagram of the IMS section of the Synapt, comprising three travelling wave ion guides labelled as TRAP, ION MOBILTY SEPARATION and TRANSFER. (b) A stacked ring ion guide.

In the MS-TOF mode, the Trap T-Wave traps and accumulates ions, after which the ions are released in a packet into the IMS T-Wave, where IMS is performed. A high-pressure helium-filled cell at the front of the IMS T-Wave cell is used to minimize scattering and/or fragmentation by collisionally cooling transferred ions. The Transfer T-Wave delivers the mobility separated ions to the

QuanTOF for detection. In the MS/MS-TOF mode, collision-induced dissociation (CID) can be performed in both the trap and transfer regions (~10^{-3} mbar).

1.2.2.2.3 TOF

For TOF analyzers, the physical property that is measured during an analysis is the flight time of the ions.[78] *m/z* values are determined by measuring the time that ions take to move through a field-free region (flight tube) between the source and the detector, according to eq 1.2:

$$m / z = t^2 \frac{2eV_s}{L^2} \tag{1.2}$$

where *m* is the mass of the ion, *z* is the charge state of the ion, e is the elementary charge, V_s is the acceleration potential, t is the flight time and L is the length of the flight tube. This equation shows that *m/z* can be calculated from a measurement of t. The lower the mass of an ion, the faster it will reach the detector. There are two types of TOF analyzers: linear TOF analyzer and reflectron TOF analyzer. The linear TOF analyzer has the drawback that ions of the same *m/z* may reach the detector at different times, due to the initial energy distribution, resulting in peak broadening and poor resolution.

In Waters Synapt G2 or G2-S mass spectrometer, a reflectron TOF analyzer is used. The reflectron TOF analyzer compensates the energy distribution of ions by using successive sets of electric grids of increasing potential which deflects the ions and reverses their flight direction sending them back through the flight tube. Depending on their kinetic energy, ions of the same *m/z* will penetrate the field at different depths; ions with more kinetic energy and hence with greater velocity will penetrate the field more deeply than ions with lower kinetic energy. Consequently, the faster ions will spend more time in the reflectron and reach the detector at the same time as the slower ions with the same *m/z*. The net effect improves mass resolution with minimal losses in sensitivity.

After reviewing the basic principles and instrumentations of ESI-MS, an introduction of three areas of studying non-covalent protein-ligand interactions by ESI-MS is given in the following sections 1.3-1.5.

1.3 MS-based methods for quantifying interactions of non-covalent complexes

1.3.1 Direct ESI-MS binding assay

The direct ESI-MS binding assay is based on the direct detection of free and ligand-bound protein ions by ESI-MS. For a 1:1 protein-ligand complex (eq 1.3), the K_a is calculated using eq 1.4:

$$\mathrm{P} + \mathrm{L} \rightleftharpoons \mathrm{PL} \tag{1.3}$$

$$\mathrm{K_a} = \frac{[\mathrm{PL}]_{eq}}{[\mathrm{P}]_{eq}[\mathrm{L}]_{eq}} \tag{1.4}$$

Here $[\mathrm{PL}]_{eq}$, $[\mathrm{P}]_{eq}$ and $[\mathrm{L}]_{eq}$, are the equilibrium concentrations of PL complex, free P and L, respectively, and can be calculated from the initial concentrations of P and L in solution, $[\mathrm{P}]_o$ and $[\mathrm{L}]_o$, and the concentration ratio of PL and P at equilibrium (eqs 1.5-1.7). The ratio (*R*) of the total abundance (*Ab*) of PL and P ions (e.g., PL^{n+}, P^{n+}) measured in the gas phase by ESI-MS is expected to be equivalent to the ratio of the concentrations of PL and P in solution at equilibrium.[79]

$$\frac{[\mathrm{PL}]_{eq}}{[\mathrm{P}]_{eq}} = \frac{\sum_n Ab(\mathrm{PL}^{\mathrm{n+}})}{\sum_n Ab(\mathrm{P}^{\mathrm{n+}})} = R \qquad (1.5)$$

$$[\mathrm{P}]_{\mathrm{o}} = [\mathrm{P}]_{eq} + [\mathrm{PL}]_{eq} \qquad (1.6a)$$

$$[\mathrm{L}]_{\mathrm{o}} = [\mathrm{L}]_{eq} + [\mathrm{PL}]_{eq} \qquad (1.6b)$$

$$[\mathrm{PL}]_{eq} = \frac{R[\mathrm{P}]_{\mathrm{o}}}{1+R} \qquad (1.7)$$

Then K_a value for the 1:1 protein-ligand complex is determined from R, $[\mathrm{P}]_o$ and $[\mathrm{L}]_o$, eq 1.8.

$$\mathrm{K_a} = \frac{R}{[\mathrm{L}]_o - \frac{R}{1+R}[\mathrm{P}]_{\mathrm{o}}} \qquad (1.8)$$

When the protein (or protein assembly) can bind to N ligands (where $N >$ 1), there are N reactions to be considered:

$$\mathrm{P + L \rightleftharpoons PL} \qquad (1.9a)$$

$$\mathrm{PL + L \rightleftharpoons PL_2} \qquad (1.9b)$$

$$\mathrm{PL_2 + L \rightleftharpoons PL_3} \qquad (1.9c)$$

$$\ldots$$

$$\mathrm{PL_{N-1} + L \rightleftharpoons PL_N} \qquad (1.9d)$$

Here, we describe only the simplest case, in which all N binding sites are equivalent, with identical binding constants. The treatment of more complicated cases has been discussed elsewhere.[80] The equilibrium concentrations, $[\mathrm{P}]_{eq}$, $[\mathrm{PL}]_{eq}$, …, $[\mathrm{PL}_N]_{eq}$, can be determined from relative abundance of the corresponding ions observed in the mass spectrum and eq 1.10a. Then, using these values, the equilibrium concentration of L can be found from eq 1.10b:

$$[\mathrm{P}]_\mathrm{o} = [\mathrm{P}]_{eq} + [\mathrm{PL}]_{eq} + [\mathrm{PL}_2]_{eq} \cdots + [\mathrm{PL}_N]_{eq} \tag{1.10a}$$

$$[\mathrm{L}]_\mathrm{o} = [\mathrm{L}]_{eq} + [\mathrm{PL}]_{eq} + 2[\mathrm{PL}_2]_{eq} + \cdots + N[\mathrm{PL}_N]_{eq} \tag{1.10b}$$

K_a can be determined from eq 1.11, which are based on the general expression $K_i = K_a (N - i + 1)/i$, where i is the number of occupied binding sites[80]:

$$K_i = \frac{[\mathrm{PL}_i]_{eq}}{[\mathrm{PL}_{(i-1)}]_{eq}[\mathrm{L}]_{eq}} \tag{1.11}$$

An average K_a can be determined from the binding constant determined for each of the binding reactions.

In practice, ESI-MS binding measurements are usually limited to R values ranging from approximately 0.05 to 20 and P and L concentrations in the 0.1 to 1000 μM range. It follows that K_a values accessible with the direct ESI-MS binding assay range from approximately 10^3 to 10^7 M^{-1}.[35] However, interactions with much larger K_a values can be probed using competitive binding and direct ESI-MS measurements.

1.3.2 Potential pitfalls of the direct ESI-MS binding assay

As shown in eq. 1.5, the successful implementation of the direct ESI-MS assay requires that the equilibrium abundance ratio of bound-to-free protein present in MS is the same as the equilibrium concentration ratio. Physical or chemical processes that alter this ratio will lead to incorrect K_a values and, potentially, obscure the true binding stoichiometry. Four common sources of error associated with the ESI-MS measurements are: (1) non-uniform response factors, (2) in-source dissociation, (3) nonspecific ligand–protein binding, and (4) ESI-induced changes in solution, pH and temperature. For the purpose of this study, the first three sources of error are briefly described below, along with current strategies for minimizing their effects on the binding measurements.

1.3.2.1 Non-uniform response factors

The abundance of each protein species measured by ESI-MS is related to the solution concentration by its response factor (*RF*), which collectively accounts for the ionization and detection efficiencies:[81]

$$\frac{[\mathrm{PL}]_{eq}}{[\mathrm{P}]_{eq}} = \frac{Ab(\mathrm{PL})\,/\,RF_{PL}}{Ab(\mathrm{P})\,/\,RF_{P}} = RF\,\frac{Ab(\mathrm{PL})}{Ab(\mathrm{P})} \qquad (1.12)$$

An essential assumption in the above expression is that the *RF* values for P and PL ions are equal. This assumption is generally valid in cases when the L is small compared to the P, such that the size and surface properties of the P and PL are similar[81-84] and is applied in the present work. However, non-uniform ESI-MS response factors are generally expected in the case of protein–protein interactions. *RF*s depend on many factors (the size and structure of P and PL, the ESI conditions and the instrumental parameters used for the measurements) and it is often possible to “tune” the experimental conditions to achieve the correct *R* value (based on the known K_a). Such an approach may be appropriate when investigating structurally-related interactions. However, it is not generally recommended and should be used with caution. Several approaches have been developed to minimize the effects of non-uniform *RF*s on the determination of K_a values.[62,85,86]

1.3.2.2 In-source dissociation

Collision-induced dissociation of the gaseous PL complexes in the ion source can alter the relative abundance of PL and P ions.[32] For a 1:1 P–L complex, in-source dissociation will necessarily decrease the magnitude of K_a. In the extreme case, where no PL ions survive to detection, in-source dissociation results in a false negative. The influence of in-source dissociation depends on the configuration of the ion source used, the choice of instrumental parameters and the gas-phase stability of the complex being investigated. Usually, the occurrence of in-source dissociation can be identified from changes in *R* resulting from changes in ion source parameters, in particular voltage differences in regions of high pressure, that influence the internal energy of the ions.

In cases where the gaseous complexes are susceptible to in-source dissociation, low temperatures (drying gas, sampling capillary), low potentials across lens elements, and short accumulation times are essential for obtaining reliable binding constants. However, these conditions normally reduce signal intensities. Thus, a balance must be found between minimizing dissociation and obtaining mass spectra with sufficient signal-to-noise ratio. Another way to minimize the in-source dissociation is through the addition of small organic molecules, such as imidazole, to the ESI solution.[32,33] The origin of the stabilizing effects of imidazole is believed to be due, at least in part, to enhanced evaporative cooling resulting from the dissociation of nonspecific imidazole adducts from the gaseous PL ions.[32] Additionally, the use of imidazole, which has a relatively high gas phase basicity and a relatively low gas phase acidity[87], may also lead to a reduction in the charge states of the protein complex ions. Moreover, it was shown that the introduction of imidazole vapor to the ion source also protects complexes against in-source dissociation.[88] However, these approaches have their limitations as the detection of very labile gas phase complexes, which rapidly dissociate at ambient temperature, by ESI-MS remains challenging.

Recently a competitive binding assay, such as the reference ligand method, was developed to quantify protein-ligand interactions that are highly labile and cannot be detected by ESI-MS.[89] This method employs direct ESI-MS analysis in conjunction with a reference ligand (L_{ref}). The L_{ref} binds specifically to P, at the same binding site as L, with known affinity and forms a stable protein-ligand complex in the gas phase. The fraction of P bound to L_{ref}, which is determined directly from the ESI mass spectrum, is sensitive to the fraction of P bound to L in solution and enables the affinity of P for L to be determined. This method has proven particularly useful for the analysis of PL interactions that are kinetically unstable in the gas phase at room temperature.[90]

1.3.2.3 Nonspecific binding

It is well established that, during the ESI process, free L can bind nonspecifically to P and PL due to concentration effects, resulting in false positives. Consequently, the observation of gaseous ions corresponding to a particular PL complex does not, by itself, establish the presence of that interaction in solution. The observation of multiple ligands bound to the target protein with a Poisson-like distribution is a tell-tale sign of occurrence of nonspecific ligand binding. Changes in the magnitude of K_a with changes in ligand concentration may also alert to the occurrence of nonspecific ligand binding.

The formation of nonspecific protein–ligand complexes can be understood in the context of the CRM of ESI (Figure 1.8).

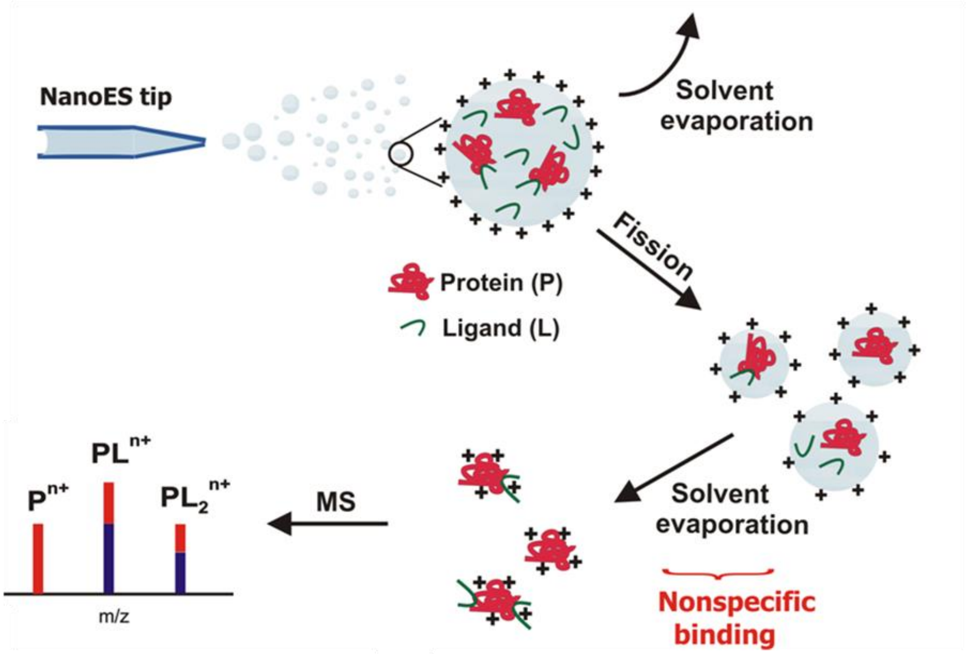


Figure 1.8. Cartoon of the CRM of ESI depicting the formation of nonspecific protein–ligand interactions (false positives).

According to the CRM, the initial ESI droplets undergo solvent evaporation until they come close to Rayleigh limit, at which point they undergo fission, releasing several small multiply charged nanodroplets (often referred to as offspring droplets) containing no analyte or one or more molecules of analyte. Solvent evaporation from the nanodroplets ultimately yields gaseous ions. If a nanodroplet contains two or more analyte molecules, nonspecific intermolecular interactions can occur as the droplet evaporates to dryness, leading to the formation of nonspecific complexes. Nonspecific binding of L to P and PL obscures the true binding stoichiometry in solution and introduces error into the K_a values measured by ESI-MS.

Generally, the formation of nonspecific protein–ligand complexes can be minimized by decreasing the initial concentration of ligand. For very weak ligand interactions ($K_a < 10^4$ M^{-1}), however, high initial concentrations of ligand are required to produce detectable level of complexes. In such cases, nonspecific binding is often unavoidable. A number of strategies have been developed to correct ESI mass spectra for the occurrence of nonspecific binding.[91-96] The most straightforward approach is the reference protein method (Figure 1.9) which involves the addition of a reference protein (P_{ref}) that does not bind specifically to the protein and ligand of interest to the solution.[92]

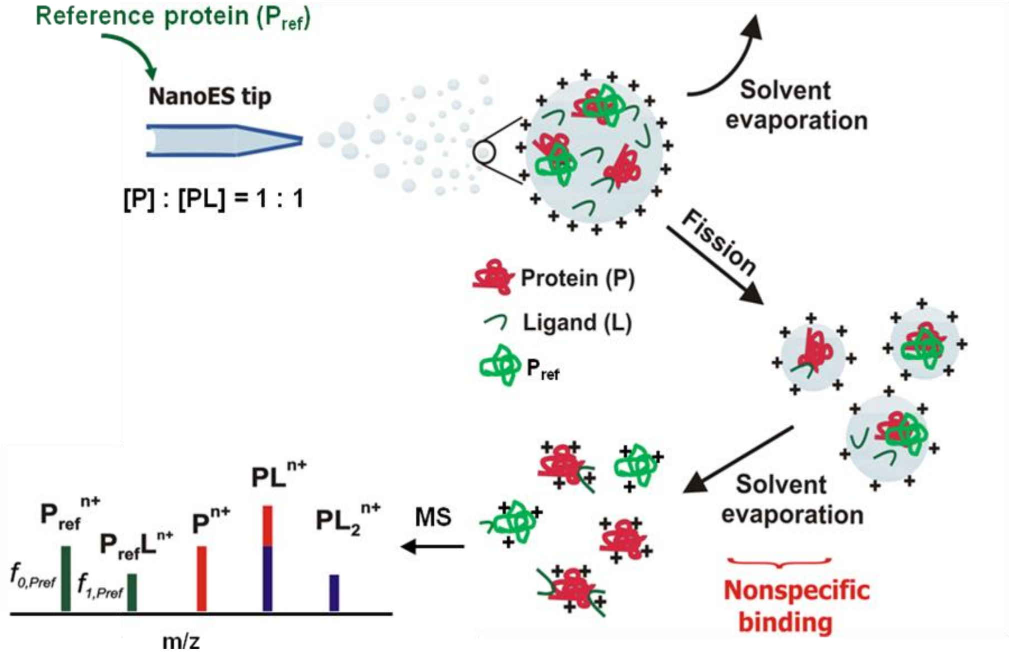


Figure 1.9. Cartoon of the CRM of ESI depicting the reference protein method.

The method is based on the assumption that nonspecific ligand binding is random, as suggested by the observation that the distribution of nonspecifically bound molecules often resembles that of a Poisson distribution, and affects equally all protein species present in the ESI droplets. The occurrence of nonspecific protein-ligand binding is monitored by the appearance of ions corresponding to nonspecific complexes of P_{ref} and L in the mass spectrum.[92] The fraction abundance of P_{ref} undergoing nonspecific ligand binding provides a quantitative measure of the contribution of nonspecific binding to the measured intensities of protein and specific protein-ligand complexes. The “true” abundance of a given PL_i species can be calculated from the apparent (measured) abundance of the PL_i species and the distribution of nonspecific $P_{ref}L_i$ species using the following expression:

$$Ab(PL_i) = [Ab_{app}(PL_i) - f_{1,Pref}Ab(PL_{i-1}) - \cdots f_{i,Pref}Ab(P)]/f_{0,Pref} \quad (1.13)$$

where $f_{i,Pref}$ is the fractional abundance of P_{ref} bound to i molecules of L. Notably, the method has been shown to correctly account for the nonspecific binding of neutral, acidic and basic biomolecules to proteins during ESI-MS analysis.[92-94] However, the reliability of the P_{ref} method for quantifying nonspecific proteins-metal ion interactions has not yet been tested and is the focus of Chapter 2.

1.4 MS-based methods for quantifying non-covalent complexes kinetics

Quantification of the kinetic parameters - the association and dissociation rate constants (k_{on} and k_{off}, respectively) - under specific solution conditions (e.g. pH, temperature, ionic strength) is important in understanding the structure and function of protein complexes and is relevant to drug design.[97-99] There exist a number of established experimental techniques for measuring the rates of biochemical reactions, including association and dissociation reactions. These include SPR,[10,11] spectroscopic methods (e.g., atomic force spectroscopy, circular dichroism or fluorescence-based approaches),[17-21] kinetic capillary electrophoresis,[100,101] radiolabeling combined with filtration/dialysis[102,103] and NMR.[29] Many of these techniques require the labelling of one of the binding partners, their attachment to a surface or some other manipulation of the system, which can complicate the interpretation of the kinetic data and, in some instances, influence the rates of the reactions being investigated.[22]

ESI-MS has emerged as an important addition to the arsenal of techniques available for measuring the kinetics of chemical and biochemical reactions.[36-47] The ESI-MS approach is attractive as there is no requirement for labelling or immobilization since the identity of reactants and products and, possibly intermediates, can usually be established directly from the measured *m/z*.[37,104] Moreover, ESI-MS analysis allows for multiple reactions to be monitored simultaneously, a feature not associated with most kinetic assays. The determination of reaction rates by ESI-MS analysis normally follows one of two general strategies: on-line (real-time) monitoring of the reaction mixture, and off-line analysis, usually following a quench step that stops the reaction. The advantage of the on-line approach is that it allows, in principle, for direct analysis of the time-dependent distribution of reactants, intermediates and products. The minimum acquisition time for an ESI mass spectrum, which is typically in the s - min range (although it varies between instruments and the nature of the sample being analyzed), places restrictions on the speed of reactions that can be reliably analyzed using the on-line approach. For this reason, real-time ESI-MS kinetic measurements are most commonly applied to relatively slow reactions, with timescales >min. However, there are examples where ESI-MS has been successfully applied to relatively fast reactions, in the ms – s range [36,37,39,41,43,44]. The measurement of fast kinetics typically requires the use of rapid mixing systems, such as a continuous-flow [36,37,39,41], rapid quenched-flow apparatus [43] or stopped-flow [44,45]. The off-line approach is generally easier to implement and affords greater flexibility in terms of the experimental conditions under which the reactions are carried out. For example, this approach is suitable for the analysis of enzyme kinetics under solution conditions that are not amenable to direct ESI-MS analysis, such as high concentrations of salts or non-volatile buffers (e.g. PBS, citrate, HEPES or TRIS) that are commonly used to stabilize proteins and ensure relevance to physiological conditions. Following the quench step, the solvent composition can be altered in order to facilitate detection of reactants or products by ESI-MS. However, a limitation of the off-line approach is that information on the distribution of species present under the reaction conditions of interest may be lost.

ESI-MS has been used to study the reaction rates for a variety of non-covalent protein interactions, including protein-protein[38,39,60], protein-small molecule[105] and protein-metal ion complexes[106], as well as for other biological complexes, such as DNA duplexes[61,62]. However, to the best of our knowledge, absolute values of k_{on} and k_{off} for protein-ligand interactions measured using this approach have not been previously reported and were the focus of Chapter 3.

1.5 MS-based methods for investigating structure and stability of non-covalent complexes in the gas phase

Besides the quantification of their solution kinetic and thermodynamic parameters by ESI-MS, there is also great interest in the use of gas phase studies to elucidate the structure and dynamics of biological complexes. In the absence of water, gas phase studies of desolvated biological complexes represent a promising experimental approach to probe directly the intrinsic (solute-solute) intermolecular interactions and, indirectly, the role of solvent in biological recognition. Two major MS based techniques, MS/MS and IMS-MS, are demonstrated to provide information about structure (topology), size and stability of gaseous proteins and protein complexes.

1.5.1 MS/MS

Although ESI-MS methods have wide applications in quantifying solution kinetic and thermodynamic properties of protein-ligand complexes, little structural information can be obtained from a single mass spectrum. This is because the soft ionization method, ESI, produces intact complex ions without fragments. With development of ion activation techniques, MS/MS has emerged as a useful tool for probing the structure and stability of biomolecules and their complexes in multiple stages of mass analysis. MS/MS experiments involve at least two stages of mass analysis on the gaseous ions. After isolation of precursor (parent) ions of interest in the first-stage of mass selection, the ions are then activated by various methods and undergo fragmentation. In the second-stage of mass analysis, the product ions from the fragmentation are detected. The resulting product ions can provide structural insights regarding the parent ion. Various ion activation techniques are available including blackbody infrared radiative dissociation (BIRD),[107-109] CID,[110-112] infrared multiphoton dissociation (IRMPD),[108,113] and electron capture dissociation (ECD).[114-117] For the purpose of this study, BIRD and CID techniques are discussed in a greater detail in the following sections.

1.5.1.1 BIRD

BIRD is a dissociation technique that allows ions to undergo unimolecular dissociation at essentially zero pressure by the exchange of energy with an ambient blackbody radiation field. BIRD experiments are performed in the ultra-low pressure, heated ion cell of the FTICR mass spectrometer (section 1.2.2.1). After ions are injected into the ICR cell, ions of interest are isolated and trapped within the ICR cell, which is essentially a collision-free environment. To perform BIRD experiments, flexible heating blankets are placed on the portion of the vacuum tube which surrounds the ion cell to control the temperature of the ion cell. Ions of interest are activated by the absorption of blackbody infrared photons emitted from the walls of the ion cell. After a certain period of time, parent ions as well as resulting product ions are excited and detected by the FTICR mass spectrometer.

The unimolecular dissociation reaction of a bimolecular ion, AB^+ (1.14), can be envisioned as a two-step mechanism.[107,109]

$$AB^+ \xrightarrow{k_{uni}} A^+ + B \tag{1.14}$$

$$AB^+ + h\nu \underset{k_{-1,rad}}{\overset{k_{1,rad}}{\rightleftharpoons}} AB^{+*} \tag{1.15a}$$

$$AB^{+*} \xrightarrow{k_d} A^{+} + B \qquad (1.15b)$$

The first step involves the excitation of AB^{+} to activated species, AB^{+*}, by the absorption of infrared photons (eq 1.15a). The activated species can either relax back into AB^{+} through the emission of infrared photons, or dissociate into products A^{+} and B (eq 1.15b). Under steady state conditions, the observed collisionless unimolecular rate constant (k_{uni}), is expressed in terms of radiative activation and deactivation rate constants ($k_{1,rad}$ and $k_{-1,rad}$) and unimolecular dissociation rate constant (k_d) , eq 1.16a.

$$k_{uni} = k_d \left(\frac{k_{1,rad}}{k_{-1,rad} + k_d} \right) \qquad (1.16a)$$

$$k_{uni} = k_{1,\,rad} \quad \text{if } k_d >> k_{-1,\,rad} \qquad (1.16b)$$

$$k_{uni} = k_d \left(\frac{k_{1,rad}}{k_{-1,rad}} \right) \quad \text{if } k_d << k_{-1,\,rad} \qquad (1.16c)$$

$$lnk_{uni} = -\frac{E_a}{RT} + \ln A \qquad (1.17)$$

For small weakly bound molecules in the zero-pressure limit, $k_d >> k_{-1,\,rad}$, and k_{uni} reflects only the rate of photon absorption ($k_{1,rad}$) (eq 1.16b).[107] In this case, no information about the dynamics of the dissociation process is obtained. For polymers such as peptides and DNA, the number of blackbody photons exchanged with chamber walls will increase approximately linearly with size. Meanwhile, dissociation rates decrease with increasing ion size for a fixed internal energy. Thus, $k_d << k_{-1,rad}, k_{1,rad}$ and k_{uni} is given by eq 1.16c, where $k_{1,rad}/k_{-1,rad}$ results in Boltzmann distribution.[107] In this case, activation energies

(E_a) and pre-exponential factors (A) can be determined from the temperature dependence of k_{uni}, eq 1.17.

The BIRD technique provides a unique route to measure accurate dissociation energetics and kinetics for large biological complex ions, including antibody-antigen,[50] enzyme-substrated/inhibitor,[49] lectin-carbohydrate,[118] and bacterial toxin multiprotein complexes.[118] Moreover, BIRD combined with a functional group replacement (BIRD-FGR) approach has been developed and used to identify individual interactions in gaseous biological complexes and to quantify their strength.[33,48,50,52] To establish whether a particular residue (on protein) or functional group (on ligand) is involved in binding, the group is modified (chemically or biochemically) in such a way that any pre-existing interaction is lost. The E_a for the loss of ligand from the modified complex is then compared to the E_a for the ligand loss from the unmodified complex. A decrease in E_a indicates that the particular residue/functional group (that was modified) was involved in stabilizing the complex. Moreover, the differences in E_a for the modified and unmodified complexes, *i.e.* $\Delta E_a = E_a(\text{unmodified}) - E_a(\text{modified})$, provides a measure of the strength of the interaction. Ideally, to identify H-bond donor/acceptor pairs, a three step approach is utilized, in which ΔE_a values are determined for complexes with modified ligand, modified protein and modified protein and ligand. For a given H-bond donor/acceptor pair, the magnitude of the ΔE_a values determined for the three complexes will be identical.[50] Using the BIRD-FGR method, the intermolecular interactions in gaseous ions of protein-ligand complexes composed of a 27 kDa single chain variable fragment of the monoclonal antibody Se155-4a (scFv), and its native trisaccharide ligand (α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-α-Man*p*-OCH_3, L1)[48,50,119], β-lactoglobulin (Lg), a water-soluble whey protein possessing a large hydrophobic cavity, and fatty acids (FAs) [33,51,52] were recently elucidated. Furthermore, insights into the solvent effect can be obtained by comparing E_a of gasoues ions to E_a of complexes in solution. In Chapter 4, we used BIRD to study the structure and kinetic stability of the streptavidin-biotin (B) interaction, which is among the most stable protein-ligand complexes known and has served as a model system for understanding the origin of high affinity protein complexes. Streptavidin is a homotetrameric protein complex (S_4) that is isolated from *Streptomyces avidinii*.[120] Each streptavidin subunit consists of 159 residues, which are organized into an 8-stranded β-barrel, one end of which forms the binding site for B.[121]

The BIRD technique also has a number of limitations. Firstly, application of the BIRD method requires access to a FTICR mass spectrometer equipped with a temperature-controlled ion cell. This is because BIRD experiments need to be performed at a very low pressure (below 10^{-8} mbar) at which infrared photon absorption and emission exchange between ions and surrounding is the dominant phenomenon compared with negligible collisional energy exchange. Additionally, the determination of Arrhenius activation parameters requires kinetic measurements performed over a range of temperatures, which can be time consuming, especially at low reaction temperatures where the kinetics are slow.

1.5.1.2 CID

CID is the most common ion activation method used in MS/MS as it can be achieved in many kinds of mass spectrometers. In CID, the selected ions of interest are subjected to energetic collisions with neutral background gases within a gas-filled chamber. Collisions between the precursor ion and a neutral target are accompanied by an increase in internal energy, which induces decomposition.

CID has been widely used for the sequencing and structural analysis of peptides, proteins and protein complexes. Valuable information such as stoichiometry, localizing post-translational modifications, and protein sub-unit architecture can be obtained for the biological complexes. A key feature of CID is that the decomposition is slower than energy randomization. Thus the energy will be distributed among all the internal modes of the ion with an equal probability, leading to preferential decomposition at the weakest sites.[53] This feature has allowed CID to be used as a tool to probe the stability of non-covalent complexes in the gas phase. Recently, Douglas and coworkers described a CID-based approach to measure the collision energies required to dissociate the gas-phase ions of complexes of small-molecule inhibitors bound to the wild type (WT) form and single point binding site mutants (Asn44Ala, Gln87Met, and Gln87Tyr) of the catalytic domain of Cex, an enzyme that hydrolyses xylan and xylo-oligosaccharides. The authors found that the dissociation energies followed the same trend as the thermodynamic stabilities in aqueous solution. It was concluded that individual intermolecular H-bonds (involving residues Asn44 and Gln87) were preserved in the gas phase. [122] More recently, Oldham and coworkers described the results of a CID study on the complexes of a series of mutants of the FK506-binding protein (FKBP) and its ligand FK506. [111] Collisional activation of (FKBP + FK506) ions composed of WT or mutant protein, generated by ESI, demonstrated that removal of native protein–ligand interactions formed between residues Asp37 and Tyr82, and FK506 significantly destabilized the gas-phase complex. The gas-phase results obtained for the complexes composed of single point mutants (Asp37Gly and Tyr82Phe) agreed qualitatively with the trend in affinities, evidence for the preservation of these interactions in the desolvated WT complex. In Chapter 5, we tested the reliability of CID-FGR approach for identifying intermolecular interactions in the gaseous ions.

Despite its widespread use, there are limitations to the use of CID as an ion activation method to study non-covalent protein-ligand interactions. Importantly, the fact that during CID, the ions do not reach thermal equilibrium precludes the direct measure of the thermodynamic parameters for the dissociation process. Another drawback, which is due to the "slow-heating" nature of the CID process, is that non-covalent interactions are broken before any covalent bond cleavage is observed. Thus any information on the location of ligand binding is lost.

1.5.2 IMS-MS

IMS is an established technique for studying shape and conformation in protein and protein complexes. IMS separations performed with an inert background gas are based on the charge, size and shape of an ion and are proportional to the ion surface area i.e. to the rotationally averaged collision cross section (Ω). There are three principal types of ion mobility instrumentation that have been successfully coupled with mass spectrometry: Drift Tube IMS (DTIMS), T-Wave IMS (TWIMS), and Field Asymmetric Waveform IMS (FAIMS). Several reviews of IMS-MS instrumentation are available.[42,46,123] Despite differences in the configuration of IMS-MS instruments, there are common features. A pulse of ions is injected into a chamber filled with a known inert gas at a certain pressure. An electric field is applied across the chamber and the time taken for the ions to travel through is measured. The ions are actively transported through the cell by the electrostatic force; at the same time this force is opposed by collisions of ions with buffer gas. The mobility (K) is dependent on the molecular shape, charge of the species and the buffer gas pressure. DTIMS-MS instruments allow direct determination of Ω by recording the drift velocity of an ion through a linear drift field.[124] K is inversely proportional to Ω according to the following relationship:

$$K_0 = \frac{3ze}{16N}\frac{1}{\Omega}\left(\frac{2\pi}{\mu k_B \mathrm{T}}\right)^{1/2} \tag{1.18}$$

where K_0 is the reduced mobility (the measured mobility K standardized for pressure and temperature to 273.15 K and 760 Torr), z is the ion charge state, e is the elementary charge, N is the gas number density, μ is the reduced mass of the ion neutral pair (μ=Mm/(M+m), where M and m are the mass of buffer gas and ion, respectively), k_B is the Boltzmann constant and T is the gas temperature.

The application of IMS-MS for structural studies of biomacromolecules has greatly increased since the introduction of a commercial quadrupole/TWIMS

TOF mass spectrometer.[77] The detail of TWIMS, which is used in this study, is outlined in section 1.2.2.2. A particular advantage of the TWIMS device over most drift tubes is that through the use of ion accumulation and radial ion confinement, the sensitivity of the mass spectrometer is not compromised when operating in mobility mode.[77,125] However, TWIMS-MS device does not directly provide Ω from drift time data. Instead a number of different calibration approaches have been suggested, which all are based on the use of protein standards, whose Ω has been measured in DTIMS-MS instruments. [126]

Information provided by IMS-MS can be enhanced with computational approaches and vice versa.[46] Experimentally determined Ω can help to discriminate between different predicted models where several structures with similar energy minima are possible. A number of algorithms have been developed using different theoretical assumptions;[127,128] however, three of them are commonly used: hard sphere projection approximation (PA), trajectory method (TM) and exact hard sphere scattering (EHSS). For large systems PA fails because the details of the scattering process are obviously not taken into account. For small systems, EHSS fails because the ion-buffer gas interaction becomes important compared to the geometry of the ion and careful calibration of the atomic radii is essential. TM works well for any size system, but calculations are very expensive.[129] Very recently, Bleildholder et al. have developed the projection superposition approximation (PSA), a reworked version of the PA which accounts for the dramatic underestimation this rapid method provides.[130-132] In the framework of the PSA, molecular collision cross sections are computed as a projection approximation modified to account for collective size and shape effects.

1.6 The present work

This thesis focuses on the development and application of MS-based techniques to study non-covalent protein-ligand complexes in solution and the gas phase.

The application of the direct ESI-MS assay for quantifying the stoichiometry and absolute affinity of protein-metal ion binding *in vitro* is described in Chapter 2. Control ESI-MS experiments performed on solutions containing calcium chloride or calcium acetate and a pair of proteins that do not bind calcium ions in solution revealed that the nonspecific association of metal ions to proteins during ESI is a random process, independent of protein size and structure. These results establish the reliability of the reference protein method for quantitatively correcting ESI mass spectra for the occurrence of nonspecific metal ion binding to proteins during ESI-MS analysis. To demonstrate the utility of the direct ESI-MS assay, when carried out using the reference protein method, the calcium binding stoichiometry of bovine α-lactalbumin (La) and the calcium ion affinity of Lg were established.

Chapter 3 demonstrates the quantification of k_{off} for the model high affinity interaction between B and S_4 at pH 7 and temperatures ranging from 15 to 45 ℃ using ESI-MS. Two different approaches to data analysis were employed, one based on the initial rate of dissociation of the (S_4 + 4B) complex, the other involving non-linear fitting of the time-dependent relative abundances of the (S_4 +

*i*B) species. The two methods were found to yield k_{off} values that are in good agreement, within a factor of two. The Arrhenius parameters for the dissociation of the biotin-streptavidin interaction in solution were established from the k_{off} values determined by ESI-MS and compared to values measured using a radiolabeled biotin assay. Importantly, the dissociation activation energies determined by ESI-MS agree, within 1 kcal mol^{-1}, with the reported value. In addition to providing a quantitative measure of k_{off}, the results of the ESI-MS measurements revealed that the apparent cooperative distribution of (S_4 + *i*B) species observed at short reaction times is of kinetic origin and that sequential binding of B to S_4 occurs in a non-cooperative fashion with the four ligand binding sites being kinetically and thermodynamically equivalent and independent.

Work described in Chapters 4 and 5 focuses on the study of the structure and kinetic stability of gaseous complexes ions. In Chapter 4, the BIRD-FGR approach was employed on the model high-affinity protein-ligand interaction between B and S_4. A comparison of E_a measured for the loss of B from the (S_4 + 4B)$^{13+}$ ions composed of WT streptavidin and five binding site mutants (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala and Tyr43Ala) suggests that at least some of the specific intermolecular interactions are preserved in the gas phase. The results of molecular dynamics simulations performed on WT (S_4 + 4B)$^{12+}$ ions with different charge configurations support this conclusion. The most significant finding here is that the gaseous WT (S_4 + 4B)$^{n+}$ ions at n = 12 – 14, owing to a much larger E_a (by as much as 13 kcal mol^{-1}) for the loss of B, are dramatically more stable kinetically at 25 ℃ than the (S_4 + 4B) complex in aqueous neutral solution. The differences in E_a values measured for the gaseous $(S_4 + 4B)^{n+}$ ions and solvated (S_4 + 4B) complex can be largely accounted for by a late dissociative transition state (*TS*) and the rehydration of B and the protein binding cavity in solution.

In Chapter 5, intermolecular interactions in the gaseous ions of two protein-ligand complexes, $(scFv + L1)^{10+}$ and $(S_4 + 4B)^{13+}$, were investigated using a CID-FGR strategy. Intermolecular interactions were identified based on decreases in the collision energy required to dissociate 50% of the reactant ion (i.e., the Ec_{50} value) upon modification of protein residues or ligand functional groups. These results were compared with interaction maps generated using the BIRD-FGR method. For the $(scFv + L1)^{10+}$ ion, it was found that deoxygenation of L1 (at the Gal C3 and C6 and Man C4 and C6 positions) or mutation of His101 (to Ala) resulted in decreases in the Ec_{50} values. These results suggest that the four hydroxyl groups, as well as His101, participate in intermolecular H-bonds. These findings are consistent with those obtained using BIRD-FGR method. However, the CID-FGR method failed to reveal the relative strengths of the intermolecular interactions or establish Man C4 OH and His101 as a H-bond donor/acceptor pair. The CID-FGR method correctly identified Tyr43 as a stabilizing contact in the $(S_4 + 4B)^{+13}$ ion, but failed to identify Trp79, Ser27 and Trp120 as being involved in stabilizing interactions. In fact, mutation of Trp79 and Trp120, led to an increase in the Ec_{50} value. Taken together, the results of this study suggest that the CID-FGR method, as implemented here, does not represent a reliable approach for identifying intermolecular interactions in the gaseous ions of protein-ligand complexes.

1.7 References

1. Cleveland, D. W.; Mao, Y.; Sullivan, K. F. *Cell* **2003**, *112*, 407-421.

2. Igakura, T.; Stinchcombe, J. C.; Goon, P. K. C.; Taylor, G. P.; Weber, J. N.; Griffiths, G. M.; Tanaka, Y.; Osame, M.; Bangham, C. R. M. *Science* **2003**, *299*, 1713-1716.

3. Waldron, K. J.; Robinson, N. J. *Nat. Rev. Micro.* **2009**, *7*, 25-35.

4. Berg, J. M. *Annual Review of Biophysics and Biophysical Chemistry* **1990**, *19*, 405-421.

5. Anderson, B. F.; Baker, H. M.; Dodson, E. J.; Norris, G. E.; Rumball, S. V.; Waters, J. M.; Baker, E. N. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1769-1773.

6. Saboury, A. A. *JICS* **2006**, *3*, 1-21.

7. Wilcox, D. E. *Inorg. Chim. Acta* **2008**, *361*, 857-867.

8. Velázquez Campoy, A.; Freire, E. *Biophys. Chem.* **2005**, *115*, 115-124.

9. Utsuno, K.; Uluda, H. *Biophys. J.* **2010**, *99*, 201-207.

10. Daghestani, H. N., Day, B.W. *Sensors* **2010**, *10*, 9630-9646.

11. De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. *Cell. Mol. Bioeng.* **2008**, *1*, 204-215.

12. Karlsson, R. *Journal of Molecular Recognition* **2004**, *17*, 151-161.

13. Schunk, P. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 541-566.

14. Homola, J.; Piliarik, M. In *Surface Plasmon Resonance Based Sensors*; Homola, J., Ed.; Springer Berlin Heidelberg: 2006; Vol. 4, p 45-67.

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

16. Fangel, J.; Pedersen, H.; Vidal-Melgosa, S.; Ahl, L.; Salmean, A.; Egelund, J.; Rydahl, M.; Clausen, M.; Willats, W. T. In *High-Throughput Phenotyping in Plants*; Normanly, J., Ed.; Humana Press: 2012; Vol. 918, p 351-362.

17. Bizzarri, A. R.; Cannistraro, S. *J. Phy. Chem. B* **2009**, *113*, 16449-16464.

18. Hill, J. J.; Royer, C. A. *Methods Enzymol.* **1997**, *278*, 390-416.

19. Málnási-Csizmadia, A.; Pearson, D. S.; Kovács, M.; Woolley, R. J.; Geeves, M. A.; Bagshaw, C. R. *Biochemistry* **2001**, *40*, 12727-12737.

20. Srisa-Art, M.; Dyson, E. C.; deMello, A. J.; Edel, J. B. *Anal. Chem.* **2008**, *80*, 7063-7067.

21. Sugawara, T.; Kuwajima, K.; Sugai, S. *Biochemistry* **1991**, *30,* 2698-2706.

22. Bao, J.; Krylova, S. M.; Wilson, D. J.; Reinstein, O.; Johnson, P. E.; Krylov, S. N. *ChemBioChem* **2011**, *12*, 2551-2554.

23. Greco, A.; Ho, J. G. S.; Lin, S.-J.; Palcic, M. M.; Rupnik, M.; Ng, K. K.-S. *Nat. Struct. Mol. Biol.* **2006**, *13*.

24. Ho, J. G. S.; Greco, A.; Rupnik, M.; Ng, K. K.-S. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18373-18378.

25. Staunton, D.; Owen, J.; Campbell, I. D. *Acc. Chem. Res.* **2002**, *36*, 207-214.

26. Gilbert, M.; Albala, J. S. *Curr. Opin. Chem. Biol.* **2002**, *6*, 102-105.

27. Zuiderweg, E. R. P. *Biochemistry* **2001**, *41*, 1-7.

28. Wishart, D. *Curr. Pharm. Biotechnol.* **2005**, *6*, 105-120.

29. Gizachew, D.; Dratz, E. *Chem. Biol. Drug Des.* **2011**, *78*, 14-24.

30. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 9340-9341.

31. Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2005**, *77*, 3060-3071.

32. Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2006**, *79*, 416-425.

33. Liu, L.; Bagal, D.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2009**, *131*, 15980-15981.

34. El-Hawiet, A.; Kitova, E. N.; Kitov, P. I.; Eugenio, L.; Ng, K. K.; Mulvey, G. L.; Dingle, T. C.; Szpacenko, A.; Armstrong, G. D.; Klassen, J. S. *Glycobiology* **2011**, *21*, 1217-1227.

35. Kitova, E.; El-Hawiet, A.; Schnier, P.; Klassen, J. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431-441.

36. Konermann, L.; Collings, B. A.; Douglas, D. J. *Biochemistry* **1997**, *36*, 5554-5559.

37. Lee, V. W. S.; Chen, Y.-L.; Konermann, L. *Anal. Chem.* **1999**, *71*, 4154-4159.

38. Sobott, F.; Benesch, J. L. P.; Vierling, E.; Robinson, C. V. *J. Biol. Chem.* **2002**, *277*, 38921-38929.

39. Simmons, D. A.; Wilson, D. J.; Lajoie, G. A.; Doherty-Kirby, A.; Konermann, L. *Biochemistry* **2004**, *43*, 14792-14801.

40. Deng, G.; Sanyal, G. *J. Pharm. Biomed. Anal.* **2006**, *40*, 528-538.

41. Pan, J.; Rintala-Dempsey, A. C.; Li, Y.; Shaw, G. S.; Konermann, L. *Biochemistry* **2006**, *45*, 3005-3013.

42. Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem* **2007**, *76*, 167-193.

43. Clarke, D. J.; Stokes, A. A.; Langridge-Smith, P.; Mackay, C. L. *Anal. Chem.* **2010**, *82*, 1897-1904.

44. Robbins, M. D.; Yoon, O. K.; Barbula, G. K.; Zare, R. N. *Anal. Chem.* **2010**, *82*, 8650-8657.

45. Miao, Z.; Chen, H.; Liu, P.; Liu, Y. *Anal. Chem.* **2011**, *83*, 3994-3997.

46. Pacholarz, K. J.; Garlish, R. A.; Taylor, R. J.; Barran, P. E. *Chem. Soc. Rev.* **2012**, *41*, 4335-4355.

47. Goodlett, D. R.; Ogorzalek Loo, R. R.; Loo, J. A.; Wahl, J. H.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 614-622.

48. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *Angew. Chem. Int. Ed.* **2004**, *43*, 4183-4186.

49. Shoemaker, G. K.; Kitova, E. N.; Palcic, M. M.; Klassen, J. S. *J. Am. Chem. Soc.* **2007**, *129*, 8674-8675.

50. Kitova, E. N.; Seo, M.; Roy, P.-N.; Klassen, J. S. *J. Am. Chem. Soc.* **2008**, *130*, 1214-1226.

51. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2010**, *132*, 17658-17660.

52. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 3054-3060.

53. Wyttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2007**, *58*, 511-533.

54. Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432-14433.

55. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V. *J. Am. Chem. Soc.* **2012**, *134*, 3429-3438.

56. Rose, R. J.; Labrijn, A. F.; van den Bremer, E. T. J.; Loverix, S.; Lasters, I.; van Berkel, P. H. C.; van de Winkel, J. G. J.; Schuurman, J.; Parren, P. W. H. I.; Heck, A. J. R. *Structure* **2011**, *19*, 1274-1282.

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

58. Drake, R. R.; Schwegler, E. E.; Malik, G.; Diaz, J.; Block, T.; Mehta, A.; Semmes, O. J. *Molecular & Cellular Proteomics* **2006**, *5*, 1957-1967.

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

60. Deroo, S.; Hyung, S.-J.; Marcoux, J.; Gordiyenko, Y.; Koripella, R. K.; Sanyal, S.; Robinson, C. V. *ACS Chem. Biol.* **2012**, *7*, 1120-1127.

61. Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

62. Gabelica, V. r.; Rosu, F. d. r.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708-6715.

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

64. Taylor, G. I.; McEwan, A. D. *J. Fluid Mech.* **1965**, *22*, 1-15.

65. Iribarne, J. V.; Thomson, B. A. *The Journal of Chemical Physics* **1976**, *64*, 2287-2294.

66. Thomson, B. A.; Iribarne, J. V. *The Journal of Chemical Physics* **1979**, *71*, 4451-4463.

67. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *The Journal of Chemical Physics* **1968**, *49*, 2240-2249.

68. Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 6926-6927.

69. Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.

70. Karas, M.; Bahr, U.; Dülcks, T. *Fresenius J Anal Chem* **2000**, *366*, 669-676.

71. Juraschek, R.; Dulcks, T.; Karas, M. *J. Am. Soc. Mass. Spectrom.* **1999**, *10*, 300-308.

72. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.

73. Marshall, A. G.; Hendrickson, C. L. *Int. J. Mass spectrom.* **2002**, *215*, 59-75.

74. Xian, F.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2012**, *84*, 708-719.

75. De Hoffmann, E.; Stroobant, V. *Mass Spectrometry Principles and Applications*; 3th ed.; John Wiley & Sons: New York, 2007.

76. Pedder, R. E. *Ardara Technologies Technical Note* **2009**, TN_3005A.

77. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. *Int. J. Mass spectrom.* **2007**, *261*, 1-12.

78. Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrometry Reviews*, **2000**, *19*, 65-107.

79. Wang, W.; Kitova, E. N.; Klassen, J. S. In *Methods Enzymol.*; Yuan, C. L., Reiko, T. L., Eds.; Academic Press: 2003; Vol. Volume 362, p 376-397.

80. Gaylord, N. G.; Gibbs, J. H. *J Polym. Sci.* **1962**, *62*, S22-S23.

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

82. Peschke, M.; Verkerk, U. H.; Kebarle, P. *J. Am. Soc. Mass. Spectrom.* **2004**, *15*, 1424-1434.

83. Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127-1137.

84. Krishnaswamy, S. R.; Williams, E. R.; Kirsch, J. F. *Protein Sci.* **2006**, *15*, 1465-1475.

85. Chitta, R. K.; Rempel, D. L.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **2005**, *1031-1038*.

86. Wilcox, J. M.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **2008**, *80*, 2365-2371.

87. Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413-656.

88. Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801-7806.

89. El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 1893-1899.

90. Liu, L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 310-318.

91. Hossain, B. M.; Konermann, L. *Anal. Chem.* **2006**, *78*, 1613-1619.

92. Sun, J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem.* **2006**, *78*, 3010-3018.

93. Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S. *Anal. Chem.* **2007**, *79*, 8301-8311.

94. Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2010**, *21*, 472-481.

95. Kitova, E. N.; Soya, N.; Klassen, J. S. *Anal. Chem.* **2011**, *83*, 5160-5167.

96. Sun, N.; Sun, J.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass. Spectrom.* **2009**, *20*, 1242-1250.

97. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. *Nat. Rev. Drug Discov.* **2006**, *5*, 730-739.

98. Swinney, D. C. *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 31-39.

99. Callender, R.; Dyer, R. B. *Chem. Rev.* **2006**, *106*, 3031-3042.

100. Chu, Y. H.; Cheng, C. C. *Cell. Mol. Life Sci.* **1998**, *54*, 663-683.

101. Hall, D. R.; Winzor, D. J. *J. Chromatogr. B* **1998**, *715*, 163-181.

102. Chilkoti, A.; Stayton, P. S. *J. Am. Chem. Soc.* **1995**, *117*, 10622-10628.

103. Klumb, L. A.; Chu, V.; Stayton, P. S. *Biochemistry* **1998**, *37*, 7657-7663.

104. Soya, N.; Fang, Y.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2011**, *21*, 547-552.

105. Peleg-Shulman, T.; Najajreh, Y.; Gibson, D. *J. Inorg. Biochem.* **2002**, *91*, 306-311.

106. Pan, J.; Konermann, L. *Biochemistry* **2010**, *49*, 3477-3486.

107. Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.

108. Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859-866.

109. Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.

110. McLuckey, S. A. *J. Am. Soc. Mass. Spectrom.* **1992**, *3*, 599-614.

111. Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851-1858.

112. Philip, R.; Gary, G. In *Practical Aspects of Trapped Ion Mass Spectrometry, Volume IV*; CRC Press: 2010, p 739-767.

113. Polfer, N. C. *Chem. Soc. Rev.* **2011**, *40*, 2211-2221.

114. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.

115. Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.

116. Shi, L.; Liu, S. Y.; Zubarev, R. A. *Progress in Chemistry* **2011**, *23*, 1710-1718.

117. Zubarev, R. A.; Good, D. M.; Savitski, M. M. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 1015-1018.

118. Felitsyn, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2001**, *73*, 4647-4661.

119. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 5902-5913.

120. L. Chaiet, F. J. W. *Arch. Biochem. Biophys.* **1964**, *106*, 1-5.

121. Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2190-2194.

122. Tešić, M.; Wicki, J.; Poon, D. K. Y.; Withers, S. G.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 64-73.

123. Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill, H. H. *J. Mass Spectrom.* **2008**, *43*, 1-22.

124. McCullough, B. J.; Kalapothakis, J.; Eastwood, H.; Kemper, P.; MacMillan, D.; Taylor, K.; Dorin, J.; Barran, P. E. *Anal. Chem.* **2008**, *80*, 6336-6344.

125. Wildgoose, J. L. et. al. In *Abstr. Am. Soc. Mass Spectrom. Meet.* 2006.

126. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protocols* **2008**, *3*, 1139-1152.

127. Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577-592.

128. Wyttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458-3464.

129. Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39*, 1633-1655.

130. Bleiholder, C.; Wyttenbach, T.; Bowers, M. T. *Int. J. Mass spectrom.* **2011**, *308*, 1-10.

131. Anderson, S. E.; Bleiholder, C.; Brocker, E. R.; Stang, P. J.; Bowers, M. T. *Int. J. Mass Spectrom.* **2012**, *330–332*, 78-84.

132. Bleiholder, C.; Contreras, S.; Do, T. D.; Bowers, M. T. *Int. J. Mass Spectrom.* **2012**, *In press*.

Chapter 2

Quantification of Protein-Metal Ion Affinities by Electrospray Ionization Mass Spectrometry[*]

2.1 Introduction

Protein-metal ion interactions play essential roles in many important biological processes. For example, it is estimated that one third of enzymes require a metal ion cofactor for optimal catalytic activity.[1] Metal ion binding is known to enhance the conformational stability of some proteins,[2] while inducing conformational changes in others.[3] Protein-metal ion interactions are also central to the transport of metal ions within cells and across cell membranes.[4] In order to fully understand the mechanisms by which metalloproteins carry out their diverse biological functions, detailed insights into the structures of the protein-metal ion complexes and the corresponding kinetic and thermodynamic parameters of the interactions are required.

There are a number of established analytical techniques used to quantify protein- metal ion interactions *in vitro*. These include absorption spectroscopy,[5] capillary electrophoresis,[6] circular dichroism,[7] and isothermal titration

---

[*] A version of this chapter has been published: Deng, L.; Sun, N.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2010**, *82*, 2170-2174

calorimetry.[8] Recently, the direct ESI-MS assay has emerged as a promising tool for quantifying noncovalent protein interactions in solution.[9-14] The ESI-MS assay is based on the detection of free and ligand-bound protein ions by ESI-MS leading to the determination of K_a values from the abundance ratio of the bound and unbound protein ions measured for a solution of known initial concentrations of protein and ligand. The advantages of the ESI-MS assay include simplicity (no labeling or immobilization), speed (measurements normally can be completed in less than ~1 min), direct insight into binding stoichiometry and the ability to measure multiple binding equilibria simultaneously. Additionally, when performed using nanoESI, which operates at solution flow rates in the nL/min range, the assay typically consumes picomoles or less of analyte per analysis.

A number of studies describing the application of ESI-MS to evaluate protein-metal ion binding have recently been reported.[3,15-19] To our knowledge, however, absolute protein-metal ion affinities have not yet been successfully determined using the direct ESI-MS assay. In fact, it has been suggested that the ESI-MS assay can only be used to evaluate relative, but not absolute, protein-metal ion affinities.[3,18] One of the most significant challenges to the reliable determination of protein-metal ion affinities using the direct ESI-MS assay is the formation of nonspecific interactions between free metal ions and proteins during the ESI process. The appearance of nonspecific complexes (false positives) obscures the true binding stoichiometry in solution and complicates the determination of the metal ion affinity. The formation of nonspecific protein-metal complexes can be minimized by limiting the initial concentration of the metal. However, for weak protein-metal interactions (K_a <10^4 M^{-1}), relatively high metal salt concentrations are necessary to produce detectable levels of complex. In such cases, nonspecific metal ion binding to proteins during the ESI process can be extensive. Fabris and coworkers developed an elegant method, based on gas phase metal-ion transfer reactions, to control the incidence of metal ion adducts in ESI-MS analysis of nucleic acids.[20] More recently, Konermann and coworkers developed an effective method, based on the use of a solution chelator, to reduce the extent of nonspecific metal ion-protein binding in ESI-MS.[21] Notably, it was shown that the tartrate anion can act as an effective chelator and undergo ion pairing with free metal ions in the ESI droplets, thereby reducing the amount of free metal in the droplets that can participate in nonspecific binding with proteins.

A possible alternative approach to the problem of nonspecific protein-metal ion binding involves correcting the ESI mass spectra for the contribution of nonspecific metal ion binding using the *reference protein method*. This method, which was originally developed to correct mass spectra for nonspecific protein-ligand binding during ESI-MS analysis, involves the addition of a P_{ref}, which does not bind specifically to the protein or ligand of interest, to the solution.[22] The fraction of P_{ref} undergoing nonspecific ligand binding provides a quantitative measure of the contribution of nonspecific ligand binding to the abundance of the protein and specific protein-ligand complex(es). The underlying assumption with the method is that nonspecific ligand binding is a random process and, in a given ESI-MS experiment, is independent of the size and structure of the proteins present. Notably, the method has been shown to correctly account for the nonspecific binding of neutral, acidic and basic biomolecules to proteins during ESI-MS analysis.[23-27] However, the reliability of the P_{ref} method for quantifying nonspecific proteins-metal ion interactions has not yet been tested.

Here, we investigate the suitability of the direct ESI-MS assay for quantifying protein-metal ion binding *in vitro*. Importantly, control experiments performed on solutions containing calcium salts and noninteracting proteins confirmed that the P_{ref} method can be used to correct ESI mass spectra for the occurrence of nonspecific metal ion binding to proteins. To demonstrate the reliability of the ESI-MS assay, when carried out in conjunction with the reference protein method, calcium ion binding measurements were performed on La and Lg.

2.2 Materials and Methods

2.2.1 Sample preparation

Calcium chloride ($CaCl_2$), calcium acetate ($Ca(OAc)_2$), bovine ubiquitin, Ubq (MW 8 565 Da), chicken egg white lysozyme, Lyz (MW 14 315 Da), bovine α- lactalbumin, La (MW 14 178 Da) and bovine β-lactoglobulin, Lg (monomer MW 18 276 Da) were purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). The carbohydrate-binding antibody single chain fragment, scFv (MW 26 539 Da), was produced using recombinant technology as described elsewhere.[30] All proteins were concentrated and dialyzed against deionized water using microconcentrators (Millipore Corp., Bedford, MA) with a molecular weight cutoff (MWCO) of 5 kDa (Ubq) or 10 kDa (Lyz, scFv and La). Purification of Lg was carried out using first a 10 kDa MWCO microconcentrator, followed by 100 kDa MWCO microconcentrator, to remove Lg aggregates. The purified protein solutions were lyophilized and weighed immediately after removal from the lyophilizer. Protein stock solutions were prepared by dissolving a known mass of protein in deionized water and stored at -20 ^{o}C until used. Calcium salt stock solutions (100 mM) were prepared by dissolving a known mass of $CaCl_2$ or $Ca(OAc)_2$ in deionized water. The ESI solutions were prepared from the stock solutions and aqueous ammonium acetate (pH 6.8) at a final concentration of 5 mM. In the case of Lg, the ESI solutions were made up with aqueous ammonium acetate (pH 8.2) at a final concentration of 10 mM.

2.2.2 Mass spectrometry

All experiments were performed on a 9.4 tesla Apex II FT-ICR mass spectrometer (Bruker, Billerica, MA) equipped with a nanoflow ESI ion source. NanoESI was performed using borosilicate tubes (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-2000 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in positive ion mode was established by applying a voltage of ~1000 V to a platinum wire inserted inside the glass tip. The solution flow rate was typically ~20 nL/min. The gaseous ions produced by nanoESI were introduced into the mass spectrometer through a stainless steel capillary (i.d. 0.43 mm) maintained at an external temperature of 66 ℃. The gas flow rate into the instrument was measured to be 0.6 L min^{-1}. The ions sampled by the capillary (~50 V) were transmitted through a skimmer (0 V) and accumulated electrodynamically in an rf hexapole. To minimize the extent of ion source dissociation, a hexapole accumulation time of 1 s was used. Ions were then ejected from the hexapole and accelerated (+2700 V) into the superconducting magnet, decelerated, and introduced into the ion cell. The trapping plates of the cell were maintained at a constant potential of ~1.4 V throughout the experiment. The typical base pressure for the instrument was ~$5\text{x}10^{-10}$ mbar. Data acquisition was controlled by an SGI R5000 computer running the Bruker Daltonics XMASS software, version 5.0. Mass spectra were obtained using standard experimental sequences with chirp broadband excitation. The time domain signal, consisting of the sum of 30 transients containing 128 K data points per transient, were subjected to one zero-fill prior to Fourier transform.[11,28]

2.2.3 Determination protein-metal ion binding stoichiometry and affinity

The procedures used for the determination of protein-ligand K_a directly from ESI mass spectra using the reference protein method are described in detail in Chapter 1.

The fractional abundance ($f_{i,P}$) of the protein-calcium ion complexes, i.e., (P + *i*Ca), identified by ESI-MS were established from the measured abundance (*Ab*) of the corresponding gas phase $(\text{P} + i\text{Ca})^{n+}$ ions using eq 2.1:

$$f_{i,P} = \frac{\sum_{n} Ab(P + i\text{C}a)^{n+})}{\sum_{i}\sum_{n} Ab(P + iCa)^{n+})} \tag{2.1}$$

The presence or absence of specific protein-calcium ion interactions in solution was established from a comparison of the distributions of the (P + *i*Ca) and (P_{ref} + *i*Ca) species measured in the same ESI mass spectrum. In the absence of specific calcium ion binding, the ratio of the corresponding $f_{i,P}$ terms (at all i values) will be identical, within experimental error. If, on the other hand, specific protein-calcium ion binding contributes to the ESI mass spectrum, the corresponding $f_{i,P}$ ratios will deviate from unity. Where specific protein-calcium ion binding was identified, the "true" abundance of the (P + *i*Ca) species present in solution was established using the P_{ref} correction method. Complete details of how the correction is carried out can be found elsewhere.[22] Shown below are the relevant expressions needed to establish the true fractional abundance, $(f_{i,P})^{true}$, of the specific (P + *i*Ca) species, where $i = 0 - 4$, from the $f_{i,P}$ and $f_{i,Pref}$ terms obtained directly from the ESI mass spectrum:

$$(f_{0,P})^{true} = f_{0,P}/f_{0,Pref} \tag{2.2a}$$

$$(f_{1,P})^{true} = [f_{1,P} - (f_{0,P})^{true}f_{1,Pref}]/f_{0,Pref} \tag{2.2b}$$

$$(f_{2,P})^{true} = [f_{2,P} - (f_{0,P})^{true}f_{2,Pref} - (f_{1,P})^{true}f_{1,Pref}]/f_{0,Pref} \tag{2.2c}$$

$$(f_{3,P})^{true} = [f_{3,P} - (f_{0,P})^{true}f_{3,Pref} - (f_{1,P})^{true}f_{2,Pref} - (f_{2,P})^{true}f_{1,Pref}]/f_{0,Pref} \tag{2.2d}$$

$$(f_{4,P})^{true} = [f_{4,P} - (f_{0,P})^{true}f_{4,Pref} - (f_{1,P})^{true}f_{3,Pref} - (f_{2,P})^{true}f_{2,Pref} - (f_{3,P})^{true}f_{1,Pref}]/f_{0,Pre} \tag{2.2e}$$

The apparent calcium ion affinity ($K_{a,app}$) for Lg was determined from the corrected (for nonspecific binding) fractional abundance of free Lg and the 1:1 (Lg + Ca) complex using eqs 2.3 and 2.4:

$$[Lg + Ca]/[Lg] = (f_{1,P})^{true}/(f_{0,P})^{true} = R \tag{2.3}$$

$$\mathrm{K}_{a,app} = \frac{R}{[Ca]_o - \dfrac{R}{1+R}[Lg]_o} \tag{2.4}$$

where $[\mathrm{Lg}]_\mathrm{o}$ and $[\mathrm{Ca}]_\mathrm{o}$ are the initial solution concentrations of Lg and the Ca^{2+}, respectively. Because Lg is known to possess three equivalent calcium ion binding sites,[29] the intrinsic calcium ion affinity (K_a) was obtained by dividing $\mathrm{K}_{\mathrm{a},app}$ by the number of binding sites, i.e., $\mathrm{K}_\mathrm{a} = (\mathrm{K}_{\mathrm{a},app})/3$.

2.3 Results and Discussion

2.3.1 Nonspecific Protein-Metal Ion binding during the ESI process

To test the suitability of the P_ref method for quantifying the occurrence of nonspecific protein-metal ion binding during the ESI process, a series of control ESI-MS experiments were performed on solutions containing a pair of proteins and calcium acetate or calcium chloride salt. Three proteins, Ubq, Lyz and scFv, which do not possess an appreciable affinity for Ca^{2+}, served as model proteins for the present study. Shown in Figure 2.1 are representative ESI mass spectra acquired in positive ion mode for solutions of Ubq (6 μM), Lyz (8 μM) and CaCl_2 at 25, 50 and 100 μM. In each of the mass spectra, ions corresponding to the free, protonated proteins were detected, i.e., $(\mathrm{Ubq} + n\mathrm{H})^{n+} \equiv \mathrm{Ubq}^{n+}$ at n = 5, 6 and $(\mathrm{Lyz} + n\mathrm{H})^{n+} \equiv \mathrm{Lyz}^{n+}$ at n = 7 - 9. Also detected were ions corresponding to Ubq and Lyz bound to one or multiple Ca^{2+} ions, i.e., $(\mathrm{Ubq} + i\mathrm{Ca}^{2+} + (n-2i)\mathrm{H})^{n+} \equiv (\mathrm{Ubq} + i\mathrm{Ca})^{n+}$ and $(\mathrm{Lyz} + i\mathrm{Ca}^{2+} + (n-2i)\mathrm{H})^{n+} \equiv (\mathrm{Lyz} + i\mathrm{Ca})^{n+}$, at the same charge states as observed for the free protein ions. As expected, the maximum number of bound Ca^{2+} ions increased with salt concentration.[28] Also shown in Figure 2.1 are the normalized distributions of the (Ubq + *i*Ca) and (Lyz + *i*Ca) species measured by ESI-MS at each of the $CaCl_2$ concentrations investigated. To aid in the comparison of the distributions, the ratios of the fractional abundance ($f_{i,P}$) of the (P + *i*Ca) species, where P = Ubq or Lyz, are listed in Table 2.1. Notably, the distributions measured at 25 and 50 μM salt concentrations are indistinguishable within experimental error. At 100 μM, the distributions are no longer identical. However, the $f_{i,P}$ ratios deviate from unity by less than ±15% and the relative abundance of free protein is identical for Ubq and Lyz, i.e., the $f_{0,Ubq}/f_{0,Lyz}$ ratio is 1.03 ± 0.04. Analogous ESI-MS measurements were performed on solutions containing the Lyz/scFv protein pair. Representative ESI mass spectra and the normalized distributions of the (P + *i*Ca) species are shown in Figure 2.2; the ratios of the corresponding $f_{i,P}$ values determined for this pair of proteins are also listed in Table 2.1. Importantly, at each salt concentration the distributions of Ca^{2+} bound nonspecifically to Lyz and to scFv are indistinguishable, within experimental error.

In their recent study, Konermann and coworkers demonstrated that the nature of the calcium salt can influence the extent of nonspecific calcium ion binding to proteins during ESI-MS.[21] To test whether the choice of calcium salt influences the effectiveness of the reference protein method, the ESI-MS measurements were repeated using $Ca(OAc)_2$. Shown in Figure 2.3 are illustrative ESI mass spectra acquired in positive ion mode for solutions of Ubq, Lyz and $Ca(OAc)_2$ (at 50 and 100 μM) along with the normalized distributions of bound Ca^{2+} ions. The mass spectra are similar in appearance to those obtained for solutions containing $CaCl_2$ with the exception of the appearance of signal corresponding to $(Lyz + Ca + OAc)^{n+}$ and $(Lyz + 2Ca + OAc)^{n+}$ ions in some of the mass spectra. The absence of the corresponding $(Lyz + OAc)^{n+}$ ions is notable and suggests that the bound acetate anion is directly associated with a Ca^{2+} ion. Curiously, CaOAc adducts were not detected for either Ubq or scFv and it is not currently understood why these adducts are only observed for Lyz. Nevertheless, to calculate the distribution of calcium ions bound nonspecifically to Lyz, the CaOAc adduct was treated as being equivalent to a single bound Ca^{2+} ion. Importantly, the distributions measured at 50 μM $Ca(OAc)_2$ are indistinguishable; while at 100 μM $Ca(OAc)_2$, the $f_{i,P}$ ratios deviate from unity by less than 15%, Table 1. Similar results were obtained for solutions containing the Lyz/scFv protein pair, Figure 2.4, Table 2.1.

Taken together, the results of these control experiments confirm that, independent of the nature of the calcium salt, nonspecific association of free Ca^{2+} ions to proteins during the ESI process occurs in a random fashion, independent of the size and structure of the proteins that are present in the ESI droplets. That is not to say that the nature of the protein-metal ion interactions is not protein structure dependent. Rather, these findings indicate that distribution of calcium ions in ESI droplets in the final stages of desolvation is essentially independent of the nature of the proteins that are present. It can, therefore, be concluded that the reference protein method can be used effectively to correct ESI mass spectra for the occurrence of nonspecific calcium ion-protein binding during ESI-MS analysis. Furthermore, it is reasonable to suggest that the reference protein method can generally be used to correct ESI mass spectra for the nonspecific binding of divalent metal ions.

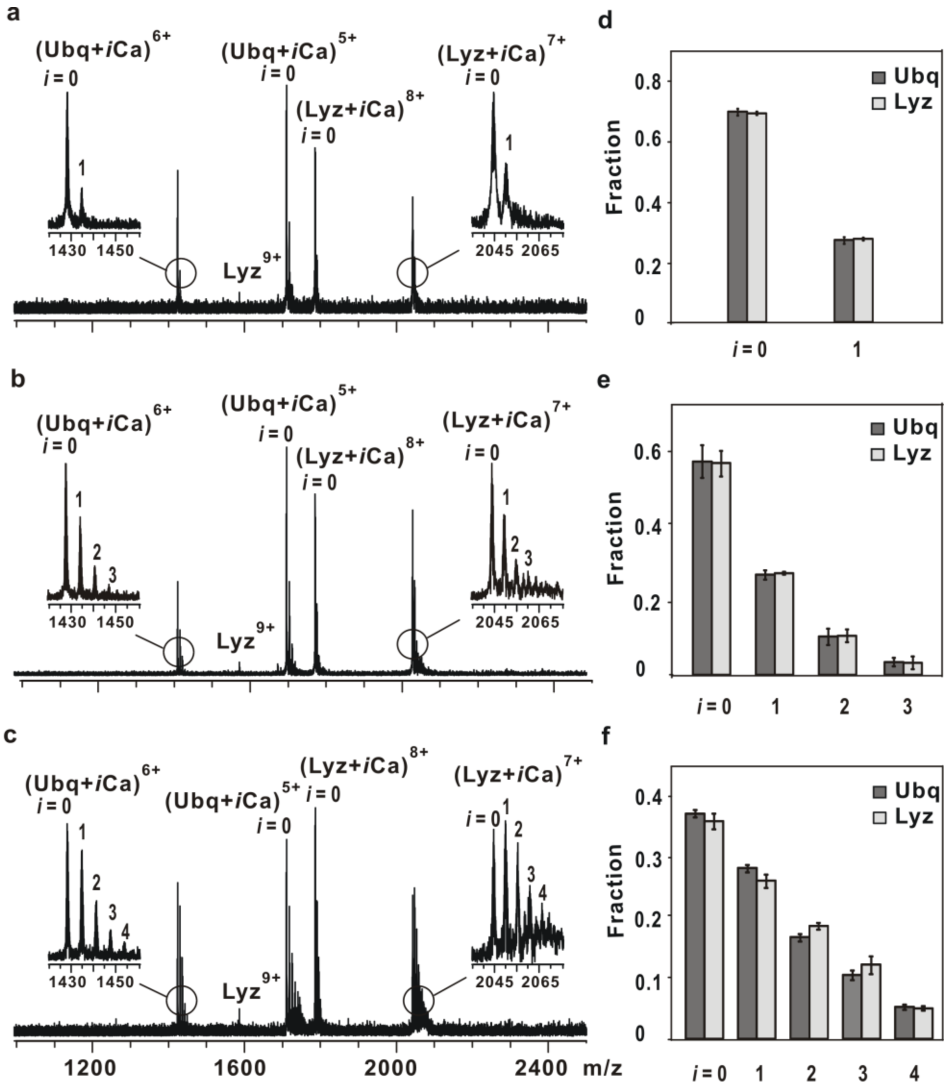


Figure 2.1. ESI mass spectra obtained in positive ion mode for solutions of Ubq (6 μM) and Lyz (8 μM) with $CaCl_2$ at (a) 25 μM, (b) 50 μM and (c) 100 μM. (d) – (f) Normalized distributions of nonspecific (P + *i*Ca) complexes, where P = Ubq and Lyz, determined from ESI mass spectra acquired for the solutions described in (a) - (c), respectively. The reported errors correspond to one standard deviation and were determined from 3 - 5 replicate measurements.

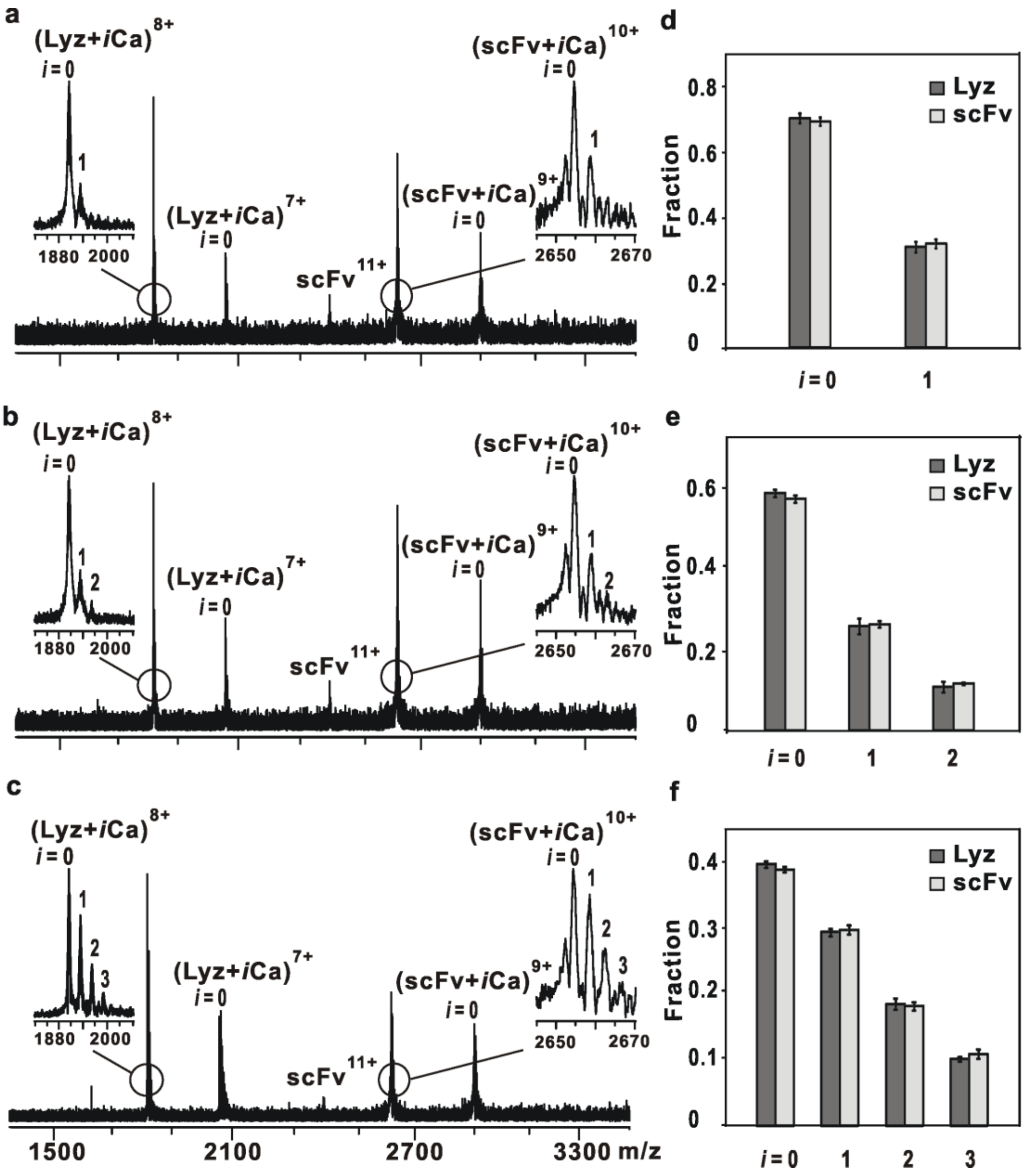


Figure 2.2. ESI mass spectra obtained in positive ion mode for solutions of Lyz (8 μM) and scFv (10 μM) with $CaCl_2$ at (a) 25 μM, (b) 50 μM and (c) 100 μM. (d) - (f) Normalized distributions of nonspecific (P + *i*Ca) complexes, where P = Lyz and scFv, determined from ESI mass spectra acquired for the solutions described in (a) - (c), respectively. The reported errors correspond to one standard deviation and were determined from 3 - 5 replicate measurements.

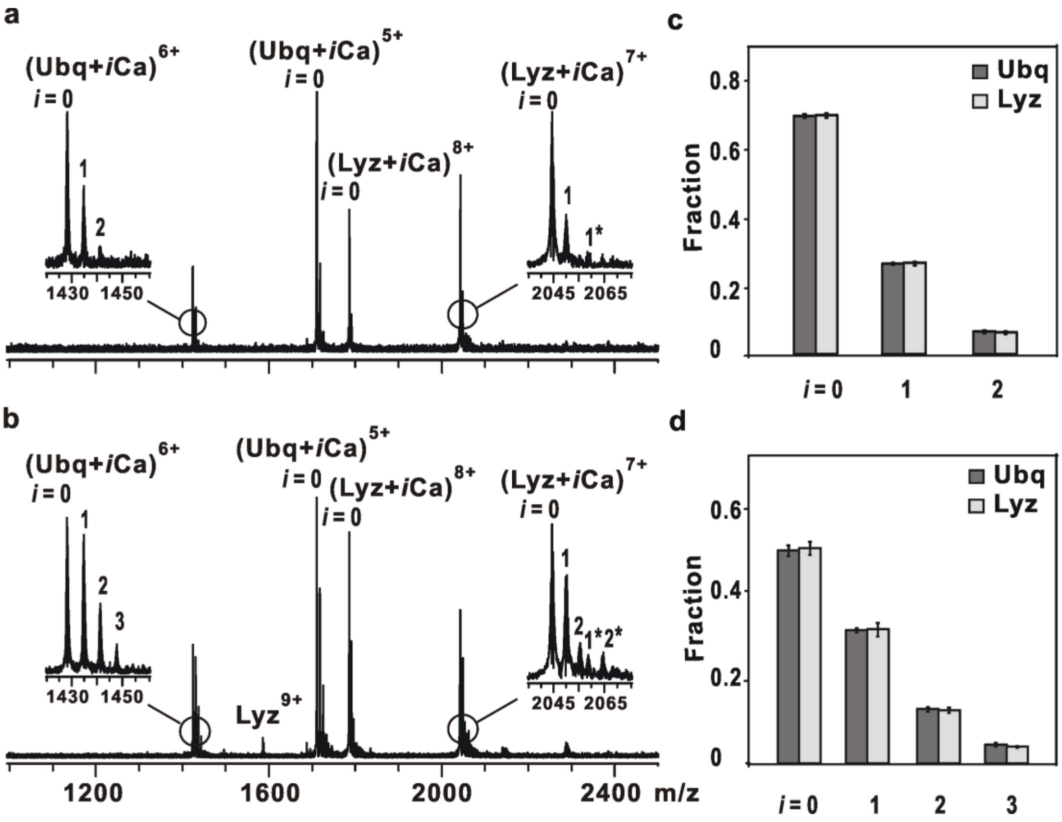


Figure 2.3. ESI mass spectra obtained in positive ion mode for solutions of Ubq (6 μM) and Lyz (8 μM) with $Ca(OAc)_2$ at (a) 50 μM and (b) 100 μM. The peaks labeled as 1* and 2* correspond to the nonspecific attachment of Ca(OAc) and $Ca_2(OAc)$, respectively, to Lyz. (c) and (d) Normalized distributions of nonspecific (P + *i*Ca) complexes, where P = Ubq and Lyz, determined from ESI mass spectra acquired for the solutions described in (a) and (b), respectively. The reported errors correspond to one standard deviation and were determined from 5 replicate measurements.

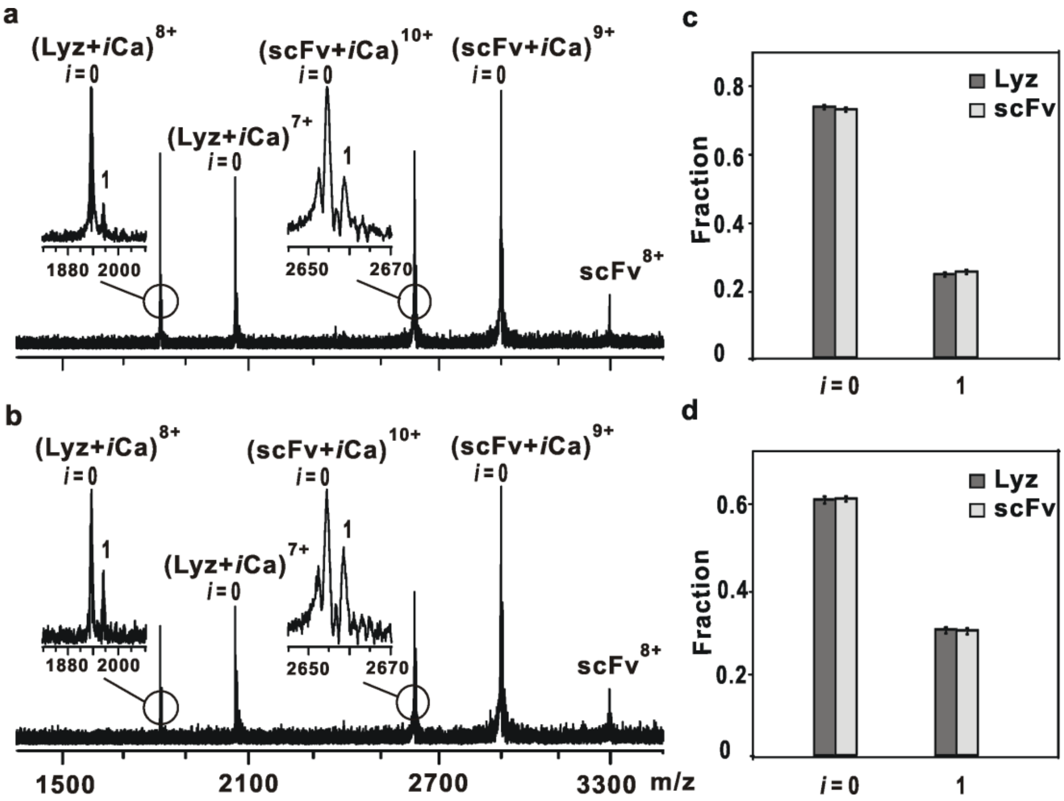


Figure 2.4. ESI mass spectra obtained in positive ion mode for solutions of Lyz (8 μM), and scFv (10 μM) with $Ca(OAc)_2$ at (a) 50 μM and (b) 100 μM. (c) and (d) Normalized distributions of nonspecific (P + *i*Ca) complexes, where P = Lyz and scFv, determined from ESI mass spectra acquired for the solutions described in (a) and (b), respectively. The reported errors correspond to one standard deviation and were determined from 5 replicate measurements.

Table 2.1. Ratio of $f_{i,P}$ terms determined for the nonspecific binding of Ca^{2+} ions to pairs of proteins (Ubq/Lyz and Lyz/scFv) in positive ion mode ESI-MS. [a,b]

$[\mathrm{Ubq}]_o$ (μM)	**$[\mathrm{Lyz}]_o$ (μM)**	**$[\mathrm{CaCl_2}]_o$ (μM)**	**$[\mathrm{Ca(OAc)_2}]_o$ (μM)**	$f_{0,\mathrm{Ubq}}/f_{0,\mathrm{Lyz}}$	$f_{1,\mathrm{Ubq}}/f_{1,\mathrm{Lyz}}$	$f_{2,\mathrm{Ubq}}/f_{2,\mathrm{Lyz}}$	$f_{3,\mathrm{Ubq}}/f_{3,\mathrm{Lyz}}$
6	8	25		1.01 ±0.02	0.98 ±0.05		
6	8	50		1.01 ±0.10	0.98 ±0.05	0.98 ±0.26	1.04 ±0.65
6	8	100		1.03 ±0.04	1.08 ±0.05	0.90 ±0.04	0.86 ±0.08
6	8		50	1.00 ±0.01	1.00 ±0.02	1.04 ±0.09	
6	8		100	0.99 ±0.04	0.99 ±0.05	1.02 ±0.07	1.13 ±0.11
$[\mathrm{Lyz}]_o$ (μM)	**$[\mathrm{scFv}]_o$ (μM)**	**$[\mathrm{CaCl_2}]_o$ (μM)**	**$[\mathrm{Ca(OAc)_2}]_o$ (μM)**	$f_{0,\mathrm{Lyz}}/f_{0,\mathrm{scFv}}$	$f_{1,\mathrm{Lyz}}/f_{1,\mathrm{scFv}}$	$f_{2,\mathrm{Lyz}}/f_{2,\mathrm{scFv}}$	$f_{3,\mathrm{Lyz}}/f_{3,\mathrm{scFv}}$
8	10	25		1.02 ±0.03	0.97 ±0.06		
8	10	50		1.02 ±0.02	0.98 ±0.08	0.92 ±0.12	
8	10	100		1.02 ±0.02	0.99 ±0.03	1.02 ±0.06	0.93 ±0.07
8	10		50	1.01 ±0.01	0.97 ±0.04		
8	10		100	1.00 ±0.02	1.01 ±0.04		

a. All measurements were performed at 25 °C. b. All ESI mass spectra were acquired with 128K data points. Only peaks with a signal-to-noise ratio > 3 were included in the determination of the $f_{i,P}$ terms. Ratios were calculated from average $f_{i,P}$ values taken from 3 - 5 replicate measurements.

2.3.2 Protein-Metal Ion Affinities determined by the direct ESI-MS assay

To demonstrate the reliability of the direct ESI-MS assay, when combined with the reference protein method, it was used to investigate the binding of Ca^{2+} to La and to Lg in solution. Bovine La possesses a single, high affinity calcium binding site, $K_a = 3x10^8$ M^{-1} (pH 7.4, 25˚C and ionic strength of 0.1 M),[6] while bovine Lg possesses three equivalent, low affinity calcium binding sites, each with a K_a of (2.9 ± 0.2) x 10^2 M^{-1} (pH 7.2, 20˚C and ionic strength of 150 mM).[29] Shown in Figure 2.5 are ESI mass spectra acquired in positive ion mode for solutions of La (10 μM), scFv (10 μM) and $CaCl_2$ at three different concentrations (25, 50 and 100 μM). The P_{ref} used for these measurements was scFv. Notably, no free La ions were detected in any of the spectra. Instead, the $(La + Ca)^{n+}$ ions, at n = 6 and 7, dominated the mass spectra; ions corresponding to the attachment of additional Ca^{2+} ions, with the number of bound Ca^{2+} ions increasing with salt concentration, were also identified. Also detected were ions corresponding to scFv bound nonspecifically to one or more Ca^{2+} ions, with the number of bound Ca^{2+} ions increasing with salt concentration. Shown in Figure 2.5 are the normalized distributions of bound calcium for La and scFv. It can be seen that the distributions of Ca^{2+} bound nonspecifically bound to scFv are essentially identical to those measured for the (La + Ca) complex, e.g., the $f_{0,(La+Ca)}/f_{0,scFv}$ ratios are 1.00 ± 0.01 (25 μM), 0.98 ± 0.02 (50 μM) and 1.01 ± 0.03 (100 μM) and the $f_{1,(La+Ca)}/f_{1,scFv}$ ratios are 1.00 ± 0.03 (25 μM), 1.08 ± 0.04 (50 μM) and 0.97 ± 0.03

(100 μM). These results indicate that the (La + Ca) complex is the only major La species present in solution. This finding agrees with the distribution of La species expected under the solution conditions used for the ESI-MS measurements.[6]

Shown in Figures 2.6a, 2.6b and Figure 2.7a are ESI mass spectra acquired in positive ion mode for solutions of Lg (15 μM), scFv (10 μM), and $CaCl_2$ at three different concentrations (100, 150 and 200 μM). In each case, ions corresponding to free Lg, and Lg bound to multiple Ca^{2+} ions were detected and the number of bound Ca^{2+} ions increased with salt concentration. The normalized distributions of calcium ions bound to Lg and scFv, as determined from the ESI mass spectra, for the three different $CaCl_2$ concentrations are shown in Figures 2.6c (100 μM), 2.6d (150 μM) and Figure 2.7b (200 μM). The distributions following correction for nonspecific binding are shown in Figures 2.6e and 2.6f and Figure 2.7c. A detailed description of the correction method can be found elsewhere[22] and the relevant equations used in the present study are given in section 2.2.3. After removing the contribution resulting from nonspecific binding, it is found that free Lg and Lg bound to a single calcium ion are the only major Lg species present in solution.

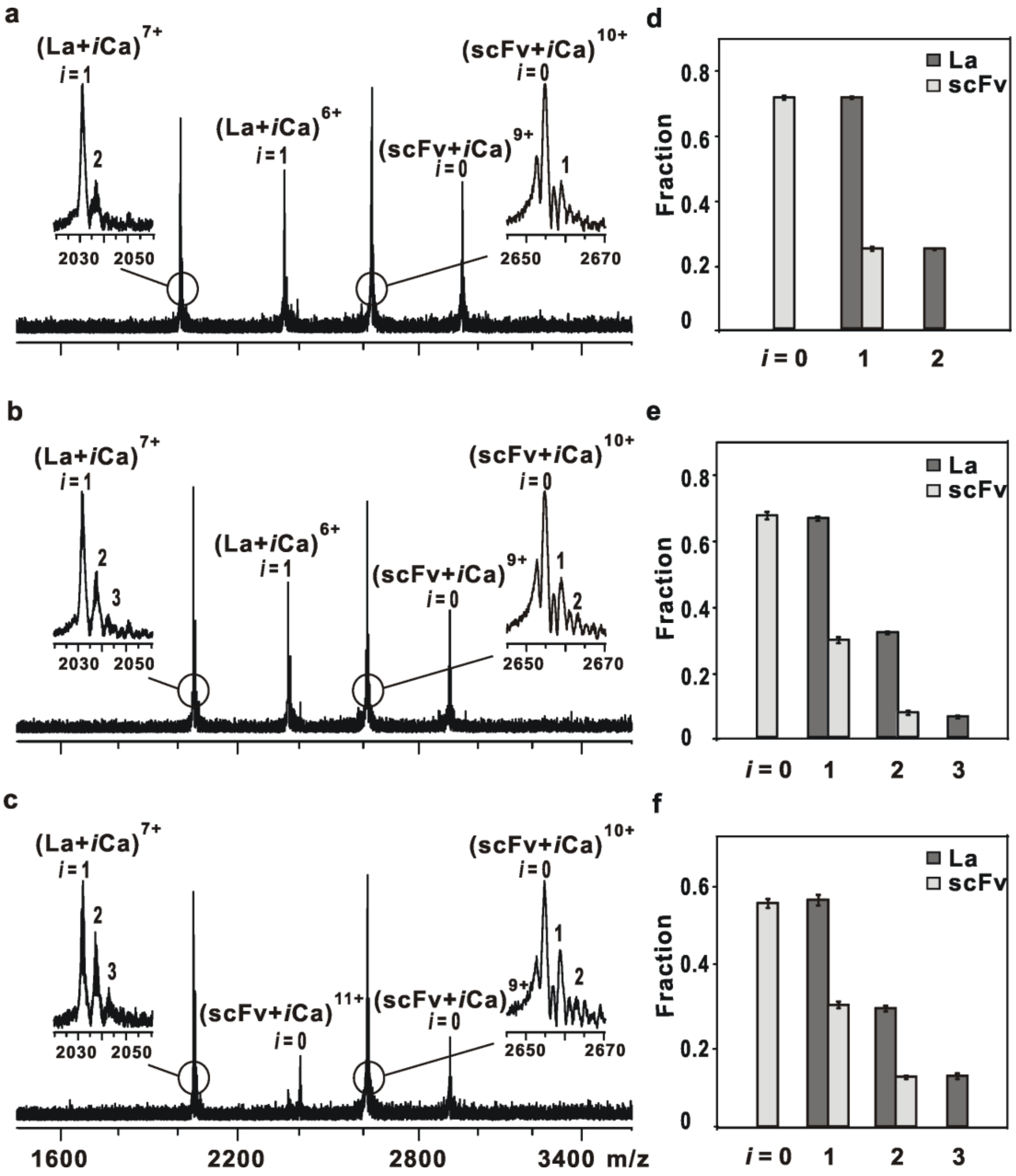


Figure 2.5. ESI mass spectra obtained in positive ion mode for solutions of scFv (10 μM), and La (10 μM) with $CaCl_2$ at (a) 25 μM and (b) 50 μM and (c) 100 μM. (d) - (f) Normalized distributions of calcium bound to La and scFv determined from ESI mass spectra acquired for the solutions described in (a) - (c), respectively. The reported errors correspond to one standard deviation and were determined from 4 or 5 replicate measurements.

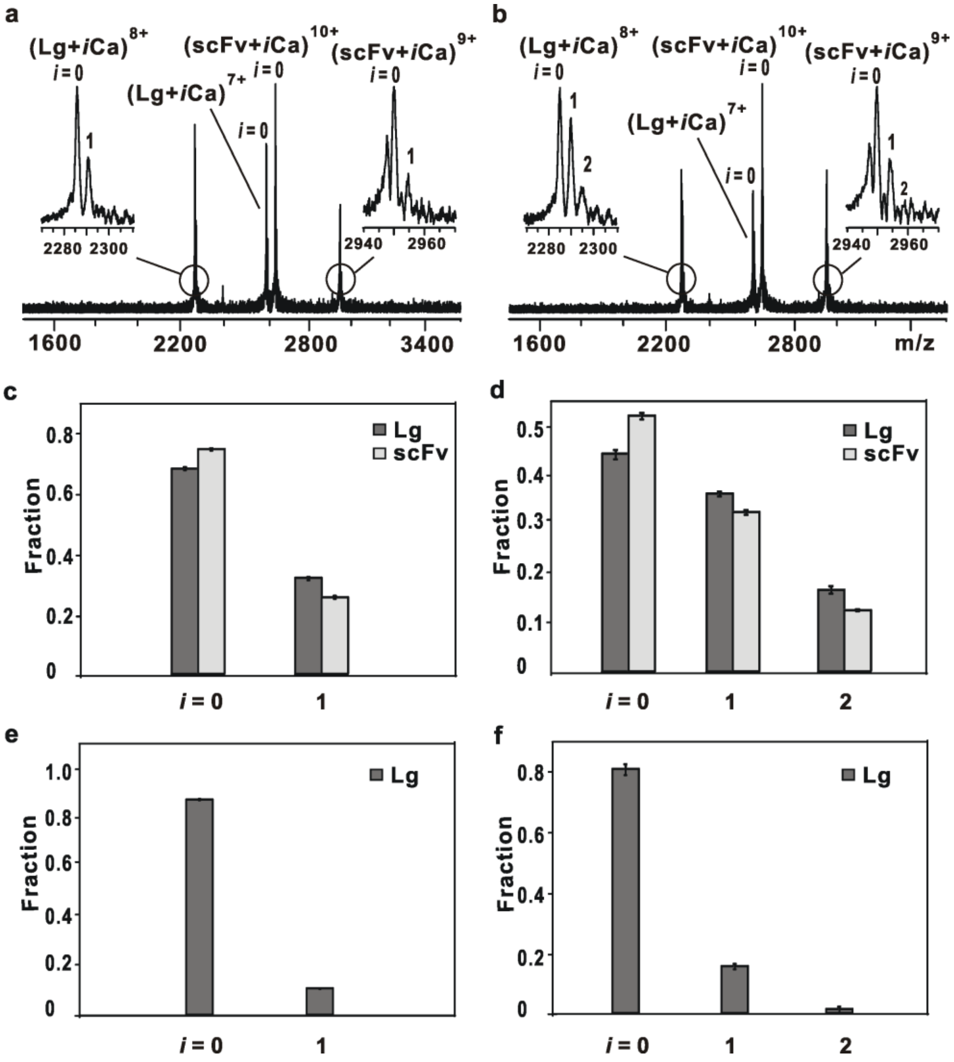


Figure 2.6. ESI mass spectra obtained for solutions of scFv (10 μM) and Lg (15 μM) with $CaCl_2$ at (a) 100 μM and (b) 150 μM. (c) and (d) Normalized distributions of calcium bound to Lg and scFv determined from ESI mass spectra acquired for the solutions described in (a) and (b), respectively. (e) and (f) Normalized distribution of calcium bound to Lg following correction for nonspecific binding using the reference protein method. The reported errors correspond to one standard deviation and were determined from 4 or 5 replicate measurements.

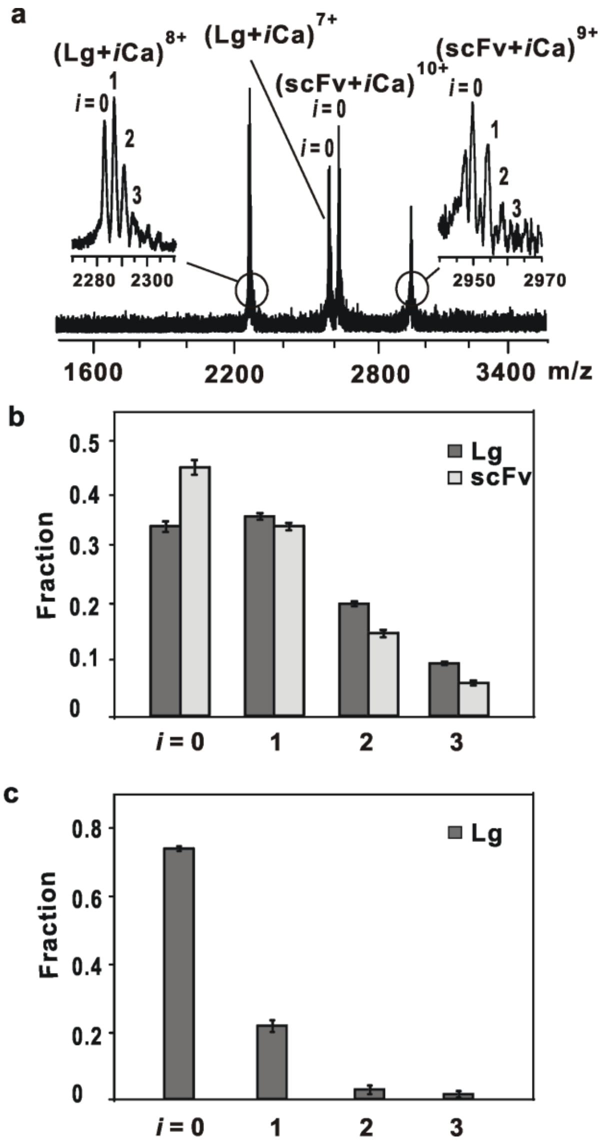


Figure 2.7. ESI mass spectrum obtained for a solution of scFv (10 μM) and Lg (15 μM) with $CaCl_2$ (200 μM). (b) Normalized distributions of calcium bound to Lg and scFv determined from ESI mass spectra acquired for the solution described in (a). (c) Normalized distribution of calcium bound to Lg following correction for nonspecific binding using the reference protein method. The reported errors correspond to one standard deviation and were determined from 5 replicate measurements.

The apparent calcium ion affinity ($K_{a,app}$) for Lg was determined from the corrected (for nonspecific binding) fractional abundance of free Lg and the 1:1 (Lg + Ca) complex using eqs 2.3 and 2.4. The $K_{a,app}$ values, determined at different salt concentrations, are similar: $(1.3 \pm 0.1) \times 10^3$ M^{-1} (100 μM), $(1.4 \pm 0.2) \times 10^3$ M^{-1} (150 μM) and $(1.5 \pm 0.1) \times 10^3$ M^{-1} (200 μM). The average $K_{a,app}$, $(1.4 \pm 0.1) \times 10^3$ M^{-1}, corresponds to an intrinsic K_a, for each of the three calcium binding sites, of $(4.7 \pm 0.3) \times 10^2$ M^{-1}. This value is in reasonable agreement (within a factor of two) with the reported value of $(2.9 \pm 0.2) \times 10^2$ M^{-1}.[29]

2.4 Conclusions

In summary, the results of the present study demonstrate that the direct ESI-MS assay can serve as a reliable method to establish the stoichiometry of calcium and, presumably, other metals binding to proteins in solution and to quantify protein-metal ion affinities. Successful implementation of the assay requires that the occurrence of nonspecific metal ion binding be minimized or corrected for, after the fact. Control experiments revealed that the nonspecific association of divalent metal ions to proteins during ESI-MS is a random process, independent of protein size and structure. These findings confirm that the reference protein method can be successfully used to quantitatively correct ESI mass spectra for the occurrence of nonspecific metal ion binding to proteins. To demonstrate the utility of the direct ESI-MS assay, when implemented with the reference protein method, it was used to establish the calcium binding stoichiometry of bovine α-lactalbumin and to quantify the calcium ion affinity of bovine β-lactoglobulin.

2.5 References

1. Waldron, K. J.; Robinson, N. J. *Nat. Rev. Micro.* **2009**, *7*, 25-35.

2. Berg, J. M. *Annu. Rev. Biophys. Chem*. **1990**, *19*, 405-421.

3. Shirran, S. L.; Barran, P. E. *J. Am. Soc. Mass Spectrom*. **2009**, *20*, 1159-1171.

4. Anderson, B. F.; Baker, H. M.; Dodson, E. J.; Norris, G. E.; Rumball, S. V.; Waters, J. M.; Baker, E.N. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1769-1773.

5. Farrell, H. M.; Thompson, M. P. *Protoplasma* **1990**, *159*, 157-167.

6. Rasmussen, B. W.; Bjerrum, M. J. *J. Inorg. Biochem.* **2003**, *95*, 113-123.

7. Hebenstreit, D.; Ferreira, F. *Allergy* **2005**, *60*, 1208-1211.

8. Saboury, A. A.; Atri, M. S.; Sanati, M. H.; Sadeghi, M. *J. Therm. Anal. Cal.* **2006**, *83*, 175-179.

9. Kitova, E. N.; Kitov, P. I.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2001**, *11*, 605-611.

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

11. Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2003**, *75*, 4945-4955.

12. Heck A. J. R.; Van Den Heuvel R. H. H. *Mass Spectrom. Rev*. **2004**, *23*, 368-389.

13. Loo, J. A.; Berhane, B.; Kaddis, C. S.; Wooding, K. M.; Xie, Y. M.; Kaufman, S. L.; Chernushevich, I. V. *J. Am. Soc. Mass Spectrom*. **2005**, *16*, 998-1008.

14. Jecklin, M. C.; Touboul, D.; Bovet, C.; Wortmann, A.; Zenobi, R. *J. Am. Soc.*

Mass Spectrom. **2008**, *19*, 332–343.

15. Loo, J. A. *Int. J. Mass Spectrom*. **2001**, *204*, 113-123.

16. Zhu, M. M.; Rempel, D. L.; Du, Z.; Gross, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 5253-5253.

17. Wortmann, A.; Rossi, F.; Lelais, G.; Zenobi, R. *J. Mass Spectrom*. **2005**, *40*, 777-784.

18. Vriendt, K. D.; Driessche, G. V.; Devreese, B.; Bebrone, C.; Anne, C.; Frere, J. M.; Galleni, M.; Beeumen, J. V. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 180-188.

19. Kaltoshov, I. A.; Zhang, M.; Eyles, S. J.; Abzalimov, R. R. *Anal. Bioanal. Chem*. **2006**, *386*, 472-481.

20. Turner, K. B.; Monti, S. A.; Fabris, D. *J. Am. Chem. Soc*. **2008**, *130*, 13353-13363.

21. Pan, J.; Xu, K.; Yang, X.; Choy, WY.; Konermann, L. *Anal. Chem*. **2009**, *81*, 5008-5015.

22. Sun J.; Kitova, E. N.; Wang, W.; Klassen, J. S. *Anal. Chem*. **2006**, *78*, 3010-3018.

23. Kitova, E. N.; Kitov, P. I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S. *Glycobiology* **2007**, *17*, 1127-1137.

24. Rademacher, C.; Shoemaker, G. K.; Kim, H.-S.; Zheng, R. B.; Taha, H.; Liu, C.; Nacario, R. C.; Schriemer, D. C.; Klassen, J. S.; Peters, T.; Lowary, T. L. *J. Am. Chem. Soc*. **2007**, *129*, 10489-105029.

25. Shoemaker, G. K.; Soya, N.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2008**,

18, 587-592.

26. Hyland, R. M.; Sun, J.; Griener, T. P.; Mulvey, G. L.; Klassen, J. S.; Donnenberg, M. S.; Armstrong, G. D. *Cell. Microbiol.* **2008**, *10*, 177-187.

27. Sun, N.; Soya, N.; Kitova, E. N.; Klassen, J.S. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 472-481.

28. Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem*. **2005**, *77*, 3060-3071.

29. Simons, J.-W.F.A.; Kosters, H.A.; Visschers, R. W.; De Jongh, H. H. J. *Arch. Biochem. Biophys.* **2002**, *406*, 143-152.

30. Zdanov, A.; Bundle, D. R.; Deng, S. - J.; MacKenzie, C. R.; Narang, S. A.; Young, M. N.; Cygler, M. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 6423-6427.

Chapter 3

Dissociation Kinetics of the Streptavidin-Biotin Interaction Measured using Direct Electrospray Ionization Mass Spectrometry Analysis[*]

3.1 Introduction

Non-covalent interactions between proteins and between proteins with other biopolymers, small molecules or metal ions are critical to most cellular processes. The abundances of protein complexes and their lifetimes reflect the rates of the corresponding association and dissociation reactions. Quantification of the kinetic parameters - the association and dissociation rate constants (k_{on} and k_{off}, respectively) - under specific solution conditions (e.g. pH, temperature, ionic strength) is important in understanding the structure and function of protein complexes and is relevant to drug design [1-3]. There exist a number of established experimental techniques for measuring the rates of biochemical reactions, including association and dissociation reactions. These include SPR[4-5], spectroscopic methods (e.g., atomic force spectroscopy, circular dichroism or fluorescence-based approaches)[6-10], kinetic capillary electrophoresis[11-12],

---

[*] A version of this chapter has been published: Deng L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **Published online Dec 18, 2012**.

radiolabeling combined with filtration/dialysis[13-14] and NMR[15]. Many of these techniques require the labelling of one of the binding partners, their attachment to a surface or some other manipulation of the system, which can complicate the interpretation of the kinetic data and, in some instances, influence the rates of the reactions being investigated.[16]

ESI-MS has emerged as an important addition to the arsenal of techniques available for measuring the kinetics of chemical and biochemical reactions.[17-28] The ESI-MS approach is attractive as there is no requirement for labelling or immobilization since the identity of reactants and products and, possibly intermediates, can usually be established directly from the measured *m/z* values.[18,29] Moreover, ESI-MS analysis allows for multiple reactions to be monitored simultaneously, a feature not associated with most kinetic assays. The determination of reaction rates by ESI-MS analysis normally follows one of two general strategies: on-line (real-time) monitoring of the reaction mixture, and off-line analysis, usually following a quench step that stops the reaction. The advantage of the on-line approach is that it allows, in principle, for direct analysis of the time-dependent distribution of reactants, intermediates and products. The minimum acquisition time for an ESI mass spectrum, which is typically in the s - min range (although it varies between instruments and the nature of the sample being analyzed), places restrictions on the speed of reactions that can be reliably analyzed using the on-line approach. For this reason, real-time ESI-MS kinetic measurements are most commonly applied to relatively slow reactions, with timescales >min. However, there are examples where ESI-MS has been successfully applied to relatively fast reactions, in the ms – s range.[17-18,20,22,24-25] The measurement of fast kinetics typically requires the use of rapid mixing systems, such as a continuous-flow[17-18,20,22], rapid quenched-flow apparatus[24] or stopped-flow[25-26]. The off-line approach is generally easier to implement and affords greater flexibility in terms of the experimental conditions under which the reactions are carried out. For example, this approach is suitable for the analysis of enzyme kinetics under solution conditions that are not amenable to direct ESI-MS analysis, such as high concentrations of salts or non-volatile buffers (e.g. PBS, citrate, HEPES or TRIS) that are commonly used to stabilize proteins and ensure relevance to physiological conditions. Following the quench step, the solvent composition can be altered in order to facilitate detection of reactants or products by ESI-MS. However, a limitation of the off-line approach is that information on the distribution of species present under the reaction conditions of interest may be lost.

ESI-MS has been used to study the reaction rates for a variety of non-covalent protein interactions, including protein-protein [19-20,30], protein-small molecule [31] and protein-metal ion complexes [32], as well as for other biological complexes, such as DNA duplexes [33-34]. However, to the best of our knowledge, absolute values of k_{on} and k_{off} for protein-ligand interactions measured using this approach have not been previously reported. Here, we describe the application ESI-MS for quantifying k_{off} for the high affinity interaction between biotin (B) and a truncated form (containing residues 13-139) of WT streptavidin. Streptavidin is a homotetrameric protein complex (S_4) that is isolated from

Streptomyces avidinii [35]. Each streptavidin subunit is organized into an 8-stranded β-barrel, with a binding site for B at one end [36]. The streptavidin-biotin interaction is one of the most stable in nature and the exceptionally high affinity (K_a of ~2.5×10^{13} M^{-1} at pH 7.4 and 25 °C) arises from an unusually small dissociation rate constant (5.4×10^{-6} s^{-1} at pH 7.4 and 25 °C) [13]. The origin of the slow dissociation kinetics has been the focus of many experimental and theoretical studies [13-14,37-40]. For example, the temperature dependence of k_{off} for the interaction between B and WT streptavidin has been compared with values measured for a variety of single site mutants in an effort to elucidate the influence of the specific amino acid side chains on the kinetic barrier to dissociation [13-14,41]. These data were measured using a radiolabeled B assay, whereby the release of bound B from (S_4 + 4B) complex was monitored in the presence of an excess of unlabeled B. In the present study, values of k_{off} for the sequential loss of B from the (S_4 + 4B) complex, at pH 7 and temperatures ranging from 15 to 45 °C, were measured using ESI-MS. Two different approaches were used to analyze the ESI-MS data, one based on the initial rate of change in the relative abundance of the (S_4 + 4B) species, the other based on non-linear fitting of the time-dependent relative abundances of all free and B-bound S_4 species. The Arrhenius parameters determined from the k_{off} values measured by ESI-MS were compared to the values obtained using the radiolabeled B assay [14]. In addition to providing a quantitative measure of the dissociation rate constants, the direct ESI-MS measurements provide a definitive answer to the question of whether the sequential binding of B to S_4 occurs in a cooperative or non-cooperative fashion.

3.2 Materials and Methods

3.2.1 Sample preparation

The plasmid for natural core streptavidin (containing residues 13-139 of WT streptavidin, MW 13 271 Da) was a gift from Prof. P. Stayton (University of Washington). Streptavidin was expressed in *E. coli* and purified using procedures described elsewhere.[42] Solutions of purified S_4 were exchanged directly into 100 mM aqueous ammonium acetate buffer using an Amicon microconcentrator with a MW cut-off of 10 kDa and lyophilized. Stock solutions of S_4 (100 μM) were prepared by dissolving a known amount of lyophilized streptavidin into 100 mM ammonium acetate and stored at -20 ℃ until needed. Biotin (B, MW 244.3 Da) was purchased from Sigma-Aldrich Canada (Oakville, Canada). The stock solution of B (800 μM) was prepared by dissolving B in Milli-Q water. All stock solutions were stored at -20 ℃ until needed.

3.2.2 Mass spectrometry

All measurements were performed using an Apex Qe 9.4 T FTICR mass spectrometer (Bruker, Billerica, MA). Nanoflow ESI (nanoESI) was performed using borosilicate tubes (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-97 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in positive ion mode was established by applying a voltage of 1.0 – 1.3 kV to a platinum wire inserted inside the glass tip. The solution flow rate was typically ~20 nL min^{-1}. Details of the instrumental and experimental conditions typically used for quantifying protein-ligand interactions can be found elsewhere.[43]

3.2.3 Kinetic measurements

For the kinetic measurements, the reaction mixtures were prepared by mixing aliquots of the stock solutions to achieve the desired concentrations of S_4 (10 μM), B (10 - 26 μM) and ammonium acetate (5 mM). The reaction mixtures were kept at constant temperature (15 - 45 ℃) using a water bath (Colora, Germany). Aliquots of the reaction mixtures were removed at specific reaction times (t) and analyzed by ESI-MS. Three ESI mass spectra were measured for each aliquot. An acquisition time of approximately 1 min was used for each mass spectrum.

3.2.4 Data analysis

Assuming that the four ligand binding sites of S_4 are kinetically equivalent and independent, *vide infra*, the apparent rate constants for ligand association and dissociation reactions ($k_{on,i}$ and $k_{off,(i+1)}$, respectively) for each (S_4 + iB) species (eq 3.1) are related to the intrinsic (microscopic) rate constants (k_{on} and k_{off}) through statistical factors, which reflect the number of free and occupied binding sites, eqs 3.2a and 3.2b:

$$(S_4 + iB) + B \underset{k_{off,(i+1)}}{\overset{k_{on,i}}{\rightleftarrows}} (S_4 + (i+1)B) \tag{3.1}$$

$$k_{on,i} = (4- i)k_{on} \tag{3.2a}$$

$$k_{off,(i+1)} = (i+1)k_{off} \tag{3.2b}$$

It follows that the rate of change of the concentration of each (S_4 + iB) species can be described by eqs 3.3a-3.3e:

$$\frac{d([\mathrm{S_4}])}{dt} = -4k_{on}[\mathrm{B}][\mathrm{S_4}] + k_{off}[\mathrm{S_4}+\mathrm{B}] \tag{3.3a}$$

$$\frac{d([\mathrm{S_4}+\mathrm{B}])}{dt} = -(3k_{on}[\mathrm{B}] + k_{off})[\mathrm{S_4}+\mathrm{B}] + 4k_{on}[\mathrm{B}][\mathrm{S_4}] + 2k_{off}[\mathrm{S_4}+2\mathrm{B}] \tag{3.3b}$$

$$\frac{d([\mathrm{S_4}+2\mathrm{B}])}{dt} = -(2k_{on}[\mathrm{B}] + 2k_{off})[\mathrm{S_4}+2\mathrm{B}] + 3k_{on}[\mathrm{B}][\mathrm{S_4}+\mathrm{B}] + 3k_{off}[\mathrm{S_4}+3\mathrm{B}] \tag{3.3c}$$

$$\frac{d([\mathrm{S_4}+3\mathrm{B}])}{dt} = -(k_{on}[\mathrm{B}] + 3k_{off})[\mathrm{S_4}+3\mathrm{B}] + 2k_{on}[\mathrm{B}][\mathrm{S_4}+2\mathrm{B}] + 4k_{off}[\mathrm{S_4}+4\mathrm{B}] \tag{3.3d}$$

$$\frac{d([\mathrm{S_4}+4\mathrm{B}])}{dt} = -4k_{off}[\mathrm{S_4}+4\mathrm{B}] + k_{on}[\mathrm{B}][\mathrm{S_4}+3\mathrm{B}] \tag{3.3e}$$

where $[\mathrm{S_4}+i\mathrm{B}]$ is the concentration of the ($\mathrm{S_4}$ + iB) species and [B] is the concentration of free B.

Because the rate of association of B to $\mathrm{S_4}$ is very high, [B] will be extremely small when sub-stoichiometric amounts of ligand are used (which is the case in the present study). Consequently, it is reasonable to assume that [B] exists at a steady-state over the course of the entire reaction. Therefore, the ik_{on}[B] terms can be approximated as ik_1, where k_1 is a pseudo first-order rate constant. Eqs 3a-3e can then be rewritten in terms of the normalized abundances ($A_{R(\mathrm{S_4}+i\mathrm{B})}$) of the ($\mathrm{S_4}$ + iB) species, eqs 3.4a-3.4e:

$$\frac{dA_{R\mathrm{S_4}}}{dt} = -4k_1 A_{R\mathrm{S_4}} + k_{off} A_{R(\mathrm{S_4}+\mathrm{B})} \tag{3.4a}$$

$$\frac{dA_{R(\mathrm{S_4}+\mathrm{B})}}{dt} = -(3k_1 + k_{off}) A_{R(\mathrm{S_4}+\mathrm{B})} + 4k_1 A_{R\mathrm{S_4}} + 2k_{off} A_{R(\mathrm{S_4}+2\mathrm{B})} \tag{3.4b}$$

$$\frac{dA_{R(\mathrm{S_4+2B})}}{dt} = -(2k_1 + 2k_{off})A_{R(\mathrm{S_4+2B})} + 3k_1 A_{R(\mathrm{S_4+B})} + 3k_{off} A_{R(\mathrm{S_4+3B})} \tag{3.4c}$$

$$\frac{dA_{R(\mathrm{S_4+3B})}}{dt} = -(k_1 + 3k_{off})A_{R(\mathrm{S_4+3B})} + 2k_1 A_{R(\mathrm{S_4+2B})} + 4k_{off} A_{R(\mathrm{S_4+4B})} \tag{3.4d}$$

$$\frac{dA_{R(\mathrm{S_4+4B})}}{dt} = -4k_{off} A_{R(\mathrm{S_4+4B})} + k_1 A_{R(\mathrm{S_4+3B})} \tag{3.4e}$$

Linear fitting

As described below, under the solution conditions used in the present study, there is a significant difference in the ligand association and dissociation rates immediately upon mixing S_4 and B. [9,13-14] Because of this and the fact that the initial concentration of B is less than the total concentration of binding sites, mixing S_4 with B initially produces a non-equilibrium distribution of (S_4 + *i*B), one that favours the S_4 and (S_4 + 4B) species. As the reaction proceeds, the relative abundance of (S_4 + 4B) will decrease and system will eventually achieve an equilibrium distribution of (S_4 + *i*B) species. Because $A_{R(\mathrm{S_4+3B})}$ is initially relatively small, the rate of change of $A_{R(\mathrm{S_4+4B})}$, described by eq 4e, can be approximated by eq 3.5:

$$\frac{dA_{R(\mathrm{S_4+4B})}}{dt} \approx -4k_{off} A_{R(\mathrm{S_4+4B})} \tag{3.5}$$

and k_{off} can be evaluated from a linear least squares fit of the plot of the natural logarithm of $A_{R(\mathrm{S_4+4B})}$ versus *t*, eq 3.6:

$$\ln A_{R(\mathrm{S_4+4B})} = -4k_{off} \times t + b \tag{3.6}$$

where b is a constant that is equal to $\ln A_{R(S_4+4B)}$ at t = 0. In this case, t = 0 corresponds to the earliest reaction time for which ESI mass spectra were acquired. It should be noted that the magnitude of b depends on the initial concentrations of S_4 and B.

Nonlinear fitting

An alternative approach to determining k_{off} involves applying non-linear regression analysis to the time-dependence of $A_{R(S_4+iB)}$ for each (S_4 + iB) species. This approach is more general than the initial rate method described above since there are no simplifying assumptions needed and k_1, in addition to k_{off}, can be determined. Moreover, this approach is not limited to data measured early in the reaction. In fact, inclusion of data measured at longer times, where the system is approaching equilibrium, enhances the reliability of the fitting procedure. Expressions for the time-dependent $A_{R(S_4+iB)}$ for each (S_4 + iB) species are obtained by solving eqs 3.4a-3.4e (as a system) using Maple 14 (Maplesoft, Waterloo, Canada). The experimental $A_{R(S_4+iB)}$ values at t = 0 (the earliest time point measured) served as a boundary conditions.

Shown in eqs 3.7a-3.7e is a set of solutions (functions) corresponding to experimental data acquired at 44.8 ℃. The experimental values of $A_{R(S_4+iB)}$ at t = 0 (the earliest time point measured) acquired at 44.8 ℃ were applied as boundary conditions: A_{RS_4} = 0.648, $A_{R(S_4+B)}$ = 0.084, $A_{R(S_4+2B)}$ = 0.056, $A_{R(S_4+3B)}$ = 0.045, $A_{R(S_4+4B)}$ = 0.167.

$$
\begin{aligned}
&A_{RS_4}(t) = k_{off}^4/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+(1/1000)\times(3001k_1-999k_{off})\times k_{off}^3\times\exp(-(k_1+\\
&k_{off})\times t)/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+(1/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247k_{off}^2\times k_1-713k_{off}^3)\times\\
&k_{off}\times\exp(-3(k_1+k_{off})\times t)/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}))+(1/1000)\times(167k_{off}^4\\
&+648k_1^4+56k_{off}^2\times k_1^2-84k_{off}\times k_1^3-45k_{off}^3\times k_1)\times\exp(-(4(k_1+k_{off})\times t)/(6k_{off}^2\times k_1^2+4k_{off}^3\times k_1\\
&+k_1^4+4k_{off}\times k_1^3+k_{off}^4)+(1/1000)\times(4196k_1^2-611k_{off}\times k_1+1193k_{off}^2)\times k_{off}^2\times\exp(-(2(k_1+\\
&k_{off})\times t)/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2
\end{aligned}
\tag{3.7a}
$$

$$
\begin{aligned}
&A_{R(S_4+B)}(t) = ((3/1000)\times(3001k_1-999k_{off})\times k_{off}^3\times\exp(-(k_1+k_{off})\times t)\times k_1/(k_1^2+2k_{off}\times k_1\\
&+k_{off}^2)^2-(1/1000)\times(3001k_1-999k_{off})\times k_{off}^4\times\exp(-(k_1+k_{off})\times t)/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+\\
&(1/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247k_{off}^2\times k_1-713k_{off}^3)\times k_{off}\times\exp(-(3(k_1+k_{off})\times t)\times k_1\\
&/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}))-(3/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247k_{off}^2\\
&\times k_1-713k_{off}^3)\times k_{off}^2\times\exp(-(3(k_1+k_{off})\times t)/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}))-\\
&(1/250)\times(167k_{off}^4+648k_1^4+56k_{off}^2\times k_1^2-84k_{off}\times k_1^3-45k_{off}^3\times k_1)\times\exp(-(4(k_1+k_{off})\times t)\\
&\times k_{off}/(6k_{off}^2\times k_1^2+4k_{off}^3\times k_1+k_1^4+4k_{off}\times k_1^3+k_{off}^4)-(1/500)\times(4196k_1^2-611k_{off}\times k_1+1193\\
&k_{off}^2)\times k_{off}^3\times\exp(-2(k_1+k_{off})\times t)/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+(1/500)\times(4196k_1^2-611k_{off}\times k_1+\\
&1193k_{off}^2)\times k_{off}^2\times\exp(-2(k_1+k_{off})\times t)\times k_1/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+4k_1\times k_{off}^4/(k_1^2+2k_{off}\times k_1+\\
&k_{off}^2)^2)/k_{off}
\end{aligned}
\tag{3.7b}
$$

$$
\begin{aligned}
&A_{R(S_4+2B)}(t) = ((3/1000)\times k_1^2\times(3001k_1-999k_{off})\times k_{off}^3\times\exp(-(k_1+k_{off})\times t)/(k_1^2+2k_{off}\times k_1\\
&+k_{off}^2)^2+(1/1000)\times k_1^2\times(4196k_1^2-611k_{off}\times k_1+1193k_{off}^2)\times k_{off}^2\times\exp(-(2(k_1+k_{off})\times t)/(k_1^2
\end{aligned}
$$

$$+2k_{off}\times k_1+k_{off}^2)^2+(3/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247k_{off}^2\times k_1-713k_{off}^3)\times k_{off}^3\times\exp(-$$

$$(3(k_1+k_{off})\times t)/((k_1^3+3k_{off}k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}))+(3/500)\times(167k_{off}^4+648k_1^4+56$$

$$k_{off}^2\times k_1^2-84k_{off}\times k_1^3-45k_{off}^3\times k_1)\times\exp(-4\,(k_1+k_{off})\times t)\times k_{off}^2/(6k_{off}^2k_1^2+4k_{off}^3\times k_1+k_1^4+$$

$$4k_{off}\times k_1^3+k_{off}^4)+(1/1000)\times(4196k_1^2-611k_{off}\times k_1+1193k_{off}^2)\times k_{off}^4\times\exp(-2\,(k_1+k_{off})\times t)$$

$$/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+6k_1^2\times k_{off}^4/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2-(3/1000)\times(3001k_1-999k_{off})\times$$

$$k_{off}^4\times\exp(-(k_1+k_{off})\times t)\times k_1/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2-(3/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247$$

$$k_{off}^2\times k_1-713k_{off}^3)\times k_{off}^2\times\exp(-3(k_1+k_{off})\times t)\times k_1/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+$$

$$k_{off}))-(1/250)\times(4196k_1^2-611k_{off}\times k_1+1193k_{off}^2)\times k_{off}^3\times\exp(-2(k_1+k_{off})\times t\times k_1/(k_1^2+2k_{off}\times$$

$$k_1+k_{off}^2)^2)/k_{off}^2 \tag{3.7c}$$

$$A_{R(\mathrm{S}_4+3\mathrm{B})}(t)\ = -(-4k_1^3\times k_{off}^4/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2-(1/1000)\times(3001k_1-999\times k_{off})\times k_{off}^3$$

$$\times\exp(-(k_1+k_{off})\times t)\times k_1^3/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+(1/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247$$

$$k_{off}^2\times k_1-713k_{off}^3)\times k_{off}^4\times\exp(-(3(k_1+k_{off})\times t)/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}$$

$$))+(1/250)\times(167k_{off}^4+648k_1^4+56k_{off}^2\times k_1^2-84k_{off}\times k_1^3-45k_{off}^3\times k_1)\times\exp(-(4(k_1+k_{off})\times$$

$$t)\times k_{off}^3/(6k_{off}^2\times k_1^2+4k_{off}^3\times k_1+k_1^4+4k_{off}\times k_1^3+k_{off}^4)+(3/1000)\times k_1^2\times(3001k_1-999k_{off})\times$$

$$k_{off}^4\times\exp(-(k_1+k_{off})\times t)/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2-(3/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247k_{off}^2$$

$$\times k_1-713k_{off}^3)\times k_{off}^3\times\exp(-(3(k_1+k_{off})\times t)\times k_1/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}))-$$

$$(1/500)\times(4196k_1^2-611k_{off}\times k_1+1193k_{off}^2)\times k_{off}^4\times\exp(-(2(k_1+k_{off})\times t)\times k_1/(k_1^2+2k_{off}\times k_1$$

$$+k_{off}^2)^2+(1/500)\times k_1^2\times(4196k_1^2-611k_{off}\times k_1+1193k_{off}^2)\times k_{off}^3\times\exp(-2(k_1+k_{off})\times t)/(k_1^2+$$

$$2k_{off}\times k_1+k_{off}^2)^2)/k_{off}^3 \tag{3.7d}$$

$$A_{R(\mathrm{S}_4+4\mathrm{B})}(t)\ = (k_1^4\times k_{off}^4/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2+(1/1000)\times(167k_{off}^4+648k_1^4+56k_{off}^2$$

$$\times k_1^2-84k_{off}\times k_1^3-45k_{off}^3\times k_1)\times\exp(-(4(k_1+k_{off})\times t)\times k_{off}^4/(6k_{off}^2\times k_1^2+4k_{off}^3\times k_1+k_1^4+$$

$$4k_{off}\times k_1^3+k_{off}^4)-(1/1000)\times(3001k_1-999k_{off})\times k_{off}^4\times\exp(-(k_1+k_{off})\times t)\times k_1^3/(k_1^2+2k_{off}\times k_1$$

$$+k_{off}^2)^2-(1/1000)\times(2676k_1^3-364k_{off}\times k_1^2+247k_{off}^2\times k_1-713k_{off}^3)\times k_{off}^4\times\exp(-(3(k_1+k_{off})$$

$$\times t)\times k_1/((k_1^3+3k_{off}\times k_1^2+3k_{off}^2\times k_1+k_{off}^3)\times(k_1+k_{off}))+(1/1000)\times k_1^2\times(4196k_1^2-611k_{off}\times k_1$$

$$+1193k_{off}^2)\times k_{off}^4\times\exp(-(2(k_1+k_{of})\times t)/(k_1^2+2k_{off}\times k_1+k_{off}^2)^2)/k_{off}^4 \tag{3.7e}$$

Origin (OriginLab, Northampton, MA) was used to fit the functions (with k_{off} and k_1 as adjustable parameters) to the experimental breakdown curves ($A_{R(S_4+iB)}$ values plotted versus t). At each temperature, k_{off} and k_1 values were calculated for each (S_4 +iB) species; the reported k_{off} and k_1 values in Table 3.1 correspond to the average of these values.

The temperature dependence of the measured k_{off} values was analyzed according to the Arrhenius equation, eq 3.8:

$$\ln(k_{off}) = -\frac{E_a}{RT} + \ln(A) \tag{3.8}$$

The E_a and A were calculated from the slope and intercept, respectively, of a linear least-squares fit of the plot of $\ln(k_{off})$ versus 1/T.

3.3 Results and Discussion

Shown in Figures 3.1a to 3.1c are representative ESI mass spectra acquired in positive ion mode for a solution (pH 7 and 22.1 °C) of S_4 (10 μM) and B (14 μM) and ammonium acetate (5 mM) immediately after mixing (i.e., reaction time ~ 0 min), and after 112 min and 1602 min (1.1 days). In each mass spectrum, signals corresponding to the protonated $(S_4 + iB)^{n+}$ ions with $0 \leq i \leq 4$, at n = 12 – 16, are evident. Initially, the S_4^{n+} and $(S_4 + 4B)^{n+}$ ions represent the dominant species present (Figure 3.1a). The observation of predominantly free and fully ligand-bound protein shortly after mixing B with S_4 is consistent with results of Sano and Cantor [44]. These authors used gel electrophoresis to analyze solutions of B and S_4, immediately after mixing, and observed only two major bands, which corresponded to S_4 and (S_4 + 4B). [44] The authors interpreted these results as evidence of cooperative ligand binding [44]. However, the observation of abundant S_4 and (S_4 + 4B) in the ESI mass spectra does not necessarily imply cooperative binding of B to S_4. Instead, mixing of S_4 and B could lead initially to a non-equilibrium distribution of streptavidin-biotin species due to the fast association kinetics and extremely slow dissociation kinetics. The latter explanation finds support in the observation that the distribution of (S_4 + *i*B) species changes at longer reaction times. For example, after 112 min, abundant signal is observed for all five of the (S_4 + *i*B) species, i.e., with *i* from 0 to 4, although S_4 remains the dominant species detected (Figure 3.1b); after 1602 min, the (S_4 + B) species dominates (Figure 3.1c). At much longer times, a constant distribution of (S_4 + *i*B) species is observed, indicating that an equilibrium distribution of (S_4 + *i*B) species was reached. As an example, shown in Figure 1d is a representative ESI mass spectrum acquired after 13080 min (9 days). Notably, the measured distribution of (S_4 + *i*B) species agrees with the distribution expected in the case of four identical and independent ligand binding sites, each with a microscopic K_a of 2.5×10^{13} M^{-1} (at pH 7.4 and 25 ℃) (Figure 3.1e) [13]. These results establish, unambiguously, that B binding to S_4 is not a cooperative process, in agreement with the findings of Jones and Kurzban [45] and Fidelio and co-workers [46], and that the binding sites are thermodynamically equivalent and independent. It is also concluded that the apparent cooperative distributions of the

(S_4 + iB) species observed at short reaction times are, in fact, of kinetic and not thermodynamic origin.

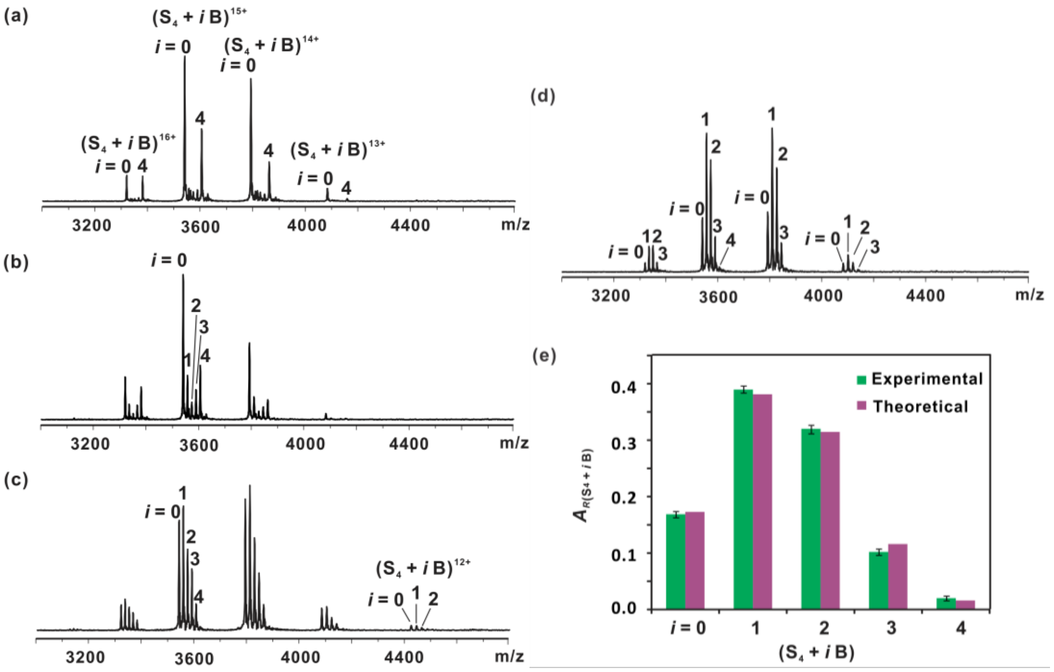


Figure 3.1. ESI mass spectra acquired for an aqueous ammonium acetate (5 mM) solution of S_4 (10 μM) and B (14 μM) at 22.1 °C and pH 6.8 and different reaction times (a) 0 sec (b) 112 min, (c) 1602 min (1.1 days) and (d) 13080 min (9 days). (e) Normalized distribution of (S_4 + i B) species, where $i = 0 – 4$, determined from the ESI mass spectrum shown in (d). The reported errors correspond to one standard deviation and were determined from 3 replicate measurements. Also shown is the calculated distribution for four equivalent ligand binding sites, each with a K_a of 2.5×10^{13} M^{-1}.

As noted in the Data Analysis section, two different approaches were used to quantify k_{off}. One approach is based on the initial rate of change of $A_{R(\mathrm{S}_4+i\mathrm{B})}$. Shown in Figure 3.2a are plots of the natural logarithm of $A_{R(\mathrm{S}_4+i\mathrm{B})}$ versus t, measured at the reaction temperatures indicated. Importantly, the plots exhibit excellent linearity. This result indicates that neglect of the ligand association reaction involving (S_4 + 3B) to $A_{R(\mathrm{S}_4+4\mathrm{B})}$ in eq 3.4e is a reasonable assumption. The k_{off} values calculated at each reaction temperature are listed in Table 3.1. Notably, the value of 5.0×10^{-5} s^{-1} determined at 36.2 ℃ agrees very well with the reported value of 4.1×10^{-5} s^{-1}, which was measured at 37 ℃ [14]. An alternative approach used to determine k_{off} involves non-linear fitting of the solutions of eqs 3.4a-3.4e to the time-dependent $A_{R(\mathrm{S}_4+i\mathrm{B})}$ values. Plotted in Figure 3.2b are the $A_{R(\mathrm{S}_4+i\mathrm{B})}$ values measured at 44.8 ℃ and the curves obtained from the non-linear fitting procedure. It can be seen that the calculated curves describe the experimental data very well. The average kinetic parameters, determined at each temperature investigated, are summarized in Table 3.1. Notably, the k_{off} values determined by the two different methods agree very well, within a factor of 2, at all of the temperatures investigated. For example, a k_{off} value of 5.0×10^{-5} s^{-1} was determined at 36.2℃ with the non-linear fitting method, which is indistinguishable from the value of 5.1×10^{-5} s^{-1} determined from the initial rates approach. With the non-linear fitting method, the k_1 ($=k_{on}$[B]) terms were also established at each temperature investigated (Table 3.1). Although [B] is not accurately known (and, in fact, varies over the course of the reaction), it is nevertheless possible to estimate k_{on} by assuming that [B] (at all reaction times) is similar in magnitude to the $[B]_{eq}$, the concentration of free B at equilibrium. Following this approach, k_{on} was estimated to be 1.3×10^{8} $M^{-1}s^{-1}$ at 22.1 ℃ using a $[B]_{eq}$ of 2.2×10^{-14} M, which was calculated for a solution of S_4 (10 μM) and B (14 μM) and a microscopic K_a of 2.5×10^{13} M^{-1} [13]. This value of k_{on} agrees reasonably well with a value of 4.5×10^{7} $M^{-1}s^{-1}$, which was determined from measurements carried out at ambient temperature (not specified) using droplet microfluidics integrated with a confocal fluorescence detection system [9].

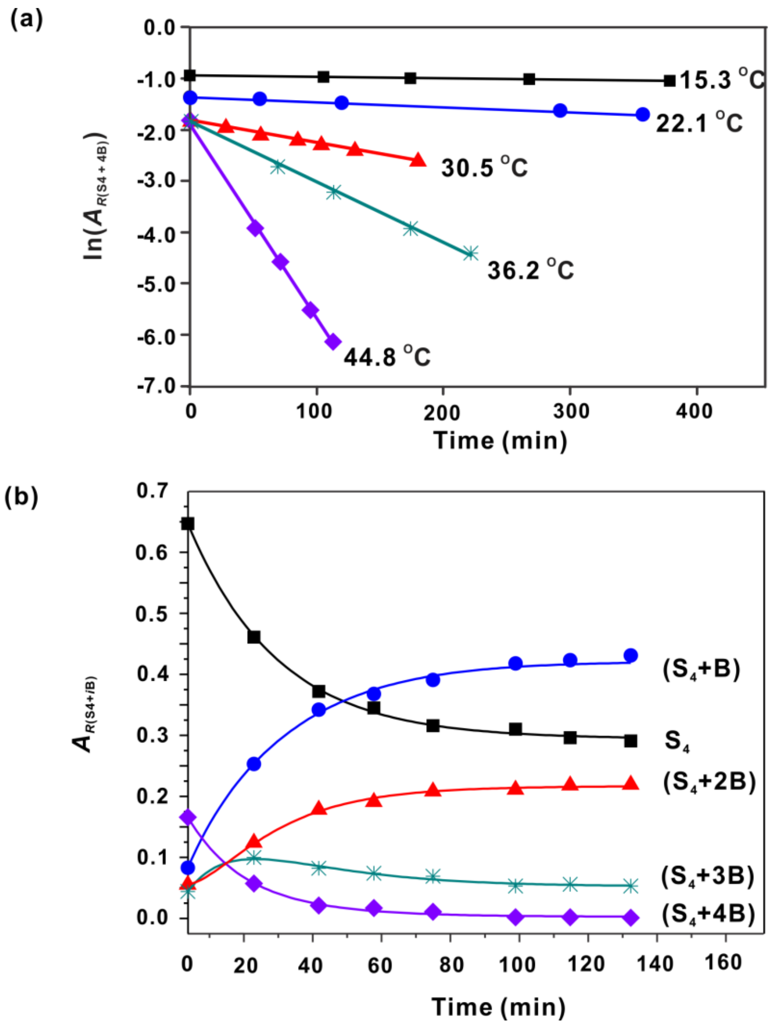


Figure 3.2. (a) Plots of the natural logarithm of $A_{R(S_4+4B)}$ versus reaction time measured by ESI-MS for neutral aqueous ammonium acetate (5 mM) solutions of S_4 (10 μM) and B (10 - 20 μM) at 15.3 ℃, 22.1 ℃, 30.5 ℃, 36.2 ℃ and 44.8 ℃. The solid curves represent linear least squares fits of the experimental data. (b) Plots of $A_{R(S_4+iB)}$ versus reaction time measured by ESI-MS neutral aqueous ammonium acetate (5 mM) solution of S_4 (10 μM), B (10 μM) at 44.8 ℃. The solid curves were determined from non-linear regression analysis to the time-dependent $A_{R(S_4+iB)}$ values.

Table 3.1. Microscopic rate constants (k_{off}) for the dissociation of the streptavidin-biotin interaction at pH 6.8 and temperatures ranging from 15 to 45 °C measured using direct ESI-MS analysis.[a,b]

$[S_4]_o$ (μM)	**$[B]_o$ (μM)**	**T (°C)**	**k_{off} (s^{-1}) [a] Linear fitting**	**k_{off} (s^{-1}) [b] Non-linear fitting**	**k_1 (s^{-1}) [b] Non-linear fitting**
10	20	15.3	$(1.1 \pm 0.1) \times 10^{-6}$	$(1.3 \pm 0.7) \times 10^{-6}$	$(2.3 \pm 0.8) \times 10^{-6}$
10	14	22.1	$(4.1 \pm 0.1) \times 10^{-6}$	$(5.5 \pm 0.3) \times 10^{-6}$	$(2.8 \pm 0.9) \times 10^{-6}$
10	10	30.5	$(1.8 \pm 0.1) \times 10^{-5}$	$(2.1 \pm 0.1) \times 10^{-5}$	$(8.7 \pm 0.3) \times 10^{-6}$
10	10	36.2	$(5.0 \pm 0.2) \times 10^{-5}$	$(5.1 \pm 0.1) \times 10^{-5}$	$(1.8 \pm 0.3) \times 10^{-5}$
10	10	44.8	$(1.6 \pm 0.1) \times 10^{-4}$	$(2.5 \pm 0.2) \times 10^{-4}$	$(8.4 \pm 0.7) \times 10^{-5}$

a. k_{off} was obtained from the slope (k_{off} = ¼ slope) of a linear least squares fit of the plot of the natural logarithm of $A_{R(S_4+4B)}$ versus t. Reported errors correspond to one standard deviation. b. At each temperature, k_{off} and k_1 were obtained from non-linear regression analysis of the time-dependent relative abundance of each of the (S_4 + iB) species. The reported k_{off} and k_1 values correspond to the average of these values and errors are one standard deviation.

Shown in Figure 3.3 are the Arrhenius plots constructed from the k_{off} values determined from the ESI-MS data using the two different data analysis approaches. The corresponding Arrhenius parameters (E_a, A) are listed in Table 3.2. Also shown in Figure 3.3 is the calculated curve based on the reported activation enthalpy and entropy, for the loss of B from the (S_4 + 4B) complex in aqueous solution at pH 7.4 [14]. Inspection of Figure 3.3 (and Table 3.2) reveals that the Arrhenius plots for the loss of B from the (S_4 + 4B) complex from linear and nonlinear fitting are similar, with E_a values of 30.4 ± 0.7 and 31.7 ± 0.8 kcal mol^{-1}, respectively. Moreover, the E_a values agree with the reported value of 31.0 ± 0.2 kcal mol^{-1} [14]. These findings indicate that both approaches to the analysis of the time-resolved ESI-MS data can provide a reliable determination of the temperature-dependence of k_{off}. However, the non-linear fitting approach is more general and, in principle, can be applied in cases where sequential ligand binding exhibits cooperativity or where multiple, non-equivalent binding sites are present.

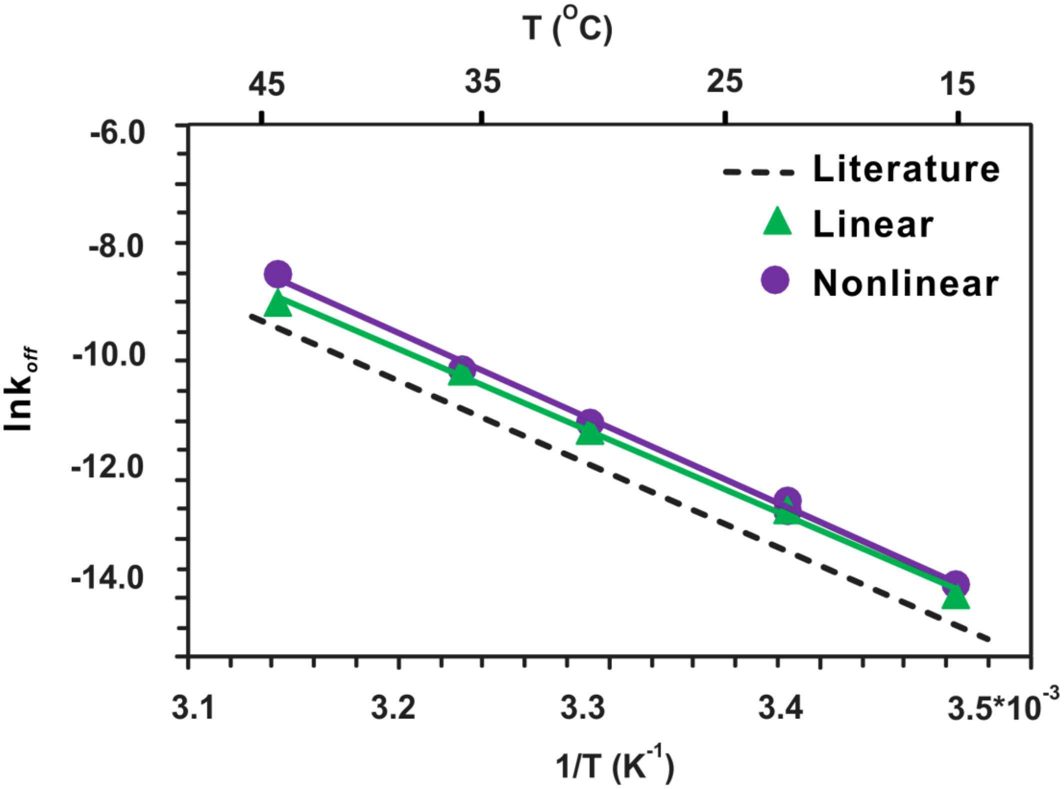


Figure 3.3. Arrhenius plots for the loss of B from the (S_4 + 4B) complex at pH 7 constructed from k_{off} values measured by ESI-MS (linear fitting (▲), non-linear fitting (●)) and corresponding plot (- - -) calculated from the activation enthalpy and entropy reported in [14].

Table 3.2. Arrhenius activation parameters (Ea, A) for the loss of B from the (S4 + 4B) complex determined from k_{off} values measured by ESI-MS (using linear and non-linear data analysis methods) at pH 6.8, and Arrhenius parameters measured at pH 7.4 using a radiolabeled biotin assay.

	E_a **(kcal mol^{-1})**	**A** **(s^{-1})**
ESI-MS (linear fitting)	30.4 ± 0.7 [a]	$10^{17.1\ \pm 0.5\ a}$
ESI-MS (non-linear fitting)	31.7 ± 0.8 [a]	$10^{18.2\ \pm 0.6\ a}$
Radiolabeled biotin assay	31.0 ± 0.2 [b]	$10^{17.3\ \pm 0.1\ b}$

a. Reported errors correspond to one standard deviation. b. Values calculated from the activation enthalpy and entropy reported in [14].

The streptavidin-biotin interaction is unusually kinetically stable and it takes several days to achieve an equilibrium distribution of (S_4 + iB) species at the temperatures investigated. However, it is important to note that this same experimental approach could be applied, in an on-line fashion, to determine k_{off} and, in principle, k_{on} for protein-ligand interactions that require less time to reach an equilibrium distribution. Given that ~1 min is typically required to acquire an ESI mass spectrum (with a high signal-to-noise ratio) for solutions of protein-ligand complexes [47], it should be possible to apply this approach to complexes that take >10 min (under the desired solution conditions) to reach equilibrium.

Although this approach will not be suitable for all protein-ligand complexes (those that exhibit both fast association and dissociation kinetics), the kinetic parameters for many protein-ligand interactions are expected to be accessible with this technique.

3.4 Conclusions

In summary, ESI-MS measurements have been used to quantify k_{off} for the sequential loss of B from the (S_4 + 4B) complex at pH 7 and temperatures ranging from 15 to 45 °C. Two different general strategies for data analysis were considered, one based on the initial rate of dissociation of the (S_4 + 4B) complex, and the other employing non-linear fitting of the time-dependent $A_{R(\mathrm{S}_4+i\mathrm{B})}$ values of the (S_4 + iB) species. The two methods were found to yield k_{off} values that agree within a factor of two. Importantly, the dissociation E_a values measured by ESI-MS agree within 1 kcal mol^{-1} with the reported value, which was measured using a radiolabeled B assay. In addition to providing a quantitative measure of k_{off} at the temperatures investigated, the ESI-MS measurements also revealed, unambiguously, that sequential B binding to S_4 occurs in a non-cooperative fashion and that the four ligand binding sites are kinetically and thermodynamically equivalent and independent.

3.5 References

1. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. *Nat. Rev. Drug Discov.* **2006**, *5*, 730-739.

2. Swinney, D. C. *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 31-39.

3. Callender, R.; Dyer, R. B. *Chem. Rev.* **2006**, *106*, 3031-3042.

4. Daghestani, H. N., Day, B.W. *Sensors* **2010**, *10*, 9630-9646.

5. De Crescenzo, G.; Boucher, C.; Durocher, Y.; Jolicoeur, M. *Cell. Mol. Bioeng.* **2008**, *1*, 204-215.

6. Bizzarri, A. R.; Cannistraro, S. *J. Phy. Chem. B* **2009**, *113*, 16449-16464.

7. Hill, J. J.; Royer, C. A. *Methods Enzymol.* **1997**, *278*, 390-416.

8. Málnási-Csizmadia, A.; Pearson, D. S.; Kovács, M.; Woolley, R. J.; Geeves, M. A.; Bagshaw, C. R. *Biochemistry* **2001**, *40*, 12727-12737.

9. Srisa-Art, M.; Dyson, E. C.; deMello, A. J.; Edel, J. B. *Anal. Chem.* **2008**, *80*, 7063-7067.

10. Sugawara, T.; Kuwajima, K.; Sugai, S. *Biochemistry* **1991**, *30*, 2698-2706.

11. Chu, Y. H.; Cheng, C. C. *Cell. Mol. Life Sci.* **1998**, *54*, 663-683.

12. Hall, D. R.; Winzor, D. J. *J. Chromatogr. B* **1998**, *715*, 163-181.

13. Chilkoti, A.; Stayton, P. S. *J. Am. Chem. Soc.* **1995**, *117*, 10622-10628.

14. Klumb, L. A.; Chu, V.; Stayton, P. S. *Biochemistry* **1998**, *37*, 7657-7663.

15. Gizachew, D.; Dratz, E. *Chem. Biol. Drug Des.* **2011**, *78*, 14-24.

16. Bao, J.; Krylova, S. M.; Wilson, D. J.; Reinstein, O.; Johnson, P. E.; Krylov, S. N. *ChemBioChem* **2011**, *12*, 2551-2554.

17. Konermann, L.; Collings, B. A.; Douglas, D. J. *Biochemistry* **1997**, *36*, 5554-5559.

18. Lee, V. W. S.; Chen, Y.-L.; Konermann, L. *Anal. Chem.* **1999**, *71*, 4154-4159.

19. Sobott, F.; Benesch, J. L. P.; Vierling, E.; Robinson, C. V. *J. Biol. Chem.* **2002**, *277*, 38921-38929.

20. Simmons, D. A.; Wilson, D. J.; Lajoie, G. A.; Doherty-Kirby, A.; Konermann, L. *Biochemistry* **2004**, *43*, 14792-14801.

21. Deng, G.; Sanyal, G. *J. Pharm. Biomed. Anal.* **2006**, *40*, 528-538.

22. Pan, J.; Rintala-Dempsey, A. C.; Li, Y.; Shaw, G. S.; Konermann, L. *Biochemistry* **2006**, *45*, 3005-3013.

23. Sharon, M.; Robinson, C. V. *Annu. Rev. Biochem* **2007**, *76*, 167-193.

24. Clarke, D. J.; Stokes, A. A.; Langridge-Smith, P.; Mackay, C. L. *Anal. Chem.* **2010**, *82*, 1897-1904.

25. Robbins, M. D.; Yoon, O. K.; Barbula, G. K.; Zare, R. N. *Anal. Chem.* **2010**, *82*, 8650-8657.

26. Miao, Z.; Chen, H.; Liu, P.; Liu, Y. *Anal. Chem.* **2011**, *83*, 3994-3997.

27. Pacholarz, K. J.; Garlish, R. A.; Taylor, R. J.; Barran, P. E. *Chem. Soc. Rev.* **2012**, *41*, 4335-4355.

28. Goodlett, D. R.; Ogorzalek Loo, R. R.; Loo, J. A.; Wahl, J. H.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 614-622.

29. Soya, N.; Fang, Y.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2011**, *21*, 547-552.

30. Deroo, S.; Hyung, S.-J.; Marcoux, J.; Gordiyenko, Y.; Koripella, R. K.; Sanyal, S.; Robinson, C. V. *ACS Chem. Biol.* **2012**, *7*, 1120-1127.

31. Peleg-Shulman, T.; Najajreh, Y.; Gibson, D. *J. Inorg. Biochem.* **2002**, *91*, 306-311.

32. Pan, J.; Konermann, L. *Biochemistry* **2010**, *49*, 3477-3486.

33. Rosu, F.; Gabelica, V.; Shin-ya, K.; De Pauw, E. *Chem. Commun.* **2003**, 2702-2703.

34. Gabelica, V. r.; Rosu, F. d. r.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708-6715.

35. L. Chaiet, F. J. W. *Arch. Biochem. Biophys.* **1964**, *106*, 1-5.

36. Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P. *Proceedings of the National Academy of Sciences* **1989**, *86*, 2190-2194.

37. Chu, V.; Stayton, P. S.; Freitag, S.; Le Trong, I.; Stenkamp, R. E. *Protein Sci.* **1998**, *7*, 848-859.

38. Stayton, P. S.; Freitag, S.; Klumb, L. A.; Chilkoti, A.; Chu, V.; Penzotti, J. E.; To, R.; Hyre, D.; Le Trong, I.; Lybrand, T. P.; Stenkamp, R. E. *Biomol. Eng* **1999**, *16*, 39-44.

39. Li, Q.; Gusarov, S.; Evoy, S.; Kovalenko, A. *J. Phy. Chem. B* **2009**, *113*, 9958-9967.

40. Deng, L.; Broom, A.; Kitova, E. N.; Richards, M. R.; Zheng, R. B.; Shoemaker, G. K.; Meiering, E. M.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 16586-16596.

41. Piran, U.; Riordan, W. J. *J. Immunol. Methods* **1990**, *133*, 141-143.

42. Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proceedings of the National Academy of Sciences* **1995**, *92*, 1754-1758.

43. El-Hawiet, A.; Kitova, E. N.; Kitov, P. I.; Eugenio, L.; Ng, K. K.; Mulvey, G. L.; Dingle, T. C.; Szpacenko, A.; Armstrong, G. D.; Klassen, J. S. *Glycobiology* **2011**, *21*, 1217-1227.

44. Sano, T.; Cantor, C. R. *J. Biol. Chem.* **1990**, *265*, 3369-3373.

45. Jones, M. L.; Kurzban, G. P. *Biochemistry* **1995**, *34*, 11750-11756.

46. Gonzalez, M.; Bagatolli, L.; Echabe, I.; Arrondo, J.; Argarana, C.; Cantor, C.; Fidelio, G. *J. Biol. Chem.* **1997**, *272*, 11288-11294

47. Kitova, E.; El-Hawiet, A.; Schnier, P.; Klassen, J. *J. Am. Soc. Mass. Spectrom.* **2012**, *23*, 431-441.

Chapter 4

Kinetic Stability of the Streptavidin-Biotin Interaction Enhanced in the Gas Phase [*]

4.1 Introduction

The timescale of noncovalent protein interactions, such as protein-receptor, antibody-antigen, and multiprotein complexes, are important factors in biochemical processes.[1] Relatedly, the lifetime of drug interactions with protein targets strongly influences their pharmacological activity.[2-3] As a result, the ability to predict and control the kinetic stabilities of protein interactions is of fundamental and practical importance. Currently, the relationship between the structures of protein complexes and their association-dissociation kinetics is poorly understood.[1-3] As is the case with structure-affinity relationships, efforts to understand structure-kinetic relationships are generally hampered by solvent effects, which are exceptionally difficult to probe experimentally. Gas-phase studies of desolvated noncovalent protein complexes afford an opportunity to investigate their intrinsic properties, i.e., their structure and kinetic stability, free

---

[*] A version of this chapter has been published: Deng, L.; Broom, A.; Kitova, E. N.; Richards, M. R.; Zheng, R. B.; Shoemaker, G. K.; Meiering, E.; Klassen, J. S. *J. Am. Chem. Soc.*, **2012**, *134(40)*, 16586-16596.

of solvent effects.[4-13] Furthermore, from a comparison of the structures and stabilities of protein complexes in their hydrated and dehydrated states, it may be possible to gain new insights into the role of water in the binding reactions.[8] Equally important, the results of gas-phase studies of proteins and their complexes serve as important benchmarks to test the reliability of existing computational methods to predict the structure and stability of biological interactions and to guide the development of new theoretical tools.

Recently, detailed investigations into the nature and strength of the intermolecular interactions present in the gaseous ions of a limited number of protein-ligand complexes have been reported.[4,6-9,14-15] The best studied complex is that of a scFv and L1.[16] The BIRD-FGR technique[17-18] was used to identify amino acid residues and sugar hydroxyl groups participating in intermolecular interactions in the gaseous, protonated (scFv + L1)$^{n+}$ ions over a range of charge states (n = 6 – 13) and to quantify these interactions.[4,6] Notably, two of the three H-bond donor-acceptor pairs identified at the lower charge states investigated are also present in the available crystal structures. This finding suggests that at least some of the specific intermolecular H-bonds in the (scFv + L1) complex are conserved upon transfer of the complex from aqueous solution to the gas phase by ESI.[6] Moreover, the strengths of the three H-bonds were found to be in good agreement with theoretical values available for model systems.[6] Recently, the BIRD-FGR approach was applied to complexes of Lg and a series of long FAs.[7-9] Kinetic data for the loss of neutral FAs from deprotonated (Lg + FA)$^{7-}$ ions, together with the results of molecular dynamics (MD) simulations, served as evidence that the acyl chains of the FAs are retained within the hydrophobic ligand binding cavity of Lg in the gas phase.[7] A comparison of the dissociation rate constants measured at 25 ℃ in the gas phase and in aqueous solution revealed that the (Lg + FA) complexes are more stable kinetically in the absence of solvent.[8] The lower kinetic stability in solution was attributed to the hydration of the FA in the *TS*, which results in a significant reduction in the dissociation E_a.[8] Furthermore, a comparison of the gas phase dissociation E_a values with enthalpies reported for the transfer of hydrocarbons from the gas phase to various organic solvents led to the surprising conclusion that the hydrophobic cavity of Lg is, in fact, relatively polar.[9]

Here, we report on the results of a gas phase study of the structure and kinetic stability of the streptavidin-B interaction, which is among the most stable protein-ligand complexes known and has served as a model system for understanding the origin of high affinity protein complexes. Streptavidin is a homotetrameric protein complex (S_4) that is isolated from *Streptomyces avidinii*.[19] Each streptavidin subunit consists of 159 residues, which are organized into an 8-stranded β-barrel, one end of which forms the binding site for B.[20] Analysis of crystal structures obtained for the ligand-bound and unbound S_4,[21-22] together with kinetic and thermodynamic data[20-21,23-29] and the results of MD simulations[30-34] suggests that three structural motifs are responsible for ligand binding: a network of intermolecular H-bonds, van der Waals interactions and the ordering of surface peptide loops upon binding of B. According to the crystal structures available for the (S_4 + 4B) complex, B is stabilized by eight H-bonds, involving residues Asn23, Ser27, Tyr43, Ser45, Asn49, Ser88, Thr90 and Asp128 of streptavidin, and the ureido oxygen and nitrogen, the tetrahydrothiophene sulfur and the carboxyl group of B (Figure 4.1). Three Trp residues in the binding cavity (Trp79, Trp92, and Trp108), together with Trp120 from the adjacent subunit, have been identified as providing an important energetic contribution to the binding of B through van der Waals contacts (Figure 4.1).[34] In addition, a flexible loop (residues 45–52) on each subunit undergoes a disordered-to-ordered conformational change upon binding of B.[21,28-29]

The unusually slow dissociation of the streptavidin-biotin interaction in aqueous solution is responsible for the exceptionally high affinity (K_a of ~2.5×10^{13} M^{-1} at pH 7.4 and 25 ℃).[19,35-36] The origin of the large E_a for ligand escape (31 kcal mol^{-1}),[25] which gives rise to the unusually small dissociation rate constant (k_{off} of 5.4×10^{-6} s^{-1} at pH 7.4 and 25 ℃)[27], has been the focus of many experimental[20-21,23-29] and theoretical studies[30-34] but has not been fully explained. Generally, a large E_a is suggestive of a late *TS*, in which many of the intermolecular interactions are broken or weakened. Based on the results of thermodynamic and kinetic studies it has been proposed that at least four of the intermolecular H-bonds, involving residues Asn23, Ser27, Tyr43 and Asp128, are broken in the *TS*.[25-26] The H-bond involving Ser45 (in the bound form) is also thought to be cleaved during dissociation but may be replaced with a new H-bond (to the ureido oxygen) in the *TS*.[24,26] It has also been proposed that the interactions involving Trp79 and Trp120, but not Trp108, are absent in the *TS*.[27]

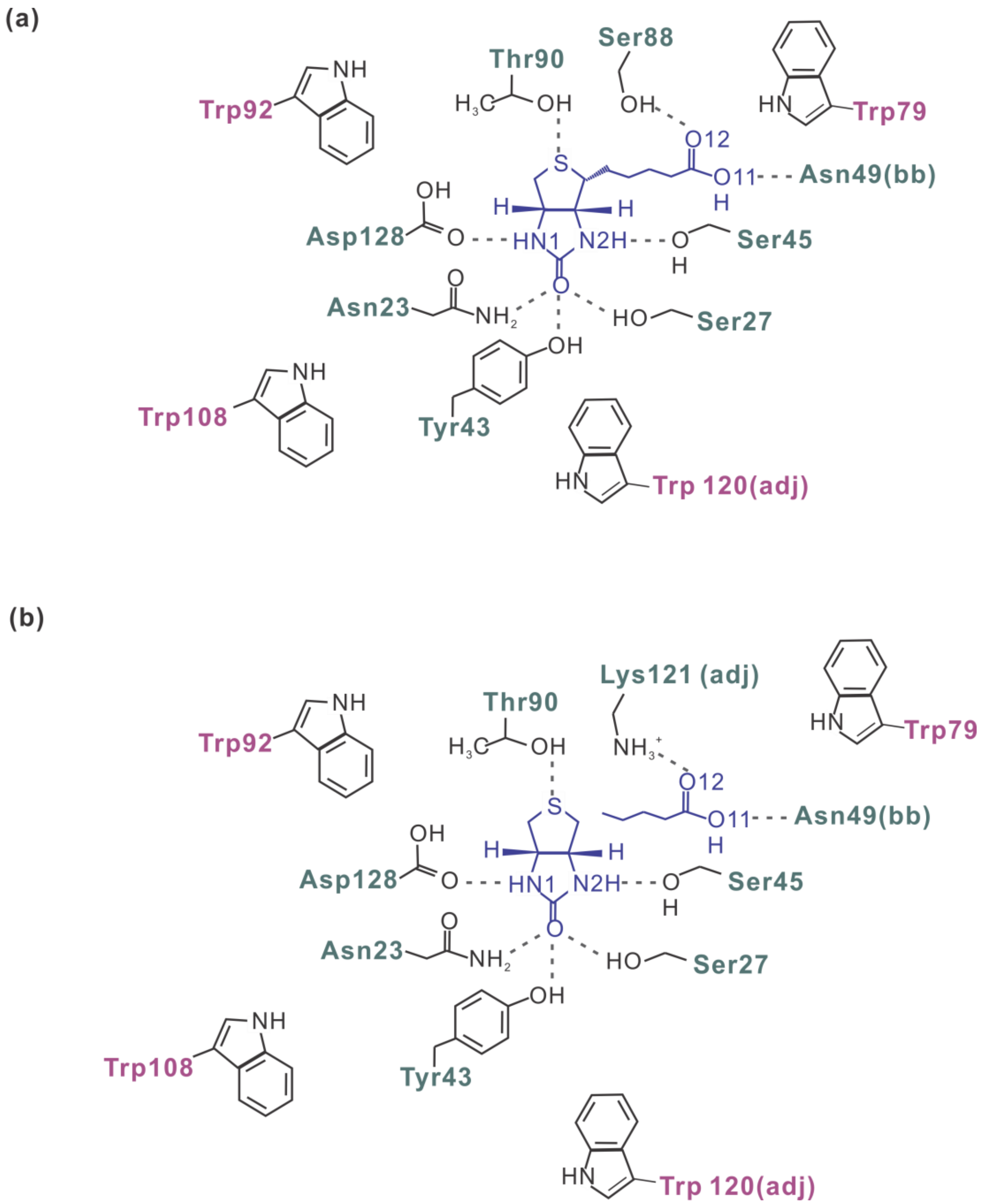


Figure 4.1. Interaction maps for (a) the WT (S_4 + 4B) complex obtained from X-ray analysis of the crystal structure, (b) the WT $(S_4 + B)^{12+}$ ion determined from MD simulations performed using fifteen different charge configurations. Here, adj represents a residue from the adjacent subunit and bb represents interaction with backbone amino.

Solvent is also expected to influence the magnitude of E_a (and k_{off}), although the nature of the solvent effects is unclear. Stayton and co-workers have suggested that the dissociative *TS* is free of substantial hydration effects.[27] However, if the loss of B from (S_4 + 4B) complex proceeds by a late *TS*, re-hydration of B and newly exposed residues in the binding cavity will serve to partially offset the energetic penalty of cleaving the protein-ligand interactions and, thereby, lower the E_a. Analysis of the crystal structures of the free and ligand-bound S_4 offers few clues as to the extent of solvent reorganization upon ligand binding.[21-22] A limited number of water molecules (two to four) can be identified in the binding cavity of some subunits of unbound WT S_4, but no water molecules are evident in the bound form.[21-22] However, the limited number or complete absence of crystallographic waters does not necessarily provide an accurate assessment of the extent of hydration of the binding cavity and changes in hydration resulting from ligand binding. Indeed, based on the results of a joint computational and crystallographic study of the earliest dissociation events, it was suggested that loss of B is initiated by the intercalation of a water molecule, which enters the binding site through an access channel at the bottom of the cavity, into the H-bond between B and Asp128.[37] The results of MD simulations reported by Houk and coworkers indicate that, for free streptavidin in an open conformation, seven water molecules reside within the binding cavity.[31] Similar results were reported by Kovalenko and coworkers.[33] Moreover, a single immobilized water molecule, acting as a water bridge between the sulphur and one of the NH groups of B, was identified in the binding cavity of the streptavidin-biotin complex.[33] Together, these computational results suggest that the loss of B is accompanied by the re-hydration of the binding cavity by six water molecules.

With the goal of more fully elucidating the influence of solvent on the streptavidin-biotin interaction and the mechanism and kinetics of ligand escape, we have, in the present study, interrogated the structure and kinetic stability of protonated ions of the ($S_4 + 4B$) complex, composed of a truncated form of WT streptavidin, in the gas phase. Ω of protonated gaseous ions of the free and ligand-bound streptavidin, i.e., WT S_4^{n+} and WT $(S_4 + 4B)^{n+}$ at n = 12 – 16, were determined from IMS measurements and the values compared to the Ω values calculated for the crystal structure reported for WT S_4 and the WT (S_4 + 4B) complex to establish whether structural changes accompany the transfer of the complexes to the gas phase by ESI. Thermal rate constants for the loss of B from the gaseous WT $(S_4 + 4B)^{n+}$ ions, at n = 12 – 16, were measured using the BIRD technique and the Arrhenius parameters established for the loss of neutral and protonated B. The E_a value for the loss of neutral B from WT $(S_4 + 4B)^{13+}$ ion was compared with those measured for $(S_4 + 4B)^{13+}$ ions composed of five binding site mutants (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala and Tyr43Ala) to establish whether specific intermolecular interactions are preserved in the gas phase. To aid in the interpretation of the gas phase data, MD simulations were carried out on desolvated WT $(S_4 + 4B)^{12+}$ ions with different charge configurations and the streptavidin-biotin interactions identified. Finally, the E_a values for the loss of B from the protonated WT $(S_4 + 4B)^{n+}$ ions were compared with those measured for the loss of B from the (S_4 + 4B) complex in aqueous solution in order to assess the possible influence of solvent reorganization on the dissociation kinetics.

4.2 Materials and Methods

4.2.1 Sample Preparation

The plasmid for natural core streptavidin (containing residues 13-139 of WT streptavidin, MW 13 271 Da) was a gift from Prof. P. Stayton (University of Washington). Notably, the kinetic data for the dissociation of the (S_4 + 4B) complex in aqueous solution were measured using natural core streptavidin.[25,27] A QuikChange II site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used to generate the single-site mutants (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala and Tyr43Ala). The PCR step was performed in a thermocycler MJ Mini PTC-1148 (Bio-Rad). The plasmid pET23a vector harbouring the streptavidin gene was used as a template. The PCR reaction mixture was treated with 1 μL of DpnI (10 units μL^{-1}) for 60 min at 37 °C to digest template-methylated non-mutated plasmid DNA. The DpnI-treated plasmids were transformed into *E. coli* XL10-Gold cells. Selected transformants were screened for the presence of the recombinant plasmid by restriction mapping and DNA sequence analysis. Positive mutants were transformed into *E. coli* BL21 (DE3) cells for expression. The WT and mutant streptavidin proteins were expressed in *E. coli* and purified using procedures described elsewhere.[38] Solutions of purified S_4 were exchanged directly into 100 mM aqueous ammonium acetate buffer using an Amicon microconcentrator with a MW cut-off of 10 kDa and lyophilized. Stock solutions of S_4 (100 μM) were prepared by dissolving a known amount of lyophilized streptavidin into 100 mM ammonium acetate and stored at -20 ℃ until needed.

Bovine β-lactoglobulin (Lg, MW 18 363 Da), human plasma transthyretin (TTR, MW 13 922 Da), egg white avidin (Avidin, MW 15 636 Da), and *Canavalia ensiformis* concanavalin A (ConA, MW 25 741 Da), and biotin (B, MW 244.3 Da) were purchased from Sigma-Aldrich Canada (Oakville, Canada). Stock solutions of Lg, TTR, Avidin and ConA (100 μM) were prepared by exchanging each protein into 100 mM ammonium acetate using Vivaspin 500 centrifugal concentrators (Sartorius Steddin Biotech) with a MW cut-off of 10 kDa. The stock solution of B (800 μM) was prepared by dissolving B in Milli-Q water. All stock solutions were stored at -20 ℃ until needed.

For the IMS measurements, the ESI solutions of the calibrant proteins (Lg, TTR, Avidin, and ConA) were prepared at a protein concentration of 10 μM in a 10 mM ammonium acetate buffer. The ESI solution of WT S_4 was prepared at concentration of 10 μM, with 4 mM imdazole and 10 mM ammonium acetate, while the ESI solution of WT (S_4 + 4B) also contained 44 μM B. For the gas-phase kinetic measurements, the ESI solutions containing S_4 (10-15 μM) and B (44-66 μM) in 10 mM ammonium acetate (pH 6.8) were prepared from aqueous stock solutions. Where a reduction of charge state for the gaseous $(S_4 + 4B)^{n+}$ was desired, imidazole (2-8 mM) was added to the solution.[39] To shift the $(S_4 + 4B)^{n+}$ ions to higher charge states, ESI was performed using a custom built device that directs a high velocity flow of air at the end of the ESI tip.[40] The choice of using protonated ions, instead of the deprotonated ions, for the present study reflects the ease with which adduct free $(S_4 + 4B)^{n+}$ ions can be produced by ESI. In negative ion mode, the deprotonated $(S_4 + 4B)^{n-}$ ions are produced with extensive acetate adducts, which complicates the interpretation of the kinetic data.

4.2.2 Mass Spectrometry

The gas-phase kinetic measurements were performed using an ApexII 9.4T FTICR mass spectrometer (Bruker, Billerica, MA). NanoESI was performed on both instruments using borosilicate tubes (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-97 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in positive ion mode was established by applying a voltage of 1.0 – 1.6 kV to a platinum wire inserted inside the glass tip. The solution flow rate was typically ~20 nL min^{-1}. Details of the instrumental and experimental conditions used for the BIRD measurements can be found elsewhere.[6-7]

The IMS measurements were performed using a Synapt G2 nanoESI-quadrupole-IMS-TOF mass spectrometer (Waters, UK). A cone voltage of 40 V was used and the source block temperature was maintained at 70 ℃. Other important voltages for ion transmission, that is the injection voltages into the trap, ion mobility, and transfer ion guides, were maintained at 10 V, 26 V, and 5 V, respectively. Argon was used in the trap and transfer ion guides at a pressure of 2.44×10^{-2} mbar and 3.36×10^{-2} mbar, respectively. The helium chamber preceding the T-wave IMS separation device was maintained at a pressure of 8.90 mbar. All IMS measurements were carried out using N_2 as the mobility gas at a pressure of 2.35 mbar. A linear IMS T-wave with a fixed wave height of 11 V and velocity of 300 m s^{-1} was used. Data acquisition and processing were carried out using Masslynx (v4.1). Mass spectra were calibrated using a sodium cesium iodide calibration file.

4.2.3 Computational Methods

MD simulations were performed using the Amber 11 program suite[41] (Accelrys, San Diego, CA). As there is no available crystal structure of the WT streptavidin-biotin tetramer, the initial geometry of the WT (S_4+4B) complex was obtained by applying a crystallographic symmetry operator on the crystal structure of WT streptavidin-biotin dimer (PDB ID: 3RY2).[22] Each of the tetrameric chains (containing residues 14-134 for the A and C chains, residues 15-136 for B and D chains) was extended to have the same length as the truncated form (containing residues 13-139) of the WT streptavidin used experimentally. This was done by aligning each tetrameric chain against the C-chain of the crystal structure for streptavidin mutant Ser27Ala (PDB ID: 1N9M)[42] using the DALI server,[43] and grafting the extended residues onto each chain of the tetramer. Currently with Amber, atomic charges and atom type parameters are available only for the charged forms of the Arg, C-terminal Ser (Cser) and N-terminal Ala residues (NAla). Consequently, it was necessary to develop charges and parameters for the neutral form of Arg, CSer and NAla. The charge parameters of the neutral form of Arg, CSer and NAla were parameterized as dipeptides (ACE-Arg-NME, ACE-Gly-CSer, and NAla-Gly-NME, respectively) using the RESP ESP charge Derive server (RED Server)[44-45] and Gaussian C.01[46], enforcing net neutrality across the targeting residues. Ions of the WT (S_4+4B) complex at the

+12 charge state were chosen for investigation. As described in more detail below, fifteen different charge distributions were considered. Topology and coordinate files for the simulations of each charge distributions were created using the Antechamber module of the AmberTools (version 11).[47] The MD simulations were performed using the Amber 03 forcefield[48] for streptavidin and a general Amber force field (GAFF)[49] for B. Simulations were performed using a 2 fs timestep with bonds to hydrogens constrained using SHAKE. The NVT ensemble was used with an Anderson thermostat (300 K, collision frequency 1 ps^{-1}) and no non-bonded cut-off (full non-bonded interactions). The system was minimized using 500 steps of steepest descent and 500 steps of conjugate gradient minimization. Following minimization, 10 ns of dynamics were needed to fully equilibrate the system, as judged by C_α root mean square deviation (RMSD). Following these 10 ns of equilibration, a further 10 ns of production were performed and used in the analysis.

4.3 Data analysis

4.3.1 Collision cross sections

Determination of Ω from drift time measurements was carried out using protocols described elsewhere.[51-52] Briefly, proteins (Lg, TTR, Avidin, and ConA) with known Ω (in He) were analysed under the same experimental conditions to establish a correlation between the measured drift times (t_D) and Ω. The t_D values were corrected for both mass-dependent and mass-independent flight times using eq 4.1:[51]

$$t_D' = t_D - c\sqrt{m/z}/1000 - 10 \cdot l_{\mathrm{transfer}} / \mathrm{WV}_{\mathrm{transfer}} \tag{4.1}$$

where $t_D^{'}$ is the corrected drift time (in ms), t_D is the measured drift time (in ms), the constant c is known as the enhanced duty cycle delay coefficient and is found within the Synapt G2 software,[51] $l_{transfer}$ is the length of the transfer T-wave region (in cm) and $WV_{transfer}$ is the transfer wave velocity (in m s^{-1}). The reported Ω (in He) for the calibrant protein ions[50] were corrected for charge and reduced mass (μ) by eq 4.2:

$$\Omega^{'} = \Omega\sqrt{\mu} / z \qquad (4.2)$$

where $\Omega^{'}$ is the corrected collision cross section in He. The $t_D^{'}$ and $\Omega^{'}$ values of each calibrant ion are summarized in Table 4.1. As shown in Figure 4.2a, the plot of $\ln(\Omega^{'})$ versus $\ln(t_D^{'})$ is linear. The slope of the curve (0.5712) corresponds to the exponential factor, X, which is sensitive to many experimental variables, including the height and velocity of the voltage 'waves' used to propel ions through the IMS separation region.[53] Using the empirically-determined value of X, the final, corrected drift times ($t_D^{''}$) were established using eq 4.3:

$$t_D^{''} = (t_D^{'})^X (z / \sqrt{\mu}) \qquad (4.3)$$

Notably, the plot of literature Ω values versus $t_D^{''}$ is linear with an R^2 value of 0.997 (Figure 4.2b). The Ω values of the protonated WT S_4^{n+} and WT $(S_4 + 4B)^{n+}$ ions were determined from the calibration curve and the corresponding $t_D^{''}$ values.

Table 4.1 Charge states, m/z values, Ω values (from the literature[50]) and drift times (t_D) are listed for the calibrant ions: β-lagtoglobulin A, transthyretin, avidin and concanavalin A. Also shown are the corrected drift times (t_D' and t_D") and corrected Ω values (Ω').

charge state	**m/z**	**literature Ω (nm^2)**	**drift time (t_D, ms)**	**corrected drift time (t_D', ms)**	**corrected Ω' (nm^2)**	**final corrected drift time (t_D'', ms)**
β-lagtoglobulin A (monomer)						
7	2624.28	16.6	14.77	13.98	12.54	5.97
8	2296.35	16.9	12.23	11.45	11.17	6.09
9	2041.26	17.8	10.99	10.18	10.46	6.41
β-lagtoglobulin A (dimer)						
11	3339.79	28.5	16.88	16.09	13.70	10.16
12	3061.51	29.0	14.99	14.20	12.78	10.33
13	2826.10	29.6	13.85	13.02	12.04	10.65
transthyretin (tetramer)						
14	3974.26	34.1	15.46	14.65	12.89	12.26
15	3711.85	34.0	13.66	12.86	11.99	12.19
16	3487.17	33.8	12.37	11.57	11.18	12.25
avidin tetramer						
15	4261.31	36.4	15.46	14.65	12.84	13.14
16	3658.92	36.4	13.80	13.00	12.04	13.09
17	3757.56	36.4	12.33	11.52	11.33	12.98
18	3550.56	36.4	11.18	10.38	10.70	12.95
concanavalin A tetramer						
20	5150.52	55.5	19.18	18.36	14.68	19.93
21	4904.59	55.5	17.53	16.71	13.98	19.83
22	4679.78	54.8	16.28	15.47	13.18	19.87

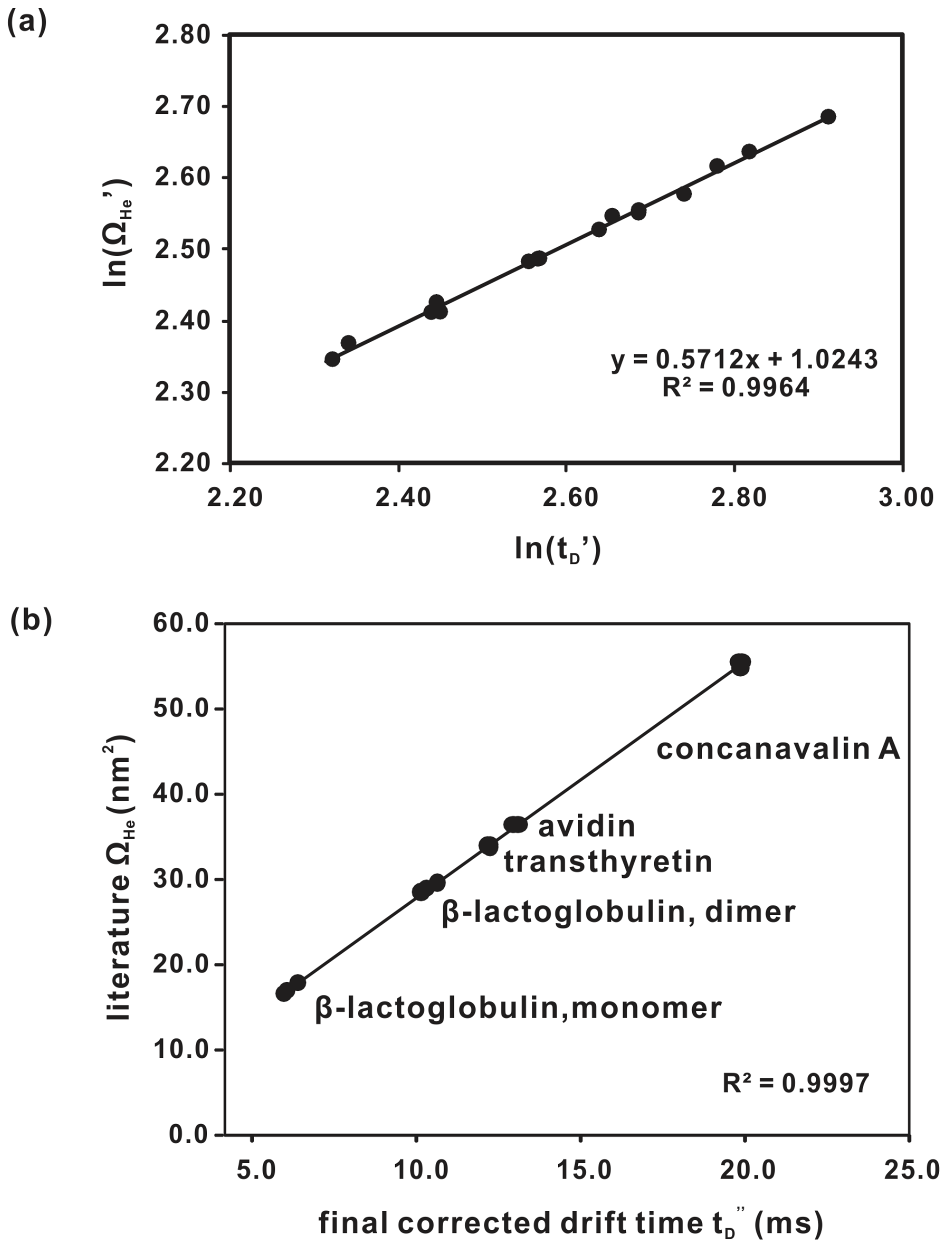


Figure 4.2. (a) Plot of ln(Ω_{He}') versus ln(t_D') for the calibrants: β-lactoglobulin A, transthyretin, avidin and concanavalin A. An exponential factor (*X*) of 0.5712 was determined from the slope of the plot. (b) Calibration curve, based on the calibrants β-lactoglobulin A, transthyretin, avidin and concanavalin A, displayed as a linear plot of literature Ω_{He} values and final corrected drift times (t_D'').

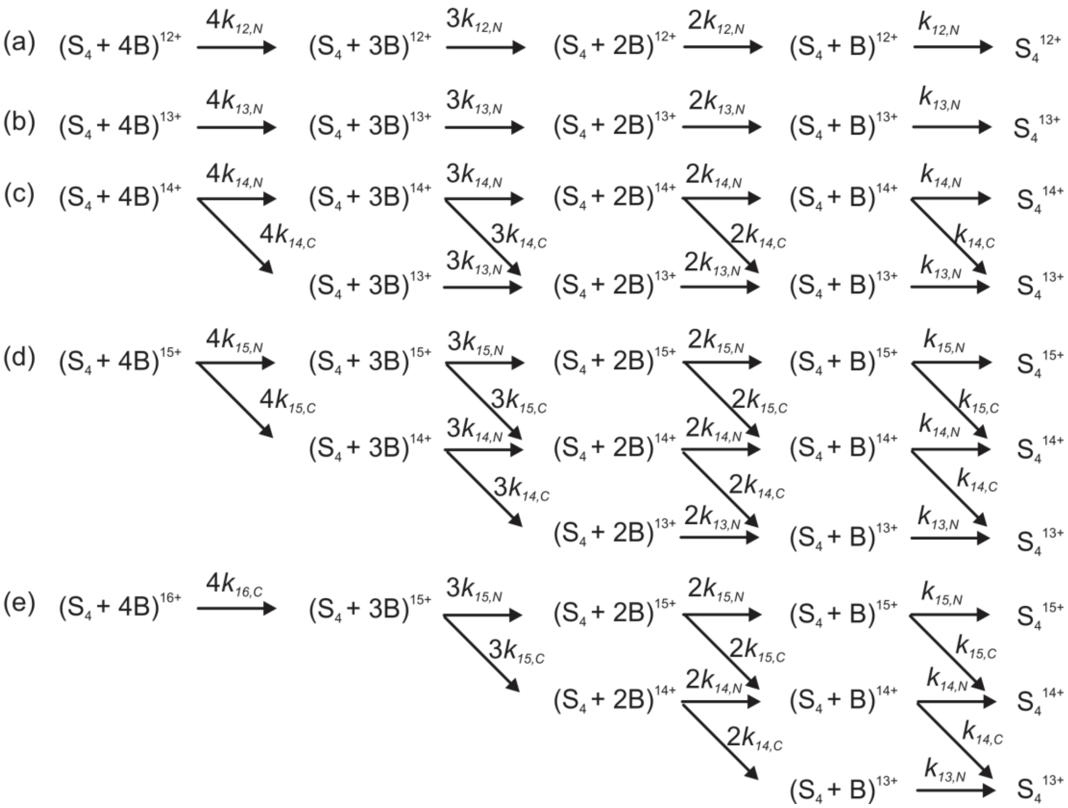
The Ω value for WT S_4 and (S_4 + 4B) in aqueous solution were estimated from available crystal structures using a modified version of Mobcal,[54] which increases the number of iterative calculations performed per input structure and is capable of accepting input coordinate files that contain up to 100,000 atoms, and the TM. The input file for S_4 was prepared using the crystal structure of WT streptavidin tetramer (PDB ID: 1SWB);[21] the input file for (S_4 + 4B) was prepared as described in the Experimental section using the crystal structure of WT streptavidin-biotin dimer (PDB ID: 3RY2).[22] The coordinates for the full tetramer were obtained by applying a crystallographic symmetry operator. All crystallographic water molecules were removed manually and hydrogens were added to the crystal structure using Leap in the Amber.[41]

4.3.2 Dissociation kinetics

The apparent rate constant ($k_{n,app}$) for the loss of B from a given protonated $(\mathrm{S_4} + 4\mathrm{B})^{n+}$ ion, at a given temperature, was determined from a linear least squares fit of the plot of the natural logarithm of the normalized abundance (A_R) of the reactant ion versus reaction time. The value of A_R was calculated using eq 4.4:

$$A_\mathrm{R} = Ab_{(\mathrm{S_4}+4\mathrm{B})^{\mathrm{n+}}} / \Sigma Ab_{(\mathrm{S_4}+i\mathrm{B})} \tag{4.4}$$

where $Ab_{(\mathrm{S_4}+4\mathrm{B})^{\mathrm{n+}}}$ is the measured abundance of the $(\mathrm{S_4} + 4\mathrm{B})^{n+}$ reactant ion and $\Sigma Ab_{(\mathrm{S_4}+i\mathrm{B})}$ is the total abundance of the reactant and product ions i.e., all $(\mathrm{S_4} + i\mathrm{B})^{n+}$ ions with $0 \leq i \leq 4$ and, if any, all $(\mathrm{S_4} + i\mathrm{B})^{(n-1)+}$ ions with $0 \leq i \leq 3$.



Scheme 4.1. Dissociation pathways observed for the gas phase ions (a) $(S_4 + 4B)^{12+}$, (b) $(S_4 + 4B)^{13+}$, (c) $(S_4 + 4B)^{14+}$, (d) $(S_4 + 4B)^{15+}$ and (e) $(S_4 + 4B)^{16+}$.

As outlined in Scheme 4.1, dissociation of the higher charge state $(S_4 + iB)^{n+}$ ions proceeds by the loss of both neutral and protonated B through parallel pathways, eq 4.5a,b:

$$(S_4 + iB)^{n+} \rightarrow (S_4 + (i\text{-}1)B)^{n+} + B \qquad (4.5a)$$

$$\rightarrow (S_4 + (i\text{-}1)B)^{(n-1)+} + B^+ \qquad (4.5b)$$

For a given protonated $(S_4 + 4B)^{n+}$ ion, $k_{n,app}$ reflects the intrinsic rate constants for the individual neutral and charged B-loss pathways, $k_{n,N}$ and $k_{n,C}$, respectively, and the number of bound ligands, eq 4.6:

$$k_{n,app} = 4(k_{n,N} + k_{n,C}) \qquad (4.6)$$

In principle, $k_{n,N}$ and $k_{n,C}$ can be determined from the relative abundance of the unbound S_4^{n+} ion measured once the reaction has gone to completion ($Ab_{S_4^{n+}}{}^{\infty}$) and $k_{n,app}$, eq 4.7

$$Ab_{S_4^{n+}}{}^{\infty} = \left(\frac{k_{n,N}}{k_{n,N} + k_{n,C}}\right)^4 = \left(\frac{4k_{n,N}}{k_{n,app}}\right)^4 \qquad (4.7)$$

However, due to ion losses associated with long trapping times, it is difficult to reliably measure $Ab_{S_4^{n+}}{}^{\infty}$, particularly at low reaction temperatures. Therefore, an alternative approach, based on the measured abundances of the reactant and product ions, was used. Using the measured abundances and the corresponding branching ratio, $k_{n,N}/(k_{n,N} + k_{n,C})$, the normalized abundance of S_4^{n+} expected at the completion of the reaction ($Ab_{S_4^{n+}}{}^{\infty}$) can be calculated at any point during the reaction, eq 4.8:

$$Ab_{S_4^{n+}}{}^{\infty} = \sum_{i=0}^{4}\left(\frac{k_{n,N}}{k_{n,N} + k_{n,C}}\right)^i Ab_{(S_4+iB)^{n+}} = \sum_{i=0}^{4}\left(\frac{4k_{n,N}}{k_{n,app}}\right)^i Ab_{(S_4+iB)^{n+}} \qquad (4.8)$$

where $Ab_{(S_4+iB)^{n+}}$ is the normalized abundance of the $(S_4 + iB)^{n+}$ ion.

Combining eqs 4.7 and 4.8 gives eq 4.9:

$$\left(\frac{4k_{n,N}}{k_{n,app}}\right)^4 = \sum_{i=0}^{4}\left(\frac{4k_{n,N}}{k_{n,app}}\right)^i Ab_{(S_4+iB)^{n+}} \qquad (4.9)$$

Using the values of $Ab_{(S_4+iB)^{n+}}$ and $k_{n,app}$ measured at different reaction times, the corresponding values of $k_{n,N}$ were determined by solving eq 4.9 using Maple (Maplesoft, Waterloo, Canada). At each temperature, $k_{n,N,ave}$ and $k_{n,C,ave}$, the average values of $k_{n,N}$ and $k_{n,C}$, were determined. For the lower charge state $(S_4 + 4B)^{n+}$ ions, with n = 12 and 13, where B is lost only at its neutral form, $k_{n,N}$ is simply equal to $1/4k_{n,app}$, while for the $(S_4 + 4B)^{16+}$ ion, which dissociates exclusively (or nearly so) by the loss of protonated B, $k_{n,C}$ is equal to $1/4k_{n,app}$.

4.3.3 Computational results.

Trajectory analysis, performed using the Visual Molecular Dynamics package,[55-56] was carried out to establish the C_α RMSD for streptavidin, the angles and distances associated with the H-bonds between B and streptavidin, and the distances associated with the intermolecular van der Waals interactions between B and four tryptophan residues, Trp 79, Trp92, Trp108 and (adjacent subunit) Trp120. In order to determine the number of H-bonds, all potential H-bonding partners between B and streptavidin were scanned at each frame, with the criterion for a H-bond being a heavy atom distance ≤ 4 Å, and a donor-hydrogen-acceptor angle ≥120°. The total number of H-bonds for each configuration was then averaged across all frames and the four subunits. Additionally, the occupancy, i.e., the fraction (*f*) of the simulation steps for which the H-bond criteria are satisfied, was evaluated. All potential van der Waals interaction atom pairs between B and four tryptophan residues (Trp 79, Trp92, Trp108 and Trp120 from adjacent subunit) were scanned at each frame, with the criterion for the presence of van der Waals interactions being the distance between atom centers for each pair of atoms between relevant tryptophan residue and B is less than or equal to the sum of the van der Walls radii for those particular atoms (van der Waals radii based on the Amber parameter set were used for the simulations). The fraction (*f*) of the simulation steps for which the van der Waals interaction criteria are satisfied was also evaluated.

The Ω for the averaged coordinates of the $(S_4 + 4B)^{12+}$ ion, at each charge configuration, was calculated using Mobcal[54] and TM. The averaged coordinates were generated by aligning the backbone atoms across all frames, following which the atomic positions were averaged. As side chain positions can often become distorted due to structural averaging, energy minimization with the protein backbone atoms fixed was performed using NAMD[56-57] in order to remove any structural artifacts introduced during the averaging step.

4.4 Results and Discussion

4.4.1 Structure and kinetic stability of $(S_4 + 4B)^{n+}$ ions in the gas phase

Protonated gas phase ions of WT $(S_4 + 4B)$ complex, e.g. $(S_4 + 4B)^{n+}$ at n = 12 - 16, were produced from aqueous ammonium acetate solutions (10 mM, pH 6.8) containing WT S_4 (10 µM), B (44 µM) and imidazole (4 mM). A representative ESI mass spectrum is shown in Figure 4.3; also shown are the corresponding arrival time distributions measured for the WT $(S_4 + 4B)^{n+}$ ions. Mobility measurements were also performed on the protonated WT S_4^{n+} ions at n = 12 - 16, which were produced from aqueous ammonium acetate solutions (10 mM, pH 6.8) containing WT S_4 (10 µM) and imidazole (4 mM), Figure 4.4. The Ω values for the S_4^{n+} and $(S_4 + 4B)^{n+}$ ions were determined from their measured drift times and the calibration curve, as described in the Data Analysis section, and are listed in Table 4.2, together with the Ω values estimated for the crystal structures of WT S_4 and WT $(S_4 + 4B)$.[21,22]

Interestingly, the Ω values measured for the S_4^{n+} and $(S_4 + 4B)^{n+}$ ions are independent of charge state, with average values of 34.4 ± 0.1 nm^2 and 35.0 ± 0.1 nm^2, respectively. These results suggest that charge-induced structural changes, at least over this range of charge states, are minimal. The average Ω value for WT $(S_4 + 4B)^{n+}$ is 35.0 ± 0.1 nm^2, which agrees within 10% with the Ω value (38.4 ± 0.4 nm^2) estimated for the $(S_4 + 4B)$ crystal structure. Similar agreement is found between the average Ω for WT S_4^{n+} (34.4 ± 0.1 nm^2) and the estimated Ω value (38.8 ± 0.4 nm^2). Given that the uncertainty in the Ω values derived from the IMS measurements is 5 - 8%[51] and the uncertainty in the calculated Ω values is 3 - 4%, the differences between the Ω values for the $(S_4 + 4B)^{n+}$ and S_4^{n+} ions and those for the corresponding crystal structures are not significant. Therefore, the IMS results would seem to suggest that the transfer of S_4 and $(S_4 + 4B)$ from neutral aqueous solution to the gas phase by ESI is not accompanied by large structural changes, or at least none that can be identified from the IMS measurements.

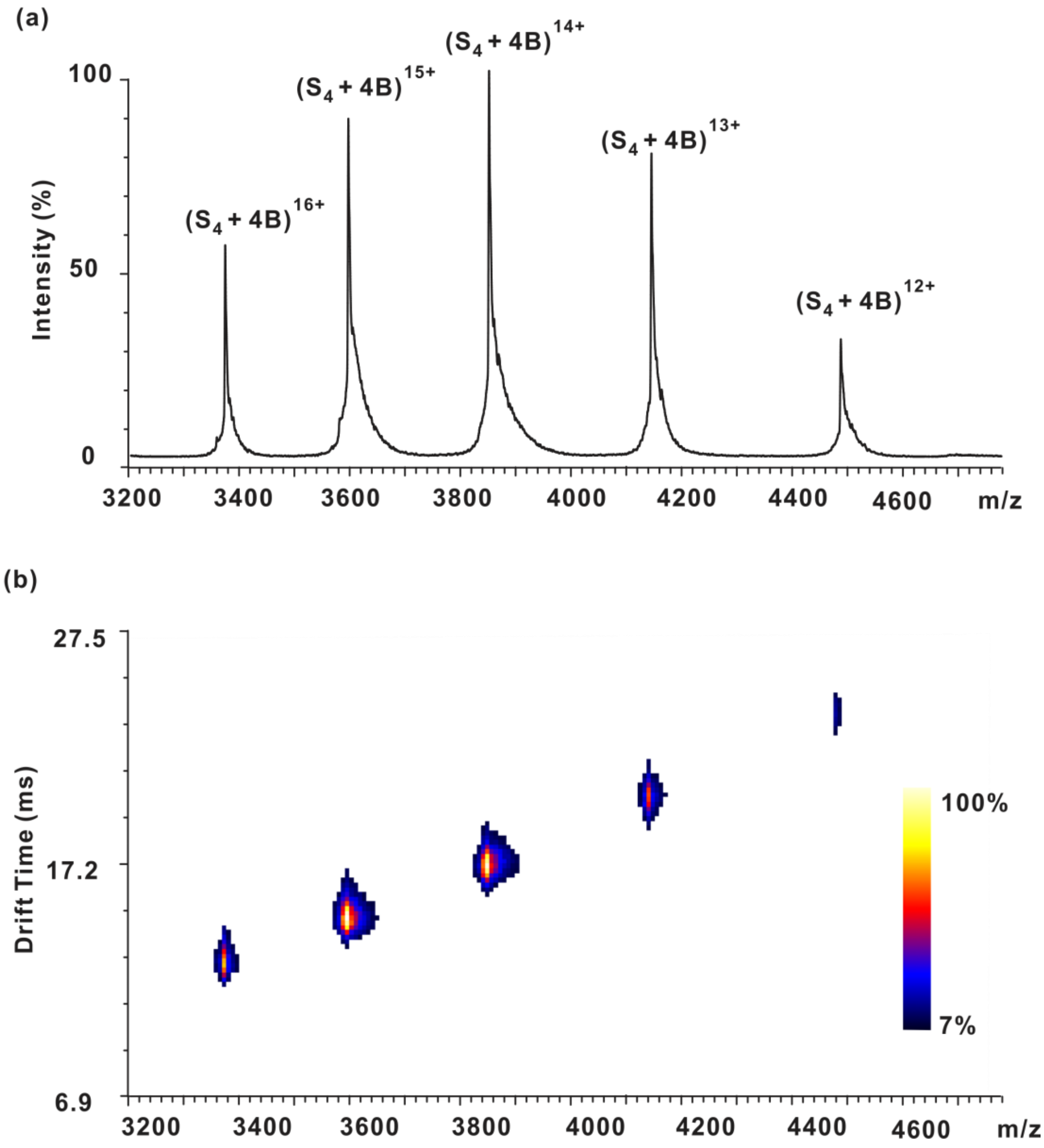


Figure 4.3. (a) ESI mass spectrum acquired for a neutral aqueous ammonium acetate (10 mM) solution of WT S_4 (10 μM), B (44 μM) and imidazole (4 mM). (b) Ion mobility heat map plotting arrival time versus *m*/*z*. The normalized ion intensities (from 7% to 100%) are represented using the indicated colour scale.

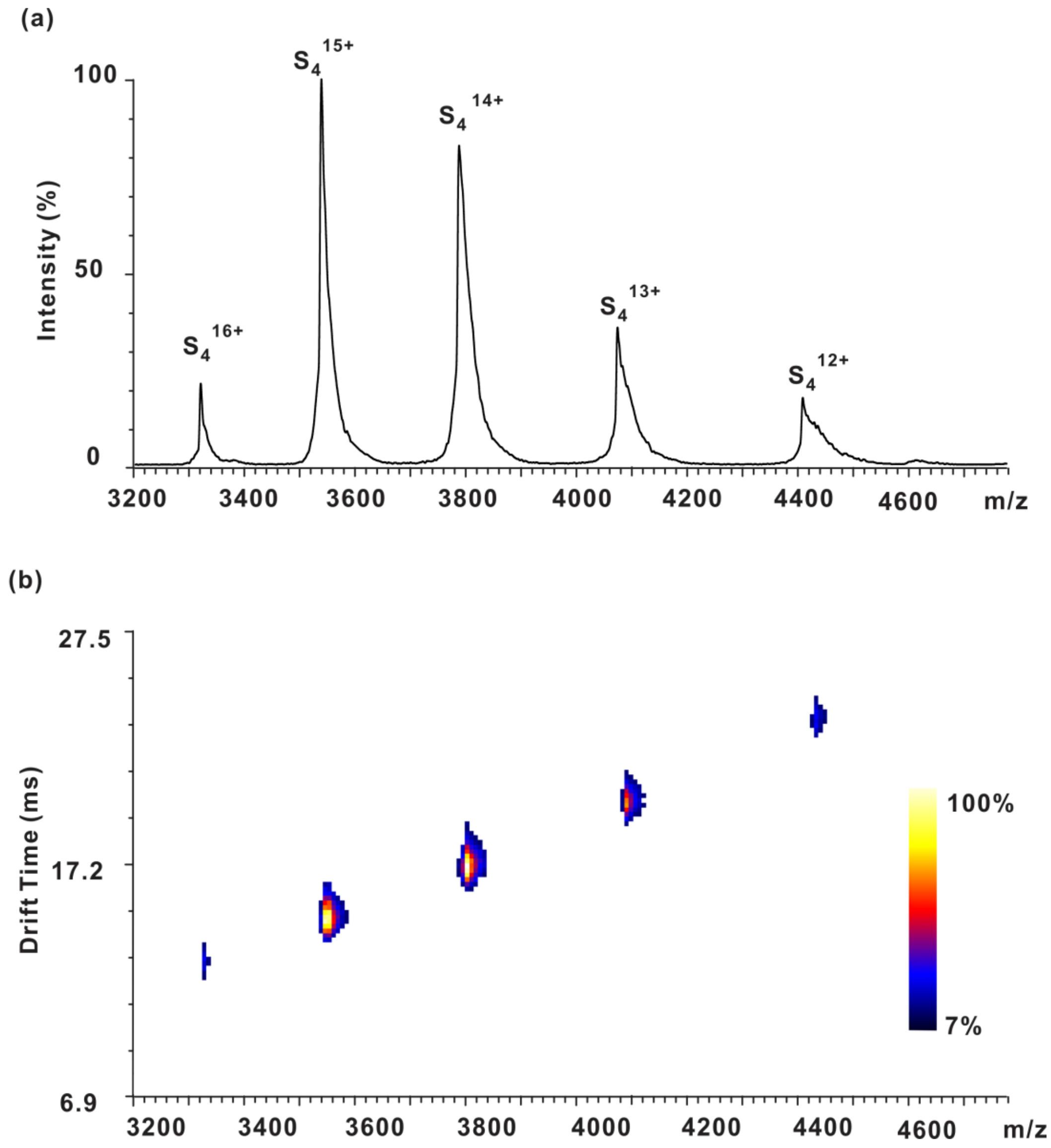


Figure 4.4. (a) ESI mass spectrum acquired for a neutral aqueous ammonium acetate (10 mM) solution of WT S_4 (10 μM) and imidazole (4 mM). (b) Ion mobility heat map plotting arrival time versus *m*/*z*. The normalized ion intensities (from 7% to 100%) are represented using the indicated colour scale.

Table 4.2. Comparison of experimentally estimated collision cross sections (Ω) for the gaseous S_4^{n+} and $(S_4 + 4B)^{n+}$ ions, where n = 12 – 16, with Ω values calculated for crystal structures reported for S_4 and the $(S_4 + 4B)$ complex.

Charge state (n)	S_4^{n+} Ω (nm^2)	$(S_4 + 4B)^{n+}$ Ω (nm^2)
12	34.54 ± 0.03 [a]	35.08 ± 0.15 [a]
13	34.42 ± 0.10 [a]	34.99 ± 0.02 [a]
14	34.29 ± 0.04 [a]	34.88 ± 0.03 [a]
15	34.27 ± 0.13 [a]	34.78 ± 0.04 [a]
16	34.45 ± 0.09 [a]	35.03 ± 0.13 [a]
Average	34.39 ± 0.13 [b]	34.95 ± 0.14 [b]
Crystal structure	38.76 ± 0.37 [b]	38.45 ± 0.36 [c]

a. Uncertainty corresponds to the standard deviation of three replicate measurements. b. Uncertainty corresponds to the standard deviation of the Ω values determined at charge states +12 to +16. c. Ω was determined from crystal structure (protein ID: 1SWB) using Mobcal and the trajectory method (TM). c. Ω was determined from crystal structure (protein ID: 3RY2) using the TM; the reported error corresponds to one standard deviation of values from replicate (10) calculations.

Time-resolved BIRD measurements were performed on the WT $(S_4 + 4B)^{n+}$ ions, with n = 12 to 16, at temperatures ranging from 80 to 135 °C. At these temperatures, BIRD results in the sequential ejection of B, with no evidence for the loss of subunits (Figure 4.5). The absence of measurable subunit loss at these temperatures is consistent with the findings of an earlier BIRD study of the protonated S_4^{n+} ions.[58] At the lower charge states, n = 12 and 13, the loss of neutral B represents the only dissociation channel observed, while the loss of protonated B represents the major dissociation channel at n = 16. At n = 14 and 15, B was ejected from the complex in both its neutral and protonated forms. The dissociation pathways observed for all of the charge states investigated are summarized in Scheme 4.1. Shown in Figure 4.6 are kinetic plots measured for the dissociation of the $(S_4 + 4B)^{n+}$ ions at the temperatures indicated. Notably, the plots exhibit excellent linearity, which is consistent with the presence of a single dominant reactant structure at each charge state. The linear plots also indicate that the loss of neutral and protonated B, from the $(S_4 + 4B)^{n+}$ ions at n = 14 and 15, proceed through parallel pathways. Arrhenius plots were constructed from the $k_{n,app}$, $k_{n,N,ave}$ and $k_{n,C,ave}$ values, which were determined as described in the Data Analysis section. Shown in Figure 4.7 are the Arrhenius plots, based on the temperature dependence of the $k_{n,app}$ values, for the loss of B from the $(S_4 + 4B)^{n+}$ ions at n = 12 – 16. The corresponding plot for the loss of B from the $(S_4 + 4B)$ complex in aqueous solution at pH 7.4, extrapolated from kinetic data determined at temperatures ranging from 4 to 37 °C,[25] is also shown for comparison. Arrhenius plots based on the $k_{n,N,ave}$ and $k_{n,C,ave}$ values are shown in Figure 4.8.

The Arrhenius parameters (E_a, A; $E_{a,N}$, A_N and $E_{a,C}$, A_C) are listed in Table 4.3 as are the parameters for the loss of B from the (S_4 + 4B) complex in aqueous solution at pH 7.4, which were calculated from the reported activation enthalpy and entropy.[25]

Table 4.3. Arrhenius activation parameters (E_a, A) measured for the loss of B from the protonated $(S_4 + 4B)^{n+}$ ions, where n = 12 – 16, in the gas phase and the corresponding parameters measured the (S_4 + 4B) complex in aqueous solution at pH 7.4.[a]

n	**E_a**	**A**	**$E_{a,N}$ [c]**	**A_N [c]**	**$E_{a,C}$ [d]**	**A_C [d]**
	(kcal mol^{-1})	**(s^{-1})**	**(kcal mol^{-1})**	**(s^{-1})**	**(kcal mol^{-1})**	**(s^{-1})**
12	44.3 ± 0.8	$10^{23.9 \pm 0.4}$	44.3 ± 0.8	$10^{23.3 \pm 0.4}$	-	-
13	43.2 ± 1.1	$10^{23.2 \pm 0.6}$	43.2 ± 1.1	$10^{22.6 \pm 0.6}$	-	-
14	38.2 ± 0.8	$10^{20.5 \pm 0.5}$	38.5 ± 1.1	$10^{19.9 \pm 0.6}$	37.3 ± 1.0	$10^{18.9 \pm 0.6}$
15	31.2 ± 0.1	$10^{17.0 \pm 0.2}$	32.2 ± 1.0	$10^{16.0 \pm 0.6}$	30.7 ± 0.6	$10^{16.0 \pm 0.3}$
16	26.1 ± 0.4	$10^{14.9 \pm 0.3}$	-	-	26.1 ± 0.4	$10^{14.2 \pm 0.3}$
Solution	31.2 ± 0.2 [b]	$10^{17.6 \pm 0.1}$ [b]				

a. Errors correspond to standard deviations. b. E_a and A values calculated at 110 ℃ from the activation enthalpy and entropy reported in reference 25. c. The parameters $E_{a,N}$, A_N correspond to the loss of neutral B from the protonated $(S_4 + 4B)^{n+}$ ions, where n = 12 – 16. d. The parameters $E_{a,C}$, A_C correspond to the loss of protonated B from the protonated $(S_4 + 4B)^{n+}$ ions, where n = 12 – 16.

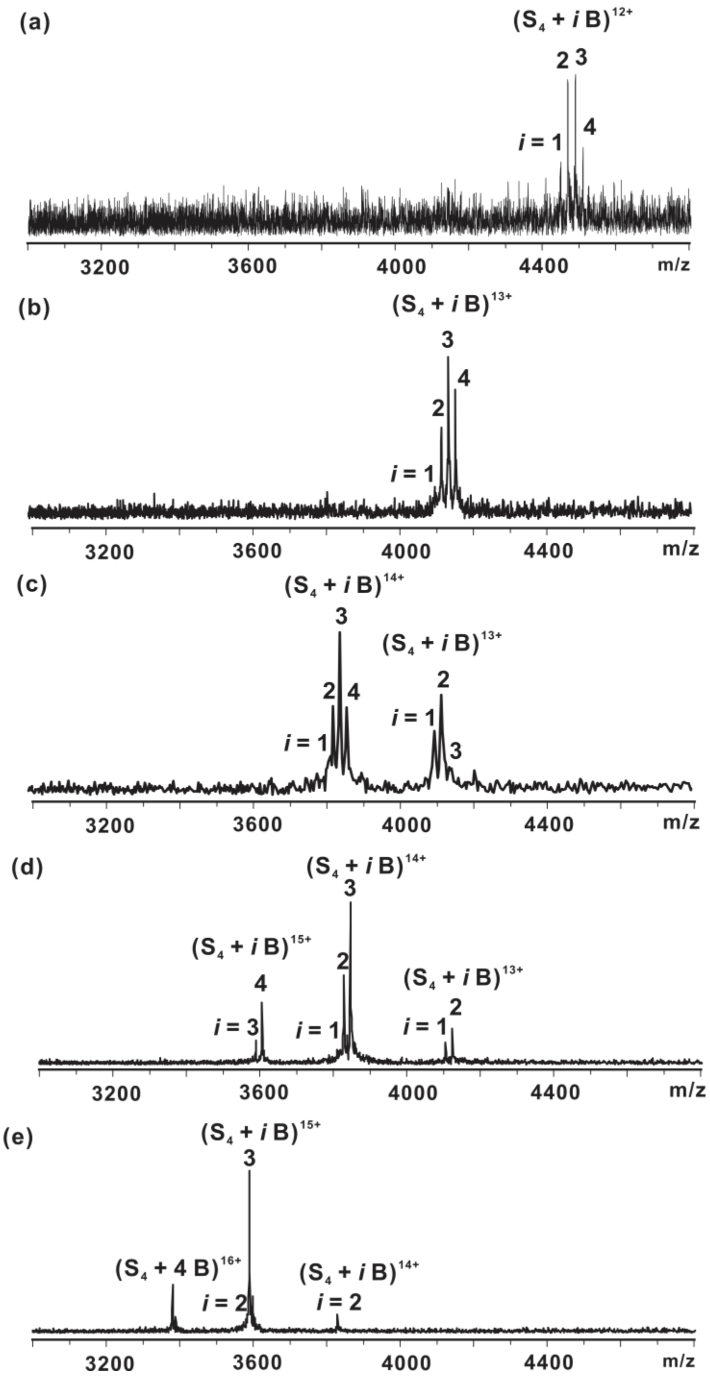


Figure 4.5. Illustrative BIRD mass spectra measured for (a) $(S_4+4B)^{12+}$ at a reaction temperature of 135 ℃ and a reaction time of 2.5 s, (b) $(S_4+4B)^{13+}$ at 125 ℃ and 6.0 s, (c) $(S_4+4B)^{14+}$ at 136 ℃ and 2.5 s, (d) $(S_4+4B)^{15+}$ at 114 ℃ and 9.0 s, and (e) $(S_4+4B)^{16+}$ 91 ℃ and 12.0 s.

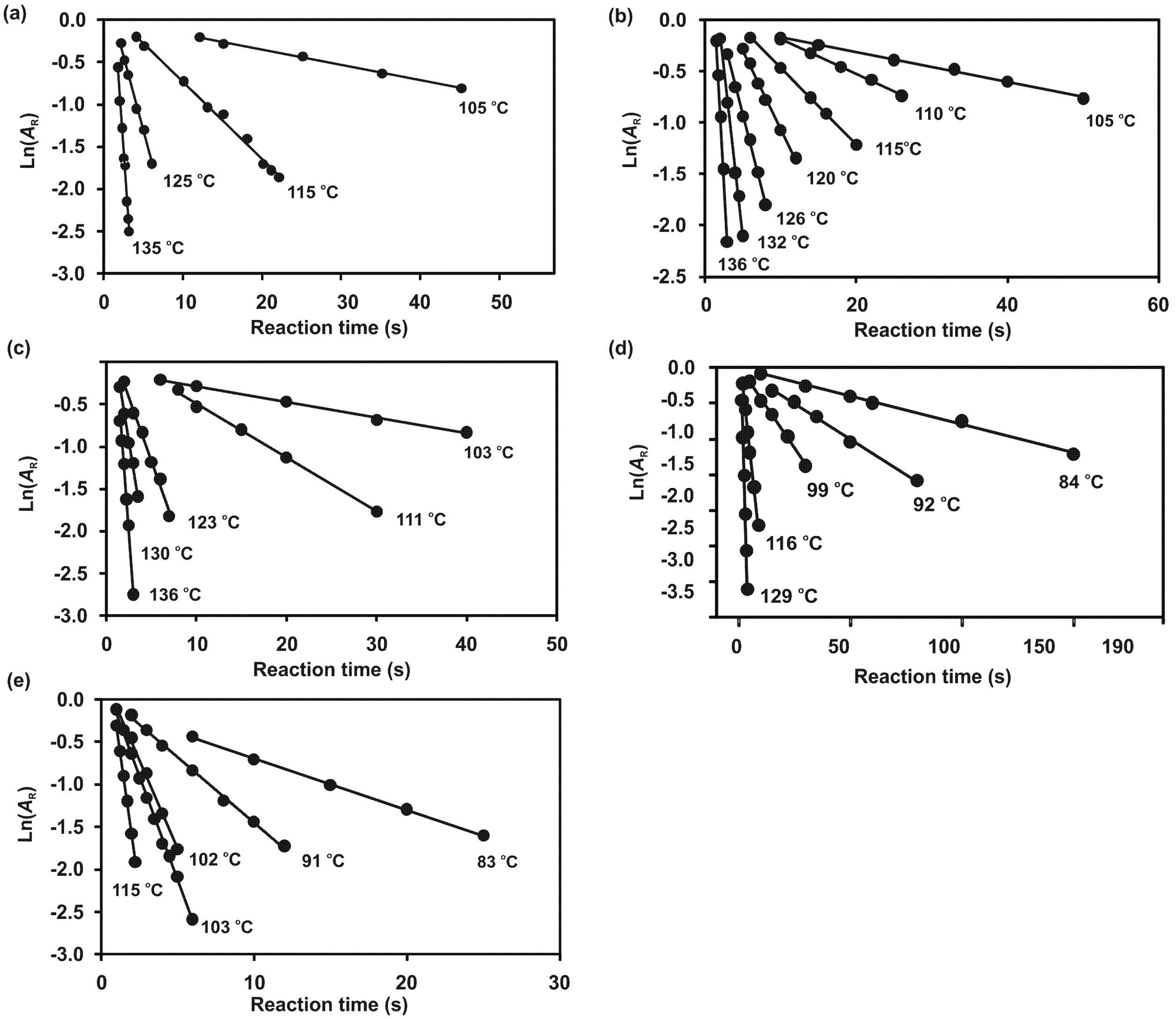


Figure 4.6. Kinetic plots of the natural logarithm of the normalized intensity (A_R) of (a) $(S_4 + 4B)^{12+}$, (b) $(S_4 + 4B)^{13+}$, (c) $(S_4 + 4B)^{14+}$, (d) $(S_4 + 4B)^{15+}$ and (e) $(S_4 + 4B)^{16+}$, versus reaction time at the temperatures indicated.

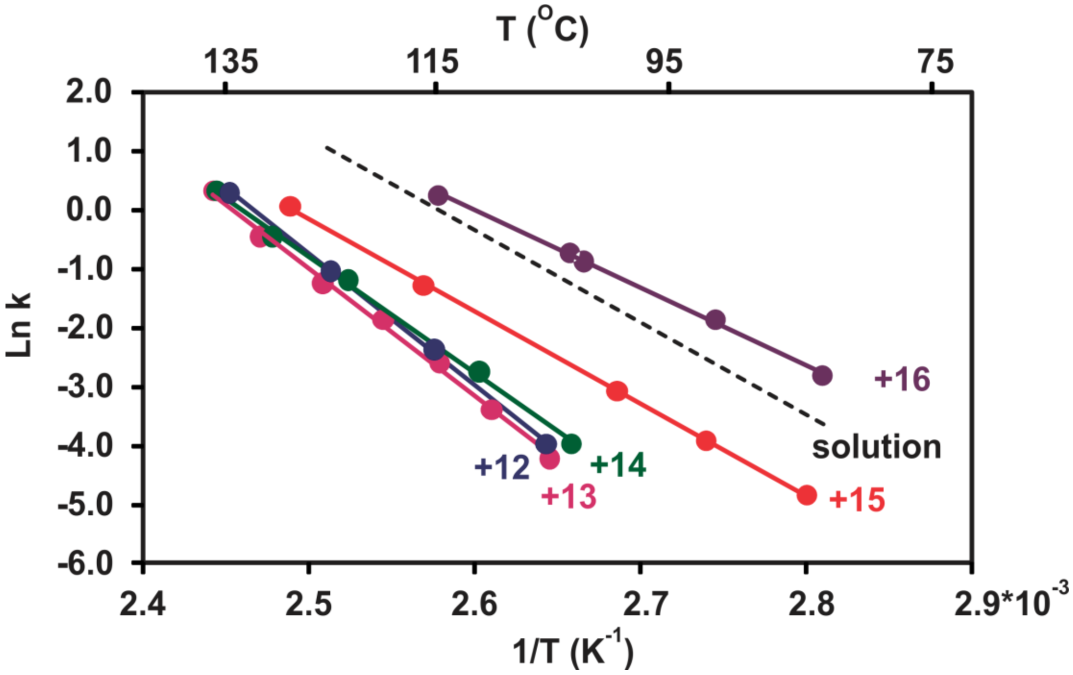


Figure 4.7. Arrhenius plots for the dissociation of the gaseous WT $(S_4 + B)^{n+}$ ions, where n = 12 – 16, and the corresponding Arrhenius plot (based on parameters calculated at 110 ℃ from the activation enthalpy and entropy values reported in reference 25) for the dissociation of the WT $(S_4 + 4B)$ complex in neutral aqueous solution at pH 7.4.

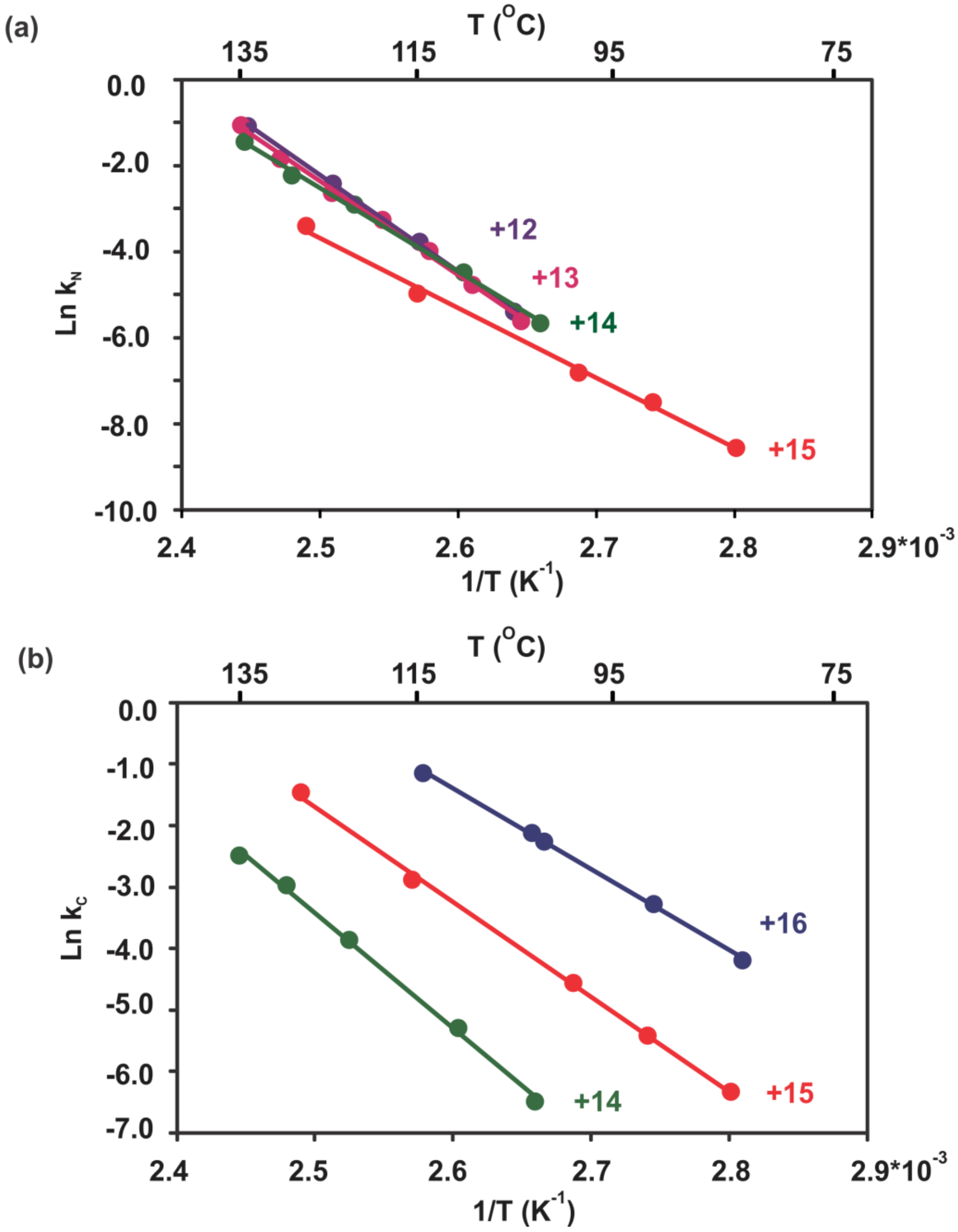


Figure 4.8. (a) Arrhenius plots for the dissociation of neutral B from the gaseous WT $(S_4 + 4B)^{n+}$ ions , where n = 12 - 15. (b) Arrhenius plots for the dissociation of protonated B from the gaseous WT $(S_4 + 4B)^{n+}$ ions, where n = 14 - 16.

Inspection of Figure 4.7 (and Table 4.3) reveals that the Arrhenius plots for the loss of B from $(S_4 + 4B)^{n+}$ ions at n = 12 and 13 are nearly indistinguishable, with E_a values of 44 and 43 kcal mol^{-1}, respectively. Taken on their own, these results suggest that B is stabilized by the same or similar intermolecular interactions at these charge states. In contrast, the kinetic stabilities of the $(S_4 + 4B)^{n+}$ ions, at n = 13 – 16, decrease with increasing charge state. The trend in kinetic stability mirrors the trend in E_a, which also decreases with increasing charge state. To a first approximation, the decrease in E_a with increasing charge state can be understood in terms of the effect of Coulombic repulsion on the pathway leading to the loss of protonated B. As discussed in detail elsewhere,[58] electrostatic repulsion between the charged products, in this case $(S_4 + 3B)^{(n-1)+}$ and B^+, will serve to reduce the energy barrier for dissociation and the magnitude of this effect is expected to increase with the charge state of the complex. Indeed, the $E_{a,C}$ values determined for the loss of protonated B from the $(S_4 + 4B)^{n+}$ ions at n = 14 – 16 decrease significantly with increasing charge state (by ~11 kcal mol^{-1} from +14 to +16). However, the $E_{a,N}$ values for the loss neutral of B from the $(S_4 + 4B)^{n+}$ ions, at n = 13 – 15, also decrease by a similar amount. These results indicate that charge-induced structural changes, which alter the intermolecular interactions or reduce their strength or number but do not lead to detectable changes in **Ω**, occur at these higher charge states.

Kinetic measurements were also performed on $(S_4 + 4B)^{13+}$ ions composed of five binding site mutants (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala and Tyr43Ala) in order to identify some of the residues that bind to B in the gas phase.

In each case, BIRD resulted in the sequential loss of neutral B, similar to what was observed for the WT $(S_4 + 4B)^{13+}$ ion. Representative kinetic plots, measured at the temperatures indicated, are shown in Figure 4.9. The Arrhenius plots are shown in Figure 4.10 and the corresponding parameters are listed in Table 4.4. Inspection of kinetic parameters reveals that replacement of Ser27 or Tyr43 with Ala results in a significant decrease, by ~5.0 kcal mol^{-1}, in E_a. According to the crystal structure, both of these residues participate in H-bonds with B (Figure 4.1). The magnitude of the reductions in E_a measured upon mutation to Ala is comparable to those observed upon replacement of H-bonding groups (e.g. hydroxyl groups) in the protonated (scFv + αGal[αAbe]αMan)$^{n+}$ ions.[6] Replacement of Trp79 or Trp120, both of which are reported to stabilize B in solution, with Phe also results in a decrease in E_a, although these changes are smaller, ~2.5 kcal mol^{-1}. Together, these results suggest that Trp79, Trp120, Ser27 and Tyr43 interact with B in the gaseous WT $(S_4 + 4B)^{13+}$ ion. In contrast, replacement of Trp108 with Phe does not produce a significant change in E_a, suggesting that either Trp108 does not interact with B in the gas phase or, if it does, the interaction is energetically similar to that resulting from the introduction of Phe at this position. Taken together, the results of this mutagenesis study suggest that, at least, some of the stabilizing intermolecular interactions in the $(S_4 + 4B)$ complex are preserved in the gaseous $(S_4 + 4B)^{13+}$ ion. This finding is consistent with those reported previously for the protonated (scFv + αGal[αAbe]αMan)$^{n+}$ ions and deprotonated (Lg + FA)$^{7-}$ ions.[6-9] Although, for both of these complexes, new (gas-phase) intermolecular interactions were identified for the ions investigated. With the goal of generating a more complete interaction map for the gaseous WT $(S_4 + 4B)^{13+}$ ion our laboratory plans to extend these measurements to additional Ala mutants.

Table 4.4. Arrhenius activation parameters (E_a, A) measured for the loss of B from the protonated $(S_4 + 4B)^{13+}$ ions, where S_4 is WT and Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala or Tyr43Ala mutants.

S_4	E_a [a] (kcal mol^{-1})	ΔE_a [b] (kcal mol^{-1})	A [a] (s^{-1})
WT	43.2 ± 1.1	-	$10^{23.2 \pm 0.6}$
Trp79Phe	40.7 ± 0.3	-2.5 ± 1.1	$10^{21.8 \pm 0.2}$
Trp108Phe	44.8 ± 0.5	1.6 ± 1.2	$10^{24.1 \pm 0.3}$
Trp120Phe	40.6 ± 0.4	-2.6 ± 1.2	$10^{21.8 \pm 0.2}$
Ser27Ala	38.5 ± 1.0	-4.7 ± 1.4	$10^{20.7 \pm 0.6}$
Tyr43Ala	37.8 ± 0.9	-5.4 ± 1.4	$10^{20.4 \pm 0.5}$

a. Errors correspond to standard deviations. b. The values correspond to: $\Delta E_a = E_a(\text{mutant}) - E_a(\text{WT})$.

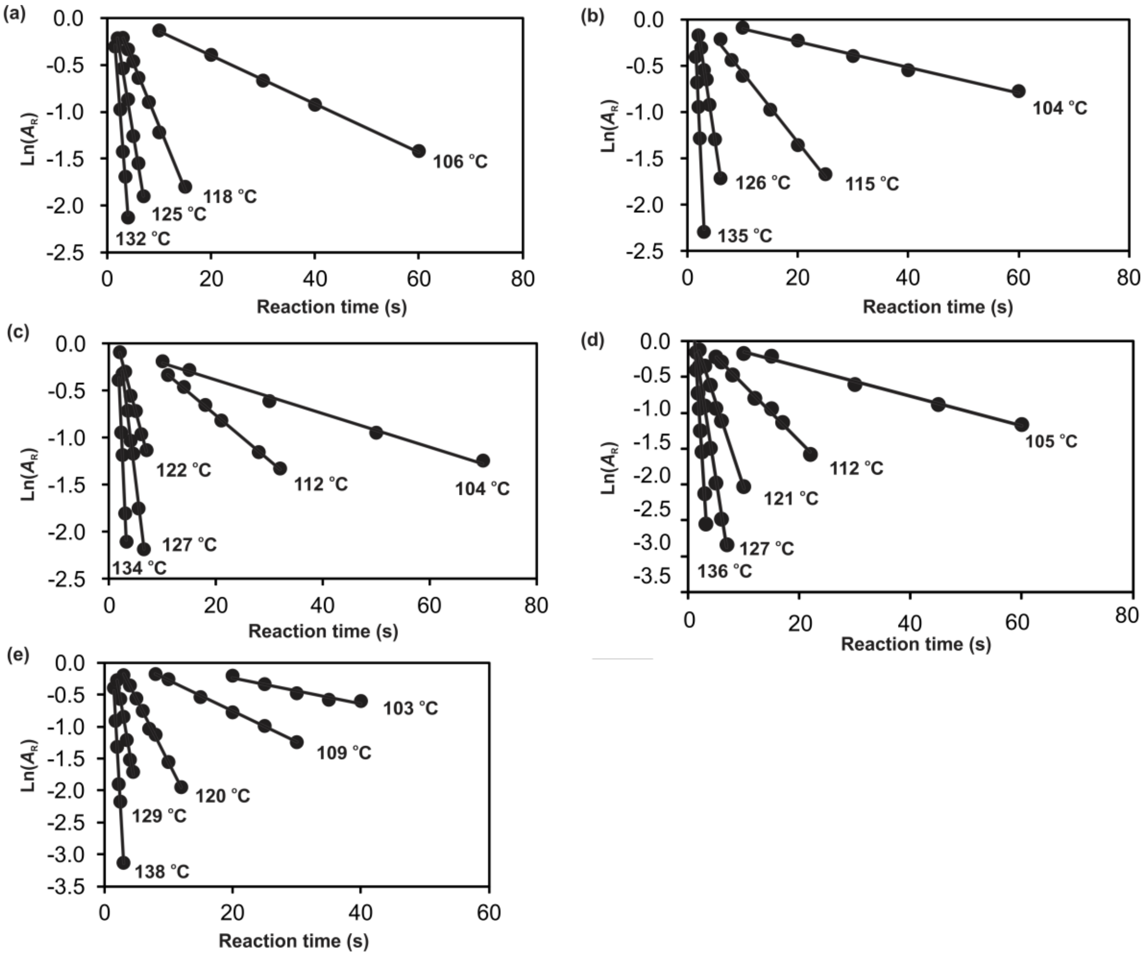


Figure 4.9. Kinetic plots of the natural logarithm of the normalized intensity (A_{R}) of (a) Trp79Phe $(S_4 + 4B)^{12+}$, (b) Trp108Phe $(S_4 + 4B)^{13+}$, (c) Trp120Phe $(S_4 + 4B)^{14+}$, (d) Ser27Ala $(S_4 + 4B)^{15+}$ and (e) Tyr43Ala $(S_4 + 4B)^{16+}$, versus reaction time at the temperatures indicated.

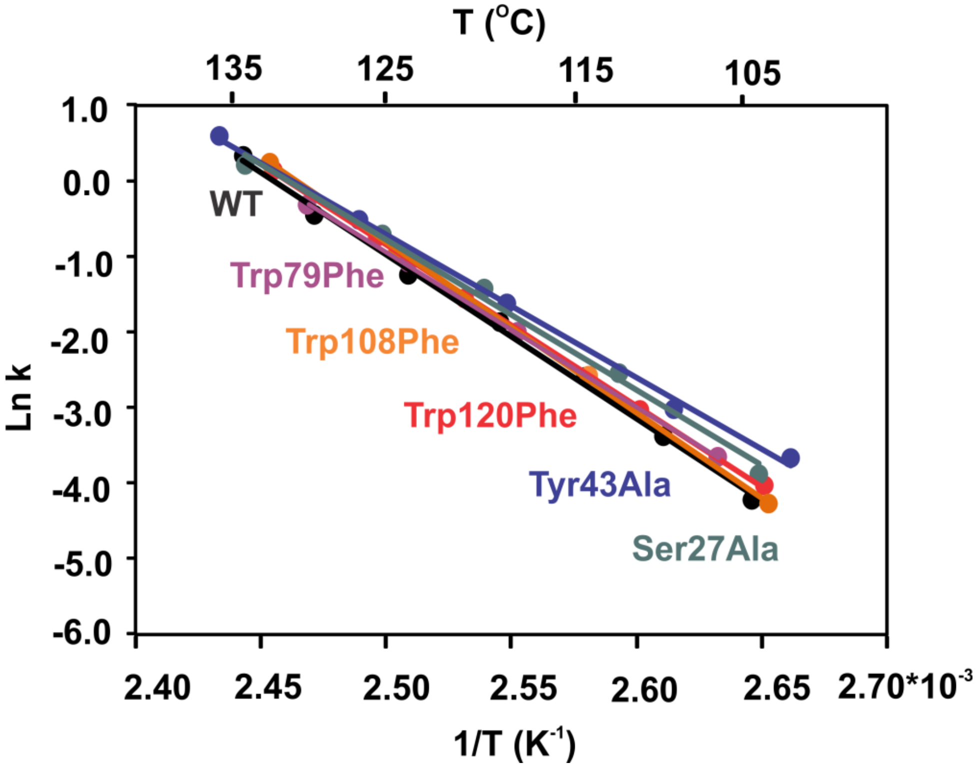


Figure 4.10. Arrhenius plots for the dissociation of gaseous $(S_4 + 4B)^{13+}$ ions composed of WT and mutant (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala and Tyr43Ala) streptavidin.

4.4.2 Computational Results

In an effort to gain additional insight into the nature of the intermolecular interactions stabilizing the gaseous WT $(S_4 + 4B)^{+12}$ and $(S_4 + 4B)^{+13}$ ions, which exhibit similar kinetic stabilities, MD simulations were performed on $(S_4 + 4B)^{12+}$ ions. The +12 charge state, with the charges partitioned equally between subunits, was selected with the view of simplifying the comparison of the intermolecular interactions in each of the subunits. Uncertainty in the appropriate location of the charges is a major challenge to implementation of MD simulations for gaseous, multiply charged ions of macromolecules and their complexes. In the present study, fifteen different charge configurations were considered. In each case, the charges were distributed equally between subunits and placed on Arg, Lys and His, which have the highest intrinsic gas phase basicity.[59] A summary of the charge sites considered for the $(S_4 + 4B)^{12+}$ ions is given in Table 4.5.

For each charge configuration, the C_α RMSD was calculated with respect to the crystal structure of the $(S_4 + 4B)$ complex, Table 4.5. In all cases the values are small, with average values of 0.372 ± 0.035 nm. The small C_α RMSD values, which are comparable in size to values reported for other protein complexes in the gas phase,[6,60] suggest relatively minor structural changes accompanying the transfer of the complex from solution to the gas phase, at least at the +12 charge state. At each charge configuration, averaged coordinates were also established and the corresponding Ω values were calculated, Table 4.5. Not surprisingly, the Ω values are somewhat sensitive to the charge configuration - the Ω values range from 36.1 to 38.1 nm^2, with an average value of 37.1 ± 0.6 nm^2. Overall, the Ω values agree quite well with the experimental value – the average value is within 8% and many of the Ω values for individual configurations are within a 1-2% of the experimental values. These results suggests that at least some of the structures obtained from the MD simulations may serve as reasonable representations of the gaseous $(S_4 + 4B)^{12+}$ ion.

Table 4.5: Analysis of results of MD simulations performed on fifteen different charge configurations for $(S_4 + 4B)^{12+}$ ions - C_α RMSD for streptavidin, collision cross sections (Ω) and number of H-bonds between biotin and streptavidin.

#	Protonated residues[a]	Cα RMSD[b] (nm)	Ω^c (nm^2)	Number of H-bonds[d]
12_1	Arg53, Arg103, Lys121	0.414	36.6	7.0
12_2	Arg59, Arg103, Lys121	0.387	36.3	7.6
12_3	Lys80, Arg103, Lys121	0.425	36.6	8.2
12_4	Arg84, Arg103, Lys121	0.358	37.1	8.0
12_5	His87, Arg103, Lys121	0.436	37.1	6.2
12_6	Arg53, Lys121, Lys132	0.352	37.5	8.6
12_7	Arg59, Lys121, Lys132	0.359	36.1	6.3
12_8	Lys80, Lys121, Lys132	0.311	36.5	7.4
12_9	Arg84, Lys121, Lys132	0.372	37.2	5.5
12_10	His87, Lys121, Lys132	0.382	37.3	7.8
12_11	Arg53, Lys121, Lys134	0.363	37.9	6.8
12_12	Arg59, Lys121, Lys134	0.361	36.6	7.5
12_13	Lys80, Lys121, Lys134	0.353	37.4	7.5
12_14	Arg84, Lys121, Lys134	0.321	37.5	6.8
12_15	His87, Lys121, Lys134	0.389	38.1	7.0
Ave		0.372	37.1	7.2
Std		0.035	0.6	0.8

a. The residues indicated correspond to the protonated residue in each subunit of the streptavidin. All remaining residues are in their neutral form. b. The C_α RMSD of streptavidin was calculated with respect to the crystal structure of the $(S_4 + 4B)$ complex. c. The Ω for the averaged coordinates at each charge configuration was calculated using Mobcal[54]. d. The total number of H-bonds between streptavidin and B for each charge configuration of the $(S_4 + 4B)^{12+}$ ions, averaged across all frames and the four subunits.

The total number of H-bonds between streptavidin and B in the $(S_4 + 4B)^{12+}$ ion at each charge configuration, averaged across all frames and the four subunits, are summarized in Table 4.5. It can be seen that there is no significant difference in the total number of H-bonds between the different configurations. Across all charge configurations, the average number of H-bonds is 7.2 ± 0.8, which is similar in number to the eight primary H-bonds identified from the crystal structure of the $(S_4 + 4B)$ complex. However, analysis of the trajectories also reveals that the H-bond donor-acceptor pairs fluctuate over time, differ between subunits and are sensitive to the charge state configuration. For example, analysis of configuration 12_1 (Table 4.5) reveals five H-bond pairs that have a fractional occupancy (*f*) > 50% in one of the subunits (Tyr43-OH---B-O, Ser45-O---B-N2H, Thr90-OH---B-S, Asp128-O---B-N1H, (adjacent subunit) Lys121-NH_3^+---B-O12) and two (Thr90-OH---B-S, (adjacent subunit) Lys121-NH_3^+---B-O12), two (Asn23-NH_2---B-O, Tyr43-OH ---B-O, Thr90-OH---B-S, (adjacent subunit) Lys121-NH_3^+---B-O12), and four H-bond pairs (Ser27-OH---B-O, Tyr43-OH---B-O, Asn49-bb---B-O11, Thr90-OH---B-S) in the other three subunits. The variation in the nature and number of the H-bonds present for the different subunits may reflect structural (crystallographic) differences between subunits, as well as structural differences arising in the heating and equilibrium process. The observation of many other H-bond pairs, with *f* < 50%, also suggests that the H-bonds are readily broken and replaced with other, energetically similar interactions.

Analysis of all the H-bonds identified among the different charge configurations for the $(S_4 + 4B)^{12+}$ ion reveals the presence of many different donor-acceptor pairs. However, only eight of these (Asn23-NH_2---B-O, Ser27-OH---B-O, Tyr43-OH---B-O, Ser45-O---B-N2H, Asn49-bb---B-O11, Thr90-OH---B-S, Asp128-O---B-N1H, (adjacent subunit) Lys121-NH_3^+---B-O12) are ever present with *f* > 50 % for a given charge configuration (Figure 4.1b). Based on this analysis it is proposed that these represent the most probable H-bonds in the gaseous $(S_4 + 4B)^{12+}$ ion. Interestingly, seven of these H-bonds are also found in the crystal structure of the $(S_4 + 4B)$ complex (Figure 4.1a); the only discrepancy being the interaction between Lys121-NH_3^+ from the adjacent subunit and B-O12, which was identified from the simulations. However, when Lys121 is in its neutral form, this interaction (which replaces the H-bond between Ser88-OH and B-O12), is absent.

The trajectories obtained for all of the charge configurations were also analyzed to establish the presence of van der Waals contacts between B and the four tryptophan residues implicated in binding in solution (Trp79, Trp92, Trp108 and Trp120 from adjacent subunit). Notably, all four of these van der Waals interactions exhibit a high *f* (>80% when averaged over all subunit and all charge configurations), suggesting that they are preserved in the gas phase (Figure 4.1b).

The results of the MD simulations performed on the $(S_4 + 4B)^{12+}$ ion suggest that the structure of the ligand binding site does not undergo significant changes when the complex is transferred to the gas phase by ESI. Furthermore, the results suggest that the specific intermolecular interactions, both H-bonds and van der Waals contacts, may survive in the gas phase; although there is also the possibility that new interactions could form. Future efforts will focus on calculating the potential of mean force (PMF)[61] for the escape of B from the binding site in the gas phase $(S_4 + 4B)^{n+}$ ions. From the PMF, the dissociation rate constants for the loss of B from the $(S_4 + 4B)^{n+}$ ions can be computed using *TS* theory.[62] This will allow for a direct comparison of the dissociation kinetics and energetics determined experimentally and predicted theoretically.

4.4.3 Kinetic stability. Solution versus gas phase

The most significant finding of the present study is the observation that the kinetics for the loss of neutral B from the protonated gaseous $(S_4 + 4B)^{n+}$ ions, at $n = 12 – 14$, are dramatically slower at 25 °C than the loss of B from the $(S_4 + 4B)$ complex in aqueous solution at pH 7.4.[25] For example, the k_{off} at 25 °C for $(S_4+4B)^{13+}$ is 3.1×10^{-9} s^{-1} (value extrapolated from the measured Arrhenius plots) compared to 5.4×10^{-6} s^{-1} in solution.[27] This is, perhaps, a surprising result given that the $(S_4 + 4B)$ complex is among the most kinetically stable protein-ligand complexes known[27,35-36] and reveals that the streptavidin-biotin interaction has the inherent capacity for even greater stability than previously recognized. As noted above, the enhanced kinetic stability of this interaction in the gas phase, compared to aqueous solution, is the result of a much larger dissociation E_a. For example, the E_a for the pathway leading to the loss of neutral B, which is not significantly affected by Coulombic repulsion, is as much as 13 kcal mol^{-1} larger than that measured in solution.[25]

As described above, the experimental and theoretical results obtained for the $(S_4 + 4B)^{n+}$ ions at the lower charge states suggest that many, if not most, of the intermolecular interactions present in solution are preserved in the gas phase. Therefore, it is reasonable to expect that solvent effects are responsible, at least in part, for the difference in E_a values measured for the gaseous $(S_4 + 4B)^{n+}$ ions and solvated $(S_4 + 4B)$ complex.[8] If the loss of B from the $(S_4 + 4B)$ complex, in its hydrated and dehydrated forms, proceeds through similar dissociation pathways in solution and the gas phase, then the difference between E_a in solution ($E_{a,aq}$) and the E_a in the gas phase ($E_{a,g}$) will reflect the difference in the hydration enthalpy of the *TS* ($\Delta H_{hydr,TS}$) and the reactant ($\Delta H_{hydr,R}$):[8]

$$E_{a,aq} - E_{a,g} \approx \Delta H_{hydr,R} - \Delta H_{hydr,TS} \tag{10}$$

Although there is no way to accurately determine the magnitude of the solvent effect, i.e., the ($\Delta H_{hydr,R} - \Delta H_{hydr,TS}$) term, the effect can be estimated in the case where dissociation proceeds through a late *TS*, wherein the ligand is fully removed from the binding cavity, or nearly so, and assuming no other structural changes accompany ligand loss. In this situation, the magnitude of the solvent effect will reflect the enthalpies of hydration of B and the amino acid residues originally involved in binding.[8]

To our knowledge, the enthalpy of hydration (ΔH_{hyd}) of deprotonated B has not been measured experimentally. However, this value can be estimated from the enthalpy of solution of B in water (ΔH_{sol}) and the enthalpy of vaporization of B to give an ideal gas (ΔH_{vap}), eq 11:[63]

$$\Delta H_{hyd} = \Delta H_{sol} - \Delta H_{vap} \tag{11}$$

The ΔH_{sol} for neutral B has been determined to be 9.7 ± 0.4 kcal mol^{-1} and the enthalpy of ionization of B is reported to be only 0.3 kcal mol^{-1}.[64] While there is no experimental ΔH_{vap} value for B it is estimated, using the Joback and Reid Group Contribution Method,[65] to be 20.3 kcal mol^{-1}. Based on these values, ΔH_{hyd} of B is estimated to be approximately -10 kcal mol^{-1}. Despite the uncertainty associated with this value, this analysis suggests that the bulk of the difference in E_a values for the loss of B in solution and the gas phase (as a neutral) can be accounted for by the hydration of B in the *TS.* The remaining energy difference potentially reflects the hydration of the exposed residues within the binding cavity in the *TS*. It is also possible that structural differences or differences in the mechanisms by which B escapes from the binding site in solution and in the gas phase are reflected in the differences in E_a values measured for the gaseous $(S_4 + 4B)^{n+}$ ions and the solvated $(S_4 + 4B)$ complex. The results of the planned mutagenesis study, *vide supra*, will allow for a more definitive conclusion regarding the structural similarities of the hydrated and dehydrated forms of the $(S_4 + 4B)$ complex. Similarly, a comparison of the PMF calculated for the escape of B from the gaseous $(S_4 + 4B)^{n+}$ ions and from the $(S_4 + 4B)$ complex in aqueous solution will help to establish similarities in the ligand escape pathways in the presence and absence of solvent. These planned computational studies will also provide an important opportunity to test the reliability of available theoretical models for describing protein-ligand interactions, both in the presence and absence of solvent.

4.5 Conclusions

In summary, the results of the first detailed study of the structure and kinetic stability of gaseous protonated ions of the (S_4 + 4B) complex are reported. Collision cross sections measured for the gaseous WT S_4^{n+} and WT $(S_4 + 4B)^{n+}$ ions, at n = 12 - 16, are found to agree within 10% of values estimated for the crystal structures for S_4 and (S_4 + 4B). These results suggest that S_4 and (S_4 + 4B) complexes do not undergo significant structural changes upon transfer from solution to the gas phase by ESI. A comparison of dissociation E_a for the B loss from the $(S_4 + 4B)^{13+}$ ions composed of WT and five binding site mutants suggests that at least some of the solution specific intermolecular interactions are preserved in the gas phase. The results of MD simulations performed on WT $(S_4 + 4B)^{12+}$ ions with different charge configurations also suggest that the solution specific intermolecular H-bonds and van der Waals contacts are preserved in the gas phase. The most significant finding of this study is the observation that the gaseous WT $(S_4 + 4B)^{n+}$ ions at n = 12 – 14, owing to a much larger E_a (by as much as 13 kcal mol^{-1}) for the loss of B, are dramatically more stable kinetically at 25 °C than the (S_4 + 4B) complex in aqueous neutral solution. It is shown that the differences in E_a values measured for the gaseous $(S_4 + 4B)^{n+}$ ions and solvated (S_4 + 4B) complex can be largely accounted for by the rehydration of B in the dissociative *TS*.

4.6 References

1. Schmidtke, P.; Luque, F. J.; Murray, J. B.; Barril, X. *J. Am. Chem. Soc.* **2011**, *133*, 18903-18910.

2. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. *Nat Rev Drug Discov* **2006**, *5*, 730-739.

3. Swinney, D. C. *Curr. Opin. Drug Discovery Dev.* **2009**, *12*, 31-39.

4. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *Angew. Chem. Int. Ed.* **2004**, *43*, 4183-4186.

5. Shoemaker, G. K.; Kitova, E. N.; Palcic, M. M.; Klassen, J. S. *J. Am. Chem. Soc.* **2007**, *129*, 8674-8675.

6. Kitova, E. N.; Seo, M.; Roy, P.-N.; Klassen, J. S. *J. Am. Chem. Soc.* **2008**, *130*, 1214-1226.

7. Liu, L.; Bagal, D.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2009**, *131*, 15980-15981.

8. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2010**, *132*, 17658-17660.

9. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 3054-3060.

10. Wyttenbach, T.; Bowers, M. T. *Annu. Rev. Phys. Chem.* **2007**, *58*, 511-533.

11. Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432-14433.

12. Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V. *J. Am. Chem. Soc.* **2012**, *134*, 3429-3438.

13. Rose, R. J.; Labrijn, A. F.; van den Bremer, E. T. J.; Loverix, S.; Lasters, I.; van Berkel, P. H. C.; van de Winkel, J. G. J.; Schuurman, J.; Parren, P. W. H. I.; Heck, A. J. R. *Structure* **2011**, *19*, 1274-1282.

14. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 5902-5913.

15. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 9340-9341.

16. Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6423-6427.

17. Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.

18. Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.

19. Chaiet, L., Wolf, F. J. *Arch. Biochem. Biophys.* **1964**, *106*, 1-5.

20. Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2190-2194.

21. Stenkamp, R. E.; Trong, I. L.; Klumb, L.; Stayton, P. S.; Freitag, S. *Protein Sci.* **1997**, *6*, 1157-1166.

22. Le Trong, I.; Wang, Z.; Hyre, D. E.; Lybrand, T. P.; Stayton, P. S.; Stenkamp, R. E. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 813-821.

23. Hyre, D. E.; Le Trong, I.; Merritt, E. A.; Eccleston, J. F.; Green, N. M.; Stenkamp, R. E.; Stayton, P. S. *Protein Sci.* **2006**, *15*, 459-467.

24. Freitag, S.; Chu, V.; Penzotti, J. E.; Klumb, L. A.; To, R.; Hyre, D.; Le Trong, I.; Lybrand, T. P.; Stenkamp, R. E.; Stayton, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8384-8389.

25. Klumb, L. A.; Chu, V.; Stayton, P. S. *Biochemistry* **1998**, *37*, 7657-7663.

26. Hyre, D. E.; Stayton, P. S.; Trong, I. L.; Freitag, S.; Stenkamp, R. E. *Protein Sci.* **2000**, *9*, 878-885.

27. Chilkoti, A.; Stayton, P. S. *J. Am. Chem. Soc.* **1995**, *117*, 10622-10628.

28. Chu, V.; Stayton, P. S.; Freitag, S.; Le Trong, I.; Stenkamp, R. E. *Protein Sci.* **1998**, *7*, 848-859.

29. Magalhães, M. L. B.; Czekster, C. M.; Guan, R.; Malashkevich, V. N.; Almo, S. C.; Levy, M. *Protein Sci.* **2011**, *20*, 1145-1154.

30. Cerutti, D. S.; Trong, I. L.; Stenkamp, R. E.; Lybrand, T. P. *J. Phys. Chem. B* **2009**, *113*, 6971-6985.

31. DeChancie, J.; Houk, K. N. *J. Am. Chem. Soc.* **2007**, *129*, 5419-5429.

32. Dixit, S. B.; Chipot, C. *J. Phys. Chem. A* **2001**, *105*, 9795-9799.

33. Li, Q.; Gusarov, S.; Evoy, S.; Kovalenko, A. *J. Phys. Chem. B* **2009**, *113*, 9958-9967.

34. Stayton, P. S.; Freitag, S.; Klumb, L. A.; Chilkoti, A.; Chu, V.; Penzotti, J. E.; To, R.; Hyre, D.; Le Trong, I.; Lybrand, T. P.; Stenkamp, R. E. *Biomol. Eng* **1999**, *16*, 39-44.

35. Green, N. M. In *Adv. Protein Chem.* 1975; Vol. 29, p 85-133.

36. Piran, U.; Riordan, W. J. *J. Immunol. Methods* **1990**, *133*, 141-143.

37. Hyre, D.; Amon, L. M.; Penzotti, J. E.; Trong, I. L.; Stenkamp, R. E.; Lybrand, T. P.; Stayton, P. S. *Nat. Struct. Biol.* **2002**, *9*, 582-585.

38. Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1754-1758.

39. Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2006**, *79*, 416-425.

40. Bagal, D.; Kitova, E. N.; Liu, L.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. *Anal. Chem.* **2009**, *81*, 7801-7806.

41. Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman P. A. *AMBER 11*; University of California, San Francisco, **2010**.

42. Ramasubbu, N.; Ragunath, C.; Mishra, P. J.; Thomas, L. M.; Gyémánt, G.; Kandra, L. *Eur. J. Biochem.* **2004**, *271*, 2517-2529.

43. Holm, L.; Rosenström, P. *Nucleic Acids Res.* **2010**, *38*, W545-W549.

44. Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. *Nucleic Acids Res.* **2011**, *39*, W511-W517.

45. Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821-7839.

46. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.1*; Gaussian, Inc., Wallingford CT, **2009**.

47. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247-260.

48. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999-2012.

49. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157-1174.

50. Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. *Anal. Chem.* **2010**, *82*, 9557-9565.

51. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protoc.* **2008**, *3*, 1139-1152.

52. Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R. *Chem. Soc. Rev.* **2010**, *39*, 1633-1655.

53. Wildgoose, J.; Giles, K.; Pringle, S.; Koeniger, S.; Valentine, S.; Bateman, R.; Clemmer, D. *Abstracts of papers,* Proceedings of the 54th ASMS Conference on Mass Spectrometry and Allied Topics, Seattle Wash.; American Society for Mass Spectrometry: Washington, DC, **2006**, Abstract 064.

54. Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082-16086.

55. University of Illinois at Urbana-Champaign. VMD sofetware Home Page. http://www.ks.uiuc.edu/Research/vmd/(accessed Dec 21, 2011).

56. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781-1802.

57. University of Illinois at Urbana-Champaign. NAMD sofetware Home Page. http://www.ks.uiuc.edu/Research/namd/ (accessed Dec 21, 2011).

58. Sinelnikov, I.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 617-631.

59. Hunter, E. P. L.; Lias, S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413-656.

60. Patriksson, A.; Marklund, E.; van der Spoel, D. *Biochemistry* **2007**, *46*, 933-945.

61. Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300-313.

62. Schenter, G. K.; Kathmann, S. M.; Garrett, B. C. *J. Chem. Phys.* **1999**, *110*, 7951-7959.

63. Nilsson, S. O.; Wadsö, I. *J. Chem. Thermodyn.* **1986**, *18*, 673-681.

64. Suurkuusk, J.; WadsÖ, I. *Eur. J. Biochem.* **1972**, *28*, 438-441.

65. Joback, K. G.; Reid, R. C. *Chem. Eng. Commun.* **1987**, *57*, 233-243.

Chapter 5

Mapping Protein-Ligand Interactions in the Gas Phase Using a Functional Group Replacement Strategy. Comparison of CID and BIRD Activation Methods[*]

5.1 Introduction

Most biological processes, including the immune response, cell-cell communication, inflammation, and bacterial and viral infection, are dependent on the formation of non-covalent protein interactions. [1] The structure and stability of these complexes are determined by the concerted action of many forces (e.g. hydrogen (H-) bonds, ionic and van der Waals interactions) between binding partners and from the displacement and reorganization of solvent molecules associated with the binding partners.[2] An understanding of these forces and the structures they lead to is essential to a complete description of biochemical processes. Gas-phase studies of desolvated protein-ligand complexes afford an opportunity to probe their intrinsic properties, free of solvent effects. In principle, insights into the role of solvent in biomolecular recognition may be gained from a comparison of the stabilities of protein-ligand complexes in their hydrated and

---

[*] A version of this chapter has been submitted for publication: Deng L.; Kitova, E. N.; Klassen, J. S. *J. Am. Soc. Mass Spectrom.*

dehydrated states. However, for this approach to be successful, the dominant stabilizing intermolecular interaction in solution must be preserved upon desolvation. BIRD[3-4], a thermal dissociation technique implemented with a FTICR mass spectrometer, combined with a FGR strategy is arguably the most reliable method for identifying individual interactions in gaseous biological complexes and quantifying the strength of these interactions. [2,5-10] To establish whether a particular residue (on protein) or functional group (on ligand) is involved in binding, the group is modified (chemically or biochemically) in such a way that any pre-existing interaction is lost. The E_a for the loss of ligand from the modified complex is then compared to the E_a for the ligand loss from the unmodified complex. A decrease in E_a indicates that the particular residue/functional group (that was modified) was involved in stabilizing the complex. Moreover, the differences in E_a for the modified and unmodified complexes, *i.e.* $\Delta E_a = E_a$(unmodified) – E_a(modified), provides a measure of the strength of the interaction. Ideally, to identify interaction pairs (e.g. H-bond donor/acceptor pairs), a three step approach is utilized, in which E_a values are determined for complexes with modified ligand, modified protein and modified protein and ligand. For a given interaction pair, the E_a values determined for the three complexes are expected to be identical.

Using the BIRD-FGR method, the intermolecular interactions in the gaseous ions of three protein-ligand complexes: scFv, a model carbohyhdrate binding protein, and a trisaccharide ligand (L1) [2,5-6], Lg and FAs [7-9] and the high affinity interaction involving S_4 and B [10], were recently elucidated. Results obtained for protonated (scFv + L1)$^{n+}$ ions revealed that some of the specific intermolecular H-bonds are conserved upon transfer of the complex from solution to the gas phase by ESI[2]. Kinetic data measured for the loss of FA from deprotonated (Lg + FA)$^{7-}$ ions suggest that the acyl chains of the FAs are retained within the hydrophobic cavity of Lg in the gas phase [7]. Furthermore, a comparison of the dissociation rate constants measured at 25°C in the gas phase and in aqueous solution revealed that (Lg + FA) complexes are more stable kinetically in the absence of water [8]. Kinetic data for ligand loss from the protonated (S_4 + 4B)$^{n+}$ ions, together with the results of MD simulations, suggested that the solution specific intermolecular H-bonds and van der Waals contacts are preserved in the gas phase.[10] Moreover, the differences in dissociative E_a values measured for the gaseous complex ions and hydrated (S_4 + 4B) complex can be accounted for by the rehydration of B in the dissociative *TS*.[10]

The aforementioned studies highlight the power of the BIRD-FGR method to reveal new insights, inaccessible by other methods, into the forces responsible for the structure and kinetic stability of protein-ligand complexes. However, application of the BIRD-FGR method requires access to a FTICR mass spectrometer equipped with a temperature-controlled ion cell. Additionally, the determination of Arrhenius activation parameters requires kinetic measurements performed over a range of temperatures, which can be time consuming, especially at low reaction temperatures where the kinetics are slow. Consequently, there is interest in adapting the BIRD-FGR method for use with other, more readily available activation techniques, such as CID[11]. For example, Douglas and coworkers described a CID-based approach to measure the collision energies required to dissociate the gas-phase ions of complexes of small-molecule inhibitors bound to the WT form and single point binding site mutants (Asn44Ala, Gln87Met, and Gln87Tyr) of the catalytic domain of Cex, an enzyme that hydrolyses xylan and xylo-oligosaccharides. The authors found that the dissociation energies followed the same trend as the thermodynamic stabilities in aqueous solution and it was concluded that specific intermolecular H-bonds (involving residues Asn44 and Gln87) were preserved in the gas phase. [12] More recently, Oldham and coworkers described the results of a CID study on the complexes of a series of mutants of the FK506-binding protein (FKBP) and its ligand FK506. [13] Collisional activation of (FKBP + FK506) ions composed of WT or mutant protein revealed that the removal of native protein–ligand interactions formed between residues Asp37 and Tyr82, and FK506 significantly destabilized the gas-phase complex. The gas-phase results obtained for the complexes composed of single point mutants (Asp37Gly and Tyr82Phe) agreed qualitatively with the trend in affinities, evidence for the preservation of these interactions in the desolvated WT complex.

The results of these CID-FGR studies are very promising. However, the reliability of this general approach for identifying intermolecular interactions in gaseous protein-ligand complexes remains unclear. With this in mind, energy-resolved CID measurements were performed on protonated ions of a series of structurally related complexes based on (scFv + L1) and (S_4 + 4B) complexes.

Intermolecular interactions were identified from decreases in the collision energy required to dissociate 50% of the reactant ion (i.e., Ec_{50}) upon modification of residues or functional groups that could participate in intermolecular interactions. The identified interactions were then compared to interaction maps generated using the BIRD-FGR approach (Figure 5.1) in order to test the reliability of the CID-FGR method.

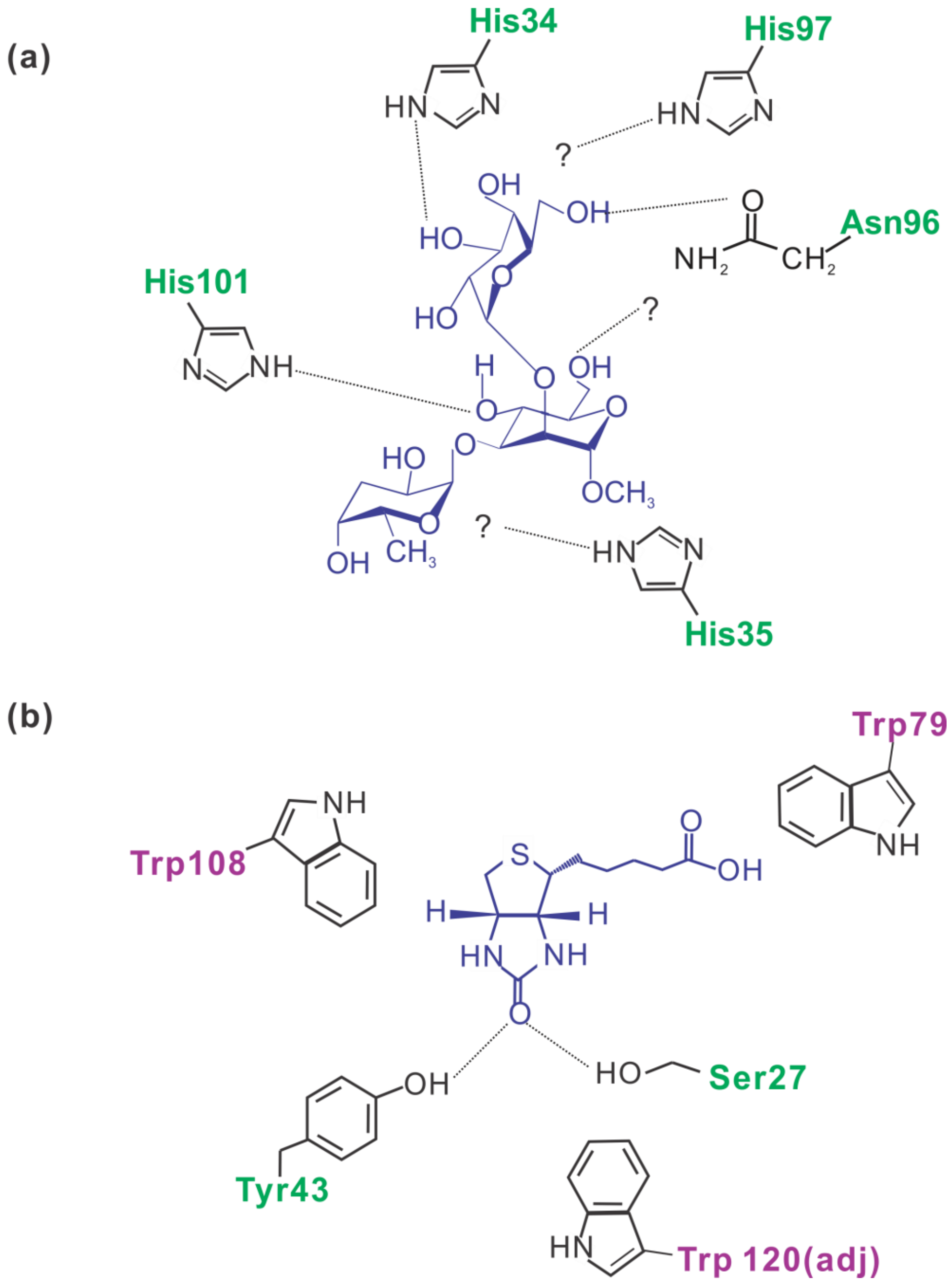


Figure 5.1. Map of intermolecular interactions determined using BRID-FGR method for the protonated (a) (scFv + L1)$^{10+}$ and (b) (S_4 + 4B)$^{13+}$ ions. The maps were adapted from those reported in [2,10], respectively, where adj represents a residue from the adjacent subunit.

5.2 Materials and Methods

5.2.1 Sample Preparation

Equine heart cytochrome *c* (cyto *c*, MW 12 358 Da), bovine β-lactoglobulin (Lg, MW 18 363 Da) and hen egg white avidin (Avidin, MW 15 636 Da) were purchased from Sigma-Aldrich Canada (Oakville, Canada). Stock solutions of cyto *c*, Lg, and Avidin (100 µM) were prepared by exchanging each protein into 100 mM ammonium acetate using Vivaspin 500 centrifugal concentrators with a 10 kDa MW cut-off (Sartorius Steddin Biotech, Gottingen, Germany). The recombinant scFv (MW 26 539 Da) of the monoclonal antibody Se155-4 and the single-point binding site mutant, His101Ala, were expressed in *E.coli* and isolated and purified using procedures described previously. [14-15] The plasmid for natural core streptavidin (containing residues 13-139 of WT streptavidin, monomer MW 13 271 Da) was a gift from Prof. P. Stayton (University of Washington). The single-point binding site mutants, Ser27Ala, Tyr43Ala Trp79Phe, Trp108Phe and Trp120Phe, were prepared using site-directed mutagenesis. The WT and mutant streptavidin proteins were expressed in *E. coli* and purified using procedures described elsewhere.[16] Solutions of purified protein were exchanged directly into 100 mM aqueous ammonium acetate buffer using an Amicon microconcentrators with a 10 kDa MW cut-off (Millipore Corp., Billerica, MA) and lyophilized. Protein (scFv or S_4) stock solutions were prepared at a concentration of 100 µM in 50 mM aqueous ammonium acetate (pH 6.8) and stored at -20 ℃ until needed.

The synthetic trisaccharide ligands for scFv, α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-α-Man*p*-$\mathrm{OCH_3}$ (L1, MW 486.5 Da), 3-deoxy-α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-α-D-Man*p*-$\mathrm{OCH_3}$ (L2, MW 470.5 Da), 6-deoxy-α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-α-D-Man*p*-$\mathrm{OCH_3}$ (L3, MW 470.5 Da), α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-4-deoxy-α-D-Man*p*-$\mathrm{OCH_3}$ (L4, MW 470.5 Da), α-D-Gal*p*-(1→2)-[α-D-Abe*p*-(1→3)]-6-deoxy-α-D-Man*p*-$\mathrm{OCH_3}$ (L5, MW 470.5 Da), were gifts from Prof. D. Bundle (University of Alberta). Biotin (B, MW 244.3 Da) was purchased from Sigma-Aldrich Canada (Oakville, Canada). Ligand stock solutions were prepared at a concentration of 200 μM in Milli-Q water and stored at -20 °C until needed.

To produce gaseous protonated $(\mathrm{scFv} + \mathrm{L})^{n+}$ ions, ESI was performed on solutions containing WT scFv (8 μM) (or WT scFv and His101Ala scFv (10 μM)), and L (5 - 15 μM) in 10 mM ammonium acetate (pH 6.8). The concentrations of protein and ligand were adjusted to minimize the formation of nonspecific (scFv + L) complexes during the ESI process. [17] The gaseous WT $(\mathrm{S_4} + 4\mathrm{B})^{n+}$ and mutant $(\mathrm{S_4} + 4\mathrm{B})^{n+}$ ions were produced by ESI performed on solutions containing WT $\mathrm{S_4}$ (8 μM) and one of the mutants $\mathrm{S_4}$ (10 μM), and B (80 μM) in 10 mM ammonium acetate (pH 6.8). Imidazole (5 mM) was added to reduce charge states of the gaseous $(\mathrm{S_4} + 4\mathrm{B})^{n+}$ ions so as to enhance the abundance of the $(\mathrm{S_4} + 4\mathrm{B})^{13+}$ ion. [18] For IMS measurements, ESI was performed on solutions (10 μM) of each calibrant protein (cyto *c*, Lg, and Avidin) in 10 mM ammonium acetate (pH 6.8).

5.2.2 Mass Spectrometry

Measurements were performed using a Synapt G2-S Q-IMS-TOF mass spectrometer (Waters, UK), equipped with a nanoESI ion source. NanoESI was performed using borosilicate tubes (1.0 mm o.d., 0.68 mm i.d.), pulled to ~5 μm o.d. at one end using a P-97 micropipette puller (Sutter Instruments, Novato, CA). The electric field required to spray the solution in positive ion mode was established by applying a voltage of 1.0 – 1.6 kV to a platinum wire inserted inside the glass tip. The solution flow rate was typically ~20 nL min^{-1}. A cone voltage of 30 V was used and the source block temperature was maintained at 80 ℃. The gaseous protein-ligand complexes, at the desired charge states, were isolated using the quadrupole mass filter. CID of the isolated ions was carried out in the Trap region, by collision with Ar gas. A series CID experiments were performed on the $(\text{scFv} + \text{L})^{+10}$ ions at different trap pressures (P_{trap} = 4, 6 and 8) to investigate the influence of collision gas pressure on the CID results. For all other measurements, an Ar pressure of 2.1×10^{-2} mbar (P_{trap} = 6) was used. The collision energy was controlled by adjusting the Trap collision energy voltage. Data were acquired at 5 V increments; at collision regions corresponding to ~50% dissociation, data were collected at 1 V increments to ensure more precise determination of Ec_{50} values. The TOF-MS was operated over a *m/z* range of 50 to 8000 at a pressure of 1.2×10^{-6} mbar. Three replicate measurements (using 1 min acquisition times) were performed at each collision energy.

To identify small differences in the Ec_{50} values for complexes composed of WT and mutant protein, the CID measurements were performed simultaneous on complexes of WT and a given mutant (e.g. WT $(scFv + L1)^{10+}$ and His101Ala $(scFv+ L1)^{10+}$; WT $(S_4 + 4B)^{13+}$ and Try43Ala $(S_4 + 4B)^{13+}$). For these measurements, ESI was performed on solutions containing both mutant and WT protein, at concentrations that yielded similar ion intensities for the ions of interest. Quadrupole isolation of each complex was carried out in an alternating fashion. This approach was adopted in order to minimize uncertainty in the Ec_{50} values resulting from fluctuations in instrumental conditions.[19] A similar approach could not be implemented for the comparison of Ec_{50} values for (scFv + L) complexes composed of L1 and the monodeoxy analogs (L2 - L5) because of the small difference in the *m/z* values. For this reason, the CID measurements were carried out on each complex individually.

IMS measurements on the WT $(scFv + L1)^{n+}$ ions were performed with a fixed wave height of 35 V and velocity of 1200 m s^{-1}. The Trap CE and Transfer CE were maintained at 5 V and 2 V, respectively. The bias on Trap DC was 40 V. Ar was used in the Trap and Transfer ion guides at a pressure of 2.97×10^{-2} mbar and 3.04×10^{-2} mbar, respectively. The helium chamber preceding the traveling wave ion mobility (TWIMS) device was operated with a He flow rate of 140 mL min^{-1}. All traveling-wave IMS measurements were performed using N_2 as the mobility gas at a flow rate of 40 mL min^{-1} and pressure of 8.90 mbar. Determination of the Ω from drift time measurements was carried out using protocols described elsewhere. [10,20-21] Briefly, proteins (cyto *c*, Lg, and Avidin) with known Ω (in He) were analysed under the same experimental conditions (as for the $(scFv + L1)^{n+}$ ions) to establish a correlation between the t_D and Ω. A linear calibration curve of literature Ω values versus corrected t_D was generated with an R^2 value of 0.994, Figure 5.2.

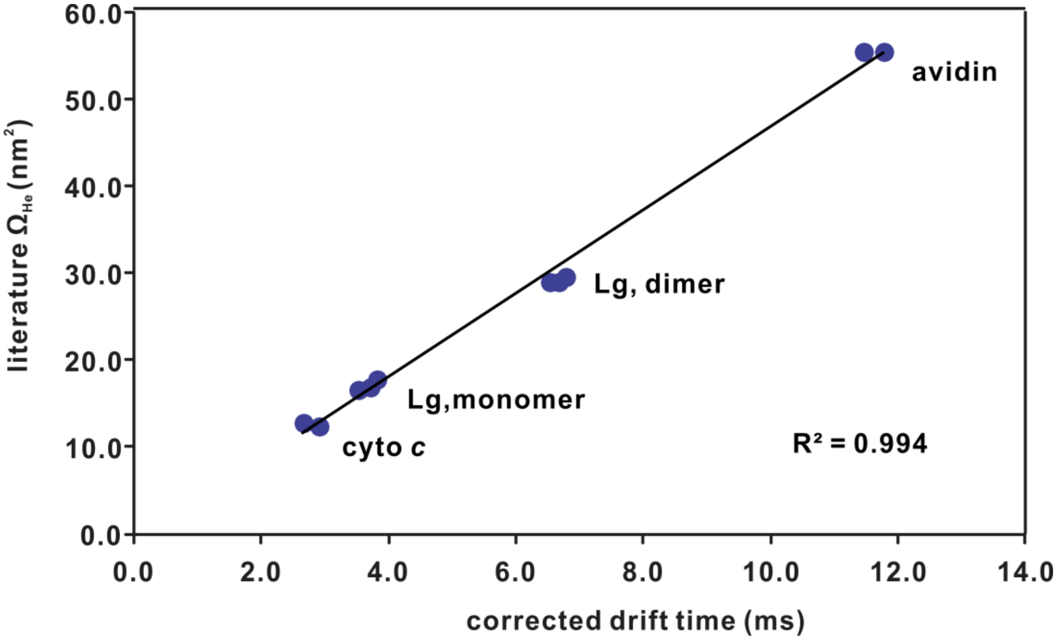


Figure 5.2. Calibration curve, based on the calibrants cyto *c*, Lg, and Avidin, displayed as a linear plot of literature Ω_{He} values and corrected drift times.

The Ω values of the protonated $(scFv + L1)^{n+}$ ions were then determined from the calibration curve and the corresponding t_D values. The Ω for the reported (scFv + L1) crystal structure (protein ID: 1MFA) [14] was calculated using Mobcal [22] and the TM.

5.2.3 CID data analysis

Dissociation profiles were constructed from the percentage precursor remaining (%precursor) as a function of collision energy. At each collision energy, %precursor was calculated using eq 5.1:

$$\%\text{precursor} = \frac{Ab_{\text{precursor}}}{Ab_{\text{precursor}} + \sum Ab_{\text{product}}} \times 100 \tag{5.1}$$

where $Ab_{\text{precursor}}$ and $\sum Ab_{\text{product}}$ are the abundances of the precursor ion and the sum of the abundances of all product ions, respectively. The dissociation profiles were treated as sigmoid curves using Origin 8 (OriginLab, Northampton, MA) by eq 5.2:

$$\%\text{precursor} = \frac{a}{1 + \mathrm{e}^{-k(\mathrm{Ec} - x_c)}} \times 100 \tag{5.2}$$

where a is the amplitude, k is the coefficient and x_c is the center of the sigmoid function. From each dissociation profile, the corresponding parameters (a, k and x_c) and associated uncertainties (δa, δk and δx_c) were determined.

The collision energy required to dissociate 50% of the reactant ion (i.e., the Ec_{50} values) were then calculated by eq 5.3:

$$\mathrm{Ec}_{50} = -k^{-1}\ln(2a - 1) + x_c \tag{5.3}$$

The uncertainties of Ec_{50} values ($\delta\mathrm{Ec}_{50}$) were calculated from error propagation and eq 5.4:

$$\delta\mathrm{Ec}_{50} = \left|\frac{\mathrm{dEc}_{50}}{\mathrm{d}k}\right|\delta k + \left|\frac{\mathrm{dEc}_{50}}{\mathrm{d}a}\right|\delta a + \left|\frac{\mathrm{dEc}_{50}}{\mathrm{d}x_c}\right|\delta x_c \tag{5.4}$$

5.3 Results and Discussion

5.3.1 Mapping the intermolecular interactions in the (scFv + L1)$^{10+}$ ion

Protonated gas-phase ions of the (scFv + L) complexes composed of WT scFv and native trisaccharide ligand (L1) or monodeoxy analogs (L2 – L5) were produced by ESI performed on aqueous ammonium acetate solutions (10 mM, pH

6.8) containing WT scFv (8 μM) and L (5 - 10 μM). Abundant protonated ions of ligand-bound and unbound scFv were detected, e.g. scFv^{n+} and (scFv + L)$^{n+}$ ions at n = 9 – 11. Notably, there was no evidence of nonspecific ligand binding to scFv under the experimental conditions used. An illustrative ESI mass spectrum acquired for a solution of scFv and L1 is shown in Figure 5.3. Also shown are the corresponding arrival time distributions measured for the (scFv + L1)$^{n+}$ ions. The narrow charge state distribution is suggestive of a compact protein structure in the gas phase. This conclusion is further supported by the results of the IMS measurements performed on the (scFv + L1)$^{n+}$ ions. The **Ω** values for the (scFv + L1)$^{n+}$ ions, at charge states 9 - 11, which were determined from their measured t_D and the calibration curve, are listed in Table 5.1, together with the **Ω** value estimated for the crystal structure of (scFv + L1) (protein ID: 1MFA).[14] The **Ω** values measured for the (scFv + L1)$^{n+}$ ions exhibit only a very subtle dependence on charge state, suggesting that charge-induced structural changes are, at least at these charge states, minimal. The average **Ω** value for the (scFv + L1)$^{n+}$ ions (22.3 ± 0.4 nm^2) agrees within 4% with the **Ω** value estimated for the (scFv + L1) crystal structure (23.2 ± 0.4 nm^2). From this it is concluded that the transfer of the (scFv + L1) complex (and presumably the other (scFv + L) complexes) from neutral aqueous solution to the gas phase by ESI (in positive ion mode) is not accompanied by significant structural changes.

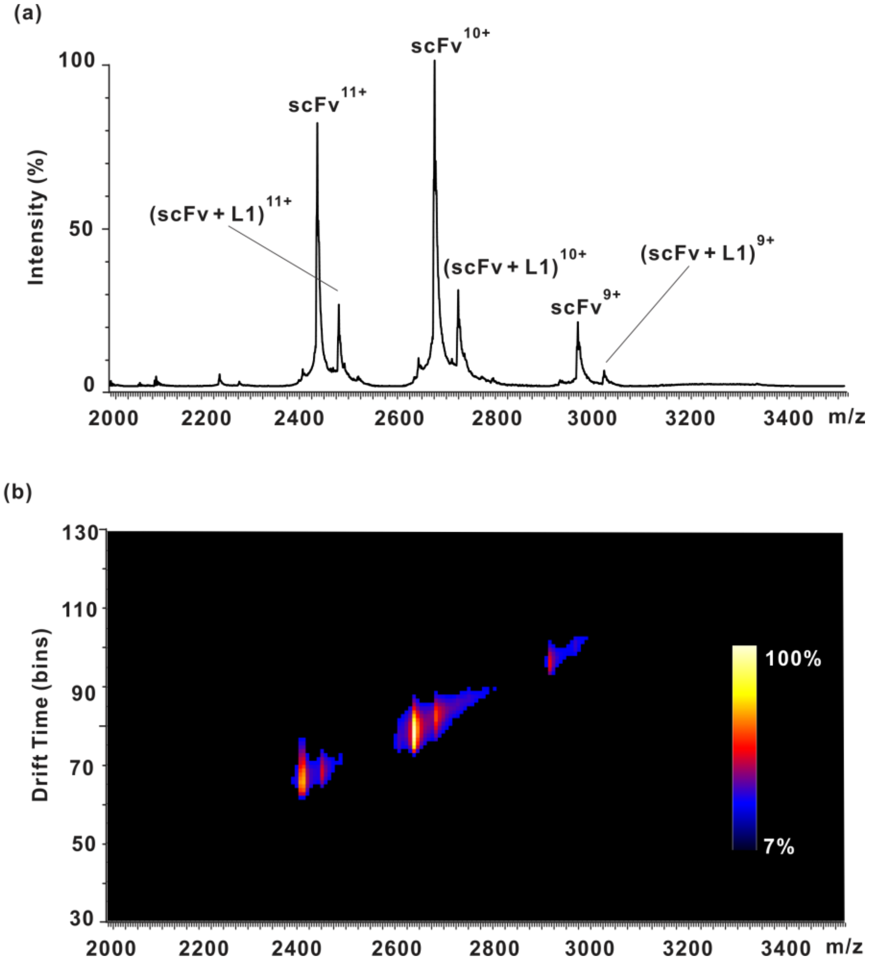


Figure 5.3. (a) ESI mass spectrum acquired for a solution of scFv (8 μM), L1 (5 μM) and 10 mM ammonium acetate (pH 6.8). (b) Ion mobility heat map of arrival time distribution versus *m*/*z*. The normalized ion intensities (from 7% to 100%) are represented using the indicated colour scale.

Table 5.1. Comparison of experimentally determined collision cross sections (Ω) for the gaseous (scFv + L1)$^{n+}$ ions, where n = 9 – 11, with Ω values calculated for crystal structure reported for the (scFv + L1) complex (protein ID: 1MFA) [14].

Charge state	**(scFv+L1)$^{n+}$**
(n)	**Ω (nm^2)**

11	22.68 ± 0.01 [a]
10	22.46 ± 0.04 [a]
9	21.86 ± 0.05 [a]
Average	22.33 ± 0.43 [b]
Crystal structure	23.24 ± 0.36 [c]

a. Uncertainty corresponds to the standard deviation of three replicate measurements. b. Uncertainty corresponds to the standard deviation of the Ω values determined at charge states +9, +10 and +11. c. Ω was determined from crystal structure (protein ID: 1MFA) [14] using Mobcal [22] and the trajectory method; the reported error corresponds to one standard deviation of values from replicate (10) calculations.

The +10 charge state of the (scFv + L) complexes was selected for all of the CID measurements. A representative CID mass spectrum for the (scFv + L1)$^{10+}$ ion is shown in Figure 5.4a; for comparison purposes, a BIRD mass spectrum is shown in Figure 5.4b. CID of the protonated (scFv + L)$^{10+}$ ions in the Trap region led exclusively to the formation of the scFv^{10+} ion (produced by the loss of neutral L); the scFv^{9+} ion (produced by the loss of charge ligand) was detected but with low abundance. The absence (L1 + H)$^{+}$ ion in any of the CID spectra, taken together with the appearance of (L1 + Na)$^{+}$ and (L1 + K)$^{+}$ ions, suggests that the scFv^{9+} ion is produced from metal ion adducts of (scFv + L1)$^{10+}$

that were not completely removed by quadrupole isolation. These results are consistent with the findings of the BIRD study. [2]

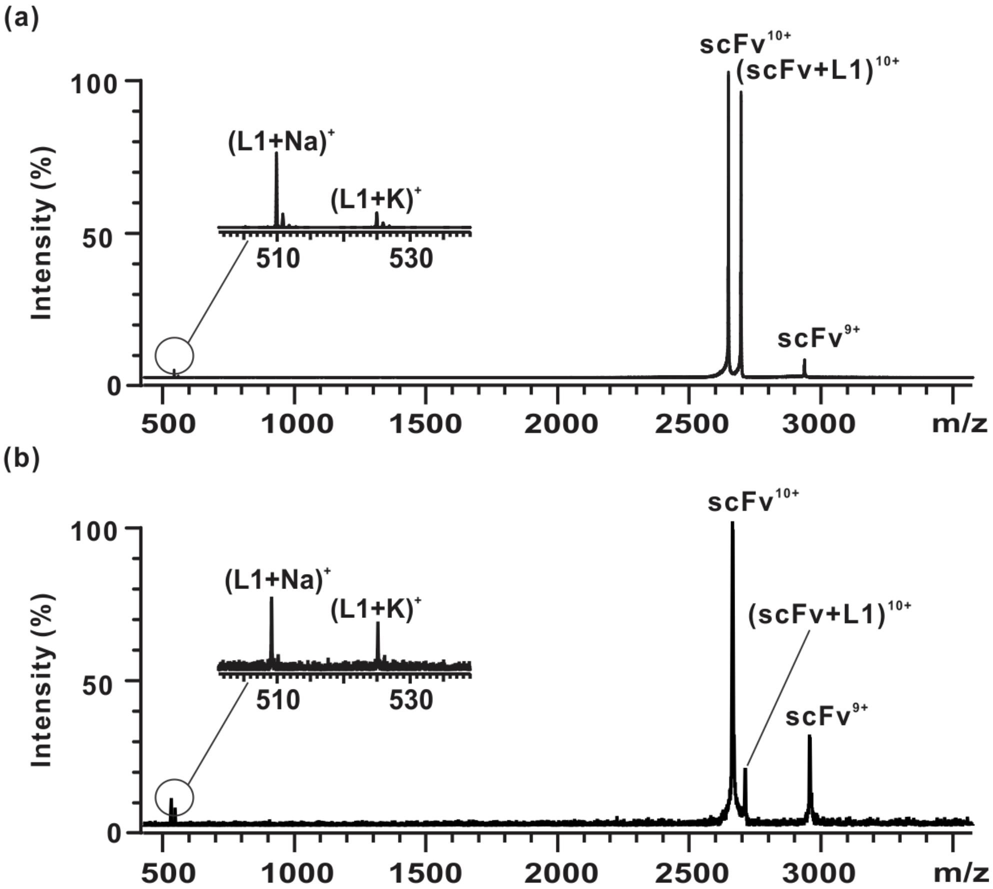


Figure 5.4. (a) CID mass spectrum acquired for the protonated (scFv + L1)$^{10+}$ ion with a collision energy of 240 eV. (b) BIRD mass spectrum obtained for the (scFv + L1)$^{10+}$ ion at reaction temperature of 154 ℃ and a reaction time of 6 s.

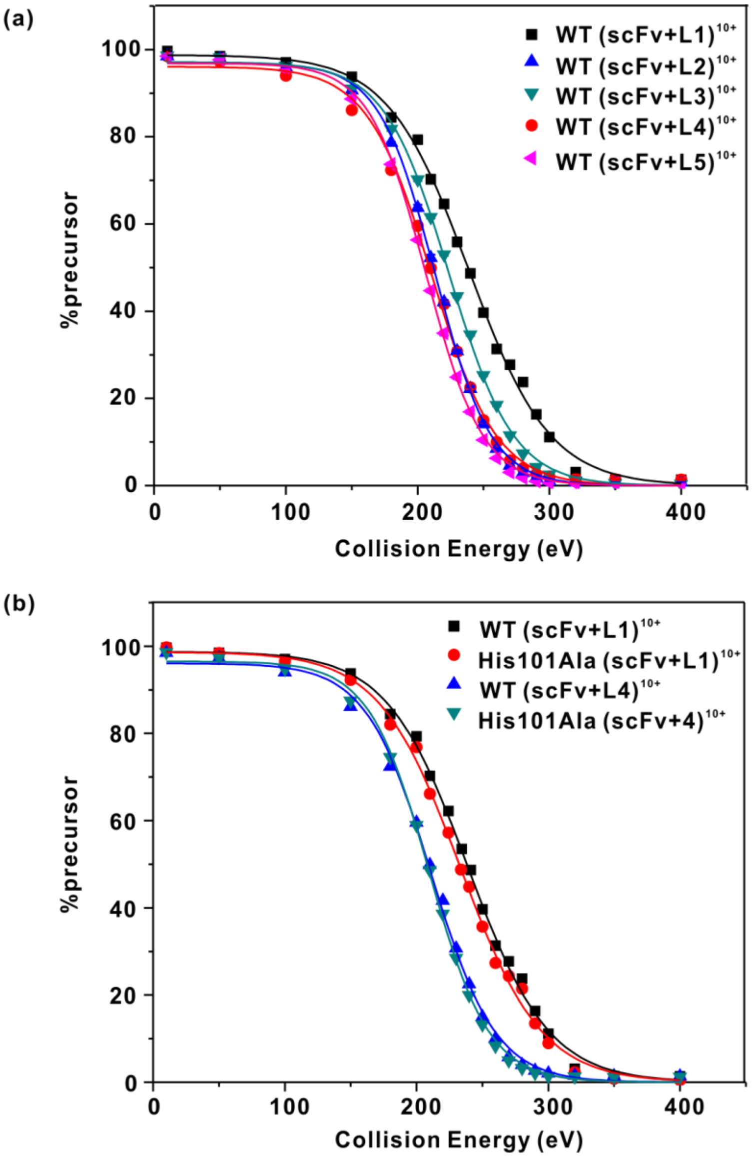


Figure 5.5. (a) CID profiles measured for the protonated WT (scFv + L)$^{10+}$ ions, where L= L1 – L5, at P_{trap} = 6 (b) CID profiles for the protonated His101Ala (scFv + L1)$^{10+}$ and His101Ala (scFv + L4)$^{10+}$ ions at P_{trap} = 6. Also shown are the CID profiles for the WT (scFv + L1)$^{10+}$ and (scFv + L4)$^{10+}$ ions. The solid curves were determined from sigmoid fitting to the collision energy-dependent %precursor values.

Shown in Figure 5.5a are CID dissociation profiles (as a function of collision energy) measured for the (scFv + L)$^{10+}$ ions composed WT scFv and L1 – L5. The Ec$_{50}$ values determined for each of the (scFv + L)$^{10+}$ ions are listed (Table 5.2). For comparison purposes, the corresponding Arrhenius parameters are also included. The stabilizing role of individual ligand hydroxyl groups was established from the difference in Ec$_{50}$ values (ΔEc$_{50}$) determined for (scFv + L1)$^{10+}$ and the corresponding (scFv + L)$^{10+}$ ions containing one of the monodeoxy analogs (L2 – L5) (Table 5.2). In all four cases, deoxygenation of L1 led to a significant decrease in Ec$_{50}$. The corresponding ΔEc$_{50}$ values are: 26.2 ± 1.8 eV (L2), 15.8 ± 1.9 eV (L3), 29.3 ± 2.1 eV (L4) and 33.1 ± 1.8 eV (L5). These results suggest that all four hydroxyl groups participate in intermolecular H-bonds that stabilize the (scFv + L1)$^{10+}$ ion. Importantly, these results are qualitatively consistent with those of the BIRD study, in which deoxygenation at each of the four sites resulted in a measurable decrease in the dissociation E$_a$.[2] However, the trend in ΔEc$_{50}$ values (L5 > L4 > L2 > L3) does not match the trend in ΔE$_a$ values (L5 ≈ L2 > L3 ≈ L4). This finding is discussed in more detail below.

Table 5.2. Comparison of Ec$_{50}$ with Arrhenius parameters (E$_a$, A) determined for the dissociation of (scFv + L)$^{10+}$ ions, where L = L1, L2, L3, L4 or L5, and WT scFv or the single point binding site mutant His101Ala.

scFv	**L**	**Ec_{50} [a]** (eV)	**ΔEc_{50} [b]** (eV)	**E_a [c]** (kcal mol^{-1})	**ΔE_a [d]** (kcal mol^{-1})	**A [c]** (s^{-1})
WT	L1	238.2 ± 1.6	-	54.7 ± 0.6	-	$10^{27.9 \pm 0.7}$
WT	L2	212.0 ± 0.8	26.2 ± 1.8	46.5 ± 0.7	8.2 ± 0.9	$10^{24.2 \pm 0.4}$
WT	L3	222.4 ± 1.0	15.8 ± 1.9	48.5 ± 0.9	6.2 ± 1.1	$10^{25.0 \pm 0.5}$
WT	L4	208.9 ± 1.4	29.3 ± 2.1	49.9 ± 1.0	4.8 ± 1.2	$10^{25.8 \pm 0.6}$
WT	L5	205.1 ± 0.9	33.1 ± 1.8	46.1 ± 0.8	8.6 ± 1.0	$10^{24.0 \pm 0.4}$
His101Ala	L1	232.3 ± 1.8	5.9 ± 2.4	50.0 ± 0.3	4.7 ± 0.7	$10^{25.5 \pm 0.2}$
His101Ala	L4	207.9 ± 1.0	30.3 ± 1.8 (24.4 ± 2.1) [e]	49.4 ± 0.8	5.3 ± 1.0 (0.6 ± 0.9) [f]	$10^{25.4 \pm 0.4}$

a. Errors correspond to the uncertainties of Ec_{50} values (δEc_{50}) that were calculated using error propagation. b. $\Delta Ec_{50} = Ec_{50}(\text{scFv+L1}) - Ec_{50}(\text{scFv+L})$; errors correspond to the uncertainties from error propagation. c. E_a and A values taken from refs [2],[6] and errors correspond to one standard deviation. d. $\Delta E_a = E_a(\text{scFv+L1}) - E_a(\text{scFv+L})$; errors correspond to the uncertainties from error propagation. e. $\Delta Ec_{50} = Ec_{50}(\text{His101Ala+L1}) - Ec_{50}(\text{His101Ala+L4})$; errors correspond to the uncertainties from error propagation. f. $\Delta E_a = E_a(\text{His101Ala+L1}) - E_a(\text{His101Ala+L4})$; errors correspond to the uncertainties from error propagation.

The aforementioned results were acquired using an Ar pressure of 2.1 x 10^{-2} mbar (P_{trap} = 6). It was of interest to assess whether the collision gas pressure has any influence on the trend in Ec_{50} values. To answer this question, measurements were repeated at Ar pressures of 1.5 x 10^{-2} mbar (P_{trap} = 4) and 2.7 x 10^{-2} mbar (P_{trap} =

8). Shown in Figure 5.6a and 5.6b are the CID dissociation profiles measured at P_{trap} 4 and 8, respectively. The Ec_{50} values and corresponding ΔEc_{50} values are listed in Table 5.3. Inspection of the results reveals that the collision gas pressure does have a modest influence the relative magnitude of the ΔEc_{50} values (P_{trap} = 4: L5 ≈ L4 > L2 > L3; P_{trap} = 8: L5 ≈ L4 > L2 ≈ L3). Notably, however, the trend in ΔEc_{50} values established at P_{trap} 4 and 8 does not match the trend in ΔE_a values.

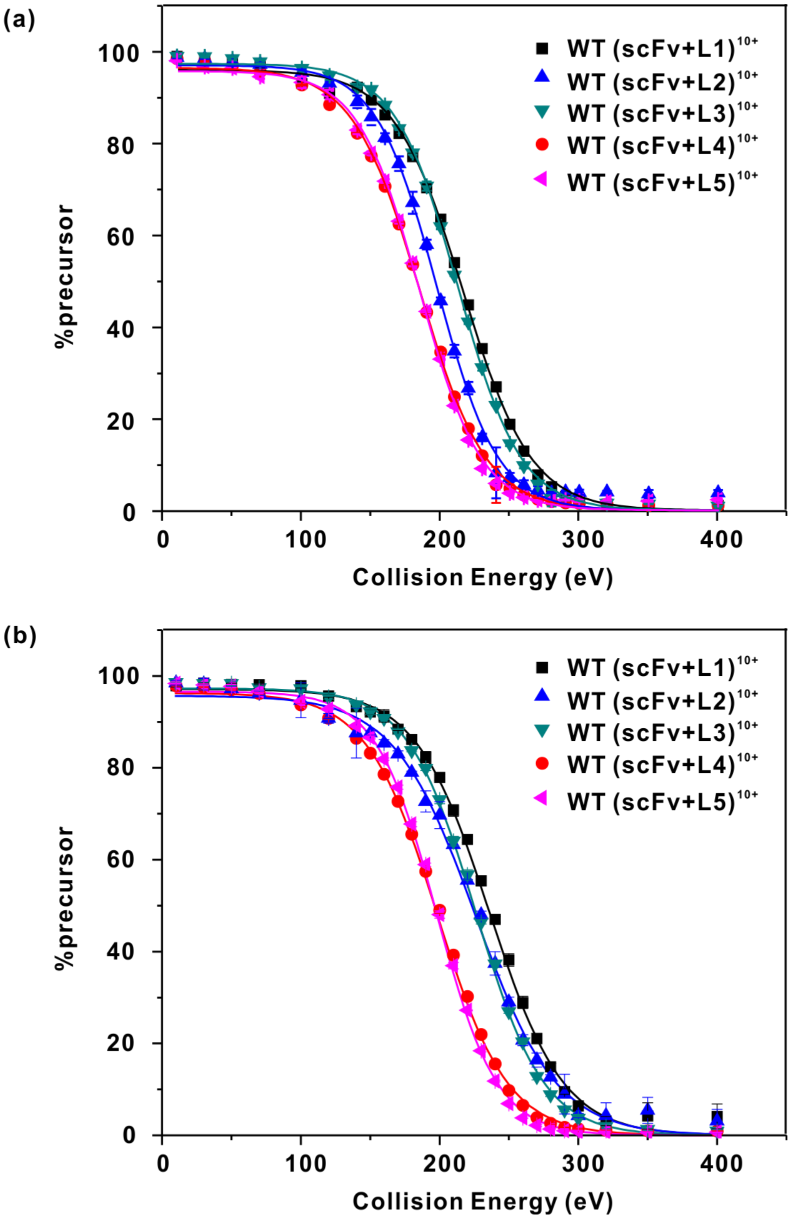


Figure 5.6. (a) CID profiles measured for the protonated WT (scFv + L)$^{10+}$ ions, where L= L1 – L5, at P_{trap} = 4. (b) CID profiles measured for the protonated WT (scFv + L)$^{10+}$ ions, where L= L1 – L5, at P_{trap} = 8. The solid curves were determined from sigmoid fitting to the collision energy-dependent %precursor values.

Table 5.3. Comparison of Ec_{50} values determined for the dissociation of protonated (scFv + L)$^{10+}$ ions at different trap pressure (P_{trap} = 4 (1.5×10^{-2} mbar) and 8 (2.8×10^{-2} mbar)).

L	$\mathbf{P_{trap}}$	$\mathbf{Ec_{50}}$ [a] **(eV)**	$\mathbf{\Delta Ec_{50}}$ [b] **(eV)**
L1	4	213.1 ± 1.0	-
L2	4	195.4 ± 0.6	17.7 ± 1.5
L3	4	210.5 ± 0.6	2.6 ± 1.2
L4	4	182.1 ± 0.9	31.0 ± 1.3
L5	4	182.2 ± 0.8	30.9 ± 1.3
L1	8	235.2 ± 1.0	-
L2	8	224.1 ± 1.7	11.1 ± 1.9
L3	8	225.5 ± 0.6	9.7 ± 1.2
L4	8	197.2 ± 0.8	38.0 ± 1.3
L5	8	197.2 ± 0.7	38.0 ± 1.2

a. Errors correspond to the uncertainties of Ec_{50} values (δEc_{50}) that were calculated using error propagation. b. $\Delta Ec_{50} = Ec_{50}$(scFv+L1) - Ec_{50}(scFv+L).

To perform the CID measurements on the (scFv + L)$^{10+}$ ions composed of the His101Ala mutant, protonated gas-phase ions (scFv + L)$^{n+}$ of both WT and His101Ala were produced simultaneously by ESI performed on aqueous ammonium acetate solutions (10 mM, pH 6.8) of WT scFv (8 μM), His101Ala scFv (10 μM), and L1 or L4 (10-15 μM). Pairs of CID profiles are shown in Figure 5.5b and the corresponding Ec_{50} values are listed (Table 5.2). The Arrhenius parameters are also included. Comparison of the dissociation profiles for the (scFv + L1)$^{10+}$ ions composed of WT and His101Ala reveals that mutation of His101 results in albeit small (5.9 eV) reduction in Ec_{50}, suggesting that the side chain of the residue is involved in stabilizing the complex (presumably through an intermolecular H-bond). It was previously shown using the BIRD-FGR method that the side chain of His101 is engaged in a H-bond with the OH group at the C4 position of Man. [2] Notably, this interaction is also identified in the crystal structure of the (scFv + L1) complex. [14] Interestingly, deoxygenation at the Man C4 position results in a further (and significant) decrease in Ec_{50} (by 24.4 eV). Taken together, these results suggest that both His101 and Man C4 OH stabilize the complex, but not through a common interaction(s). In contrast, similar ΔE_a values were obtained from BIRD measurements performed on the (scFv + L1)$^{10+}$ ions upon deoxygenation (at Man C4) or mutation of His101 (to Ala) and upon simultaneous deoxygenation and mutation. [2]

The CID results obtained for the (scFv + L)$^{10+}$ ions results reveal that the loss of intermolecular H-bonds in the (scFv + L)$^{10+}$ ion results in measurable decreases in the Ec_{50} values. However, the magnitude of the changes in Ec_{50}

values (i.e., ΔEc_{50}) does not reflect the relative E_a values. The lack of agreement between the relative Ec_{50} and E_a values can be explained by the fact that a CID mass spectrum reflects the microcannonical dissociation rate constants convoluted over the internal energy distribution (which is expected to be non-Boltzmann in nature) of the reactant ion at a given collision energy. [23-24] At low internal energies the relative ordering of rate constants (for structurally related complexes) will tend to reflect the trend in dissociation E_a, while at high internal energies, the ordering of rate constants will be dominated by the trend in the pre-exponential factors. In other words, the ΔEc_{50} value reflects the dissociation kinetics over a range of internal energy distributions. It is, therefore, not surprising that there is no temperature at which the thermal dissociation rate constants (Figure 5.7) would lead to the observed trend of Ec_{50} values for the $(scFv + L)^{10+}$ ions where L = L1 – L5.

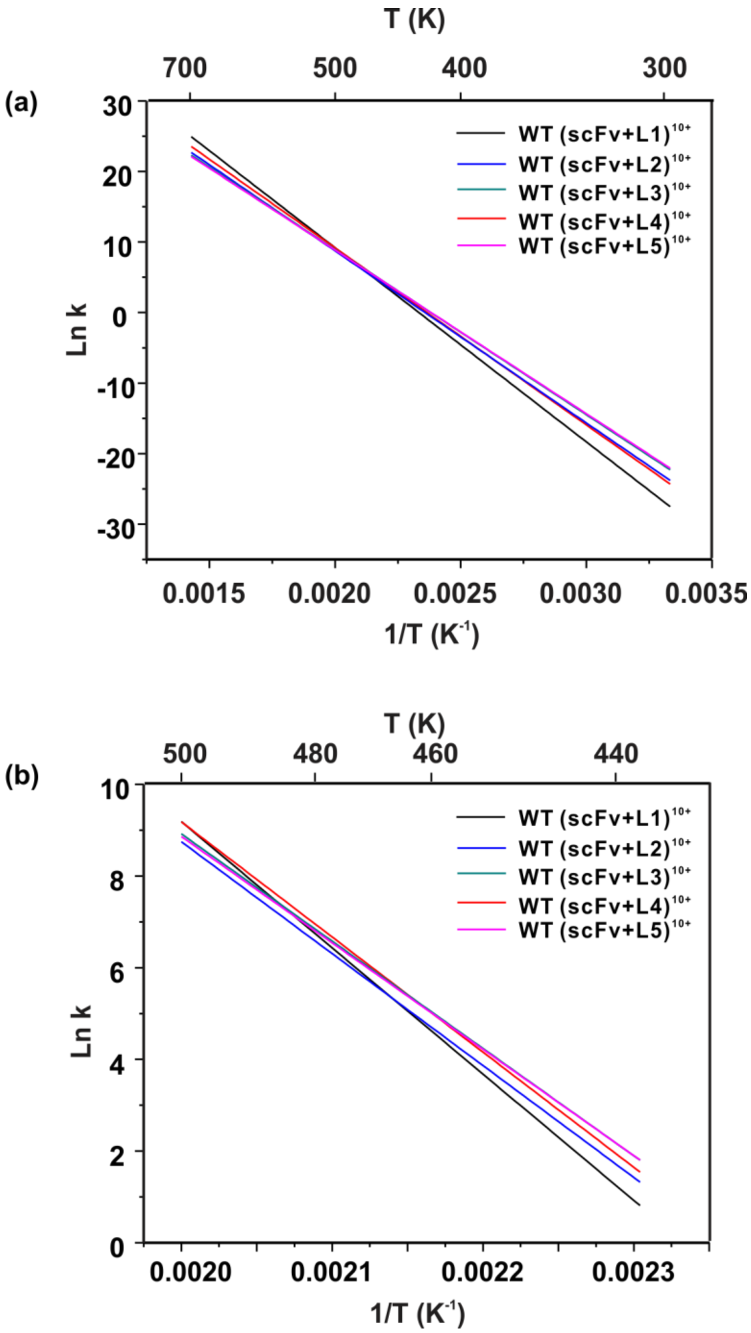


Figure 5.7. Arrhenius plots for the dissociation of WT (scFv + L)$^{10+}$ ions, where L= L1 – L5, calculated from the activation energy and A values reported in [2] in the temperature range (a) 300 K to 700K and (b) 440 K to 500 K.

Moreover, the observation that the magnitude of the ΔEc_{50} values measured upon disruption of the His101/Man C4-OH H-bond is sensitive to the nature of the modification (i.e., mutation versus deoxygenation) is evidence that modifications that are energetically equivalent (i.e., give rise to the same change in E_a), are not necessarily equivalent in terms of Ec_{50} values. The larger decrease in Ec_{50} upon deoxygenation of L1 at Man C4, compared to mutation of His101, is consistent with the reported differences in the kinetic stabilities of the WT $(scFv + L4)^{10+}$ and His101Ala $(scFv+L4)^{10+}$ ions (Figure 5.8). These kinetic differences are entropic in origin (since the E_a values are indistinguishable) and are believed to reflect the participation of the Man C4 OH group in intramolecular H-bond in the dissociative *TS*, which serves to rigidify the ligand (and decrease the entropy of activation) relative to the (scFv + L4) complex. [2]

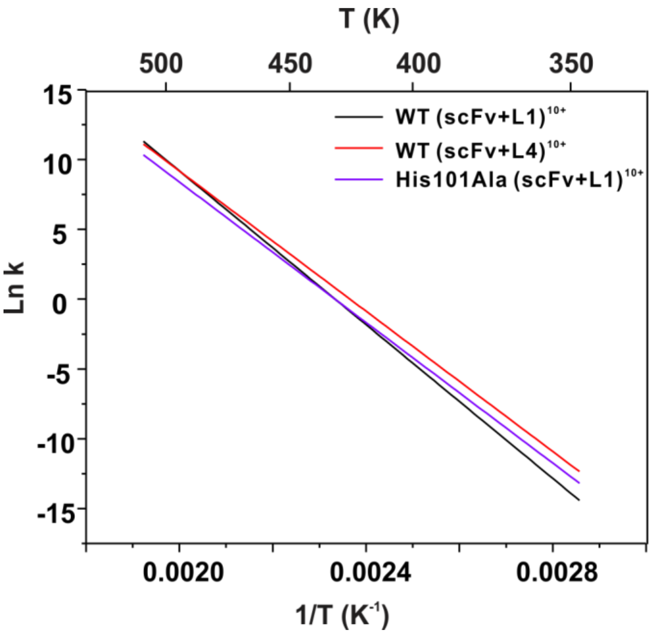


Figure 5.8. Arrhenius plots for the dissociation of WT (scFv + L1)$^{10+}$ ions, WT (scFv + L4)$^{10+}$ ions and His101Ala (scFv + L1)$^{10+}$ ions, calculated from the activation energy and A values reported in [2] in the temperature range 350 K to 520K.

5.3.2 Mapping the intermolecular interactions in the (S4 + 4B)$^{13+}$ ion

To further test the reliability of the CID-FGR strategy, the method was used to probe the intermolecular interactions in protonated (S_4 + 4B)$^{13+}$ ions. Protonated (S_4 + 4B)$^{n+}$ ions, at n = 12 to 17, were produced by ESI from aqueous ammonium acetate solutions (10 mM, pH 6.8) containing WT S_4 (8 μM) and one of the binding site single point mutants (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala or Tyr43Ala) S_4 (10 μM), B (80 μM) and imidazole (5 mM). An illustrative ESI mass spectrum is shown in Figure 5.9 for a solution containing both WT and Tyr43Ala and B. As reported elsewhere 10, the Ω values for the protonated WT $(S_4 + 4B)^{n+}$ ions, where n = 12 to 16, are independent of charge state, with an average value 35.0 ± 0.1 nm^2. This value agrees within 10% with the Ω value (38.4 ± 0.4 nm^2) estimated for the crystal structure of the $(S_4 + 4B)$ complex. Based on these results it was suggested that the transfer of $(S_4 + 4B)$ from neutral aqueous solution to the gas phase by ESI is not accompanied by large structural changes.

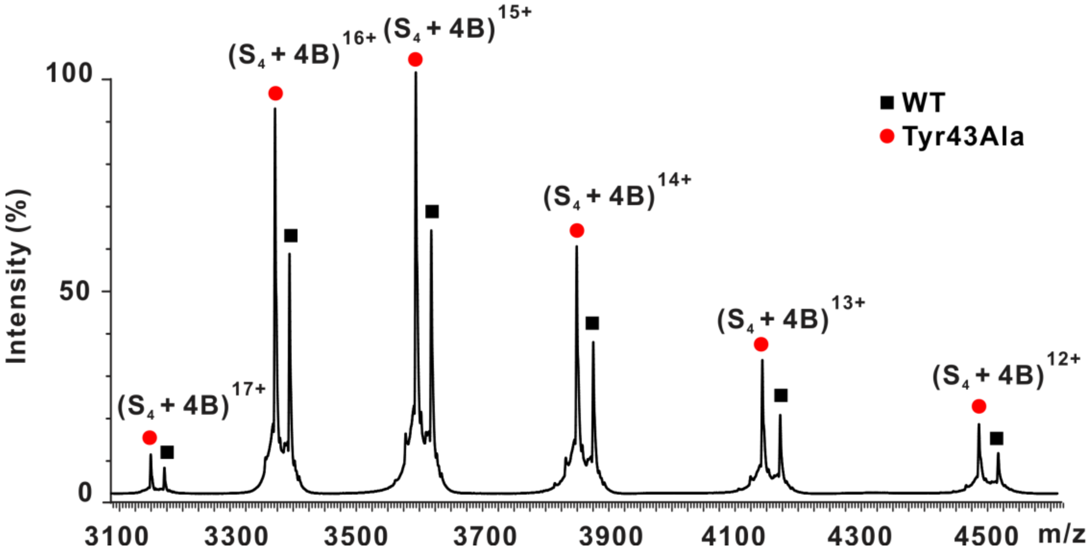


Figure 5.9. ESI mass spectrum acquired for a neutral aqueous ammonium acetate (10 mM) solution of WT S_4 (8 μM), Tyr43Ala S_4 (10 μM), B (80 μM) and imidazole (5 mM).

The +13 charge state of the $(S_4 + 4B)$ complexes was selected for the CID measurements. Representative CID mass spectra for the WT and Tyr43Ala $(S_4 + 4B)^{13+}$ ions are shown in Figure 5.10. CID of the $(S_4 + 4B)^{13+}$ ions resulted predominantly in the loss of neutral B; the loss of protonated B was observed as a minor pathway. Subsequent dissociation of the primarily product ions, $(S_4 +$

$3B)^{13+}$ and $(S_4 + 3B)^{12+}$, proceeded in a similar fashion, such that a series of $(S_4 + iB)^{13+}$ and $(S_4 + iB)^{12+}$ ions were evident in the CID mass spectra. The CID dissociation profiles (for pairs of $(S_4 + 4B)^{13+}$ ions composed of WT and each single point mutant) are shown in Figure 5.11. The Ec_{50} values are reported in Table 5.4, along with the corresponding ΔEc_{50} values, as well as the Arrhenius parameters [10].

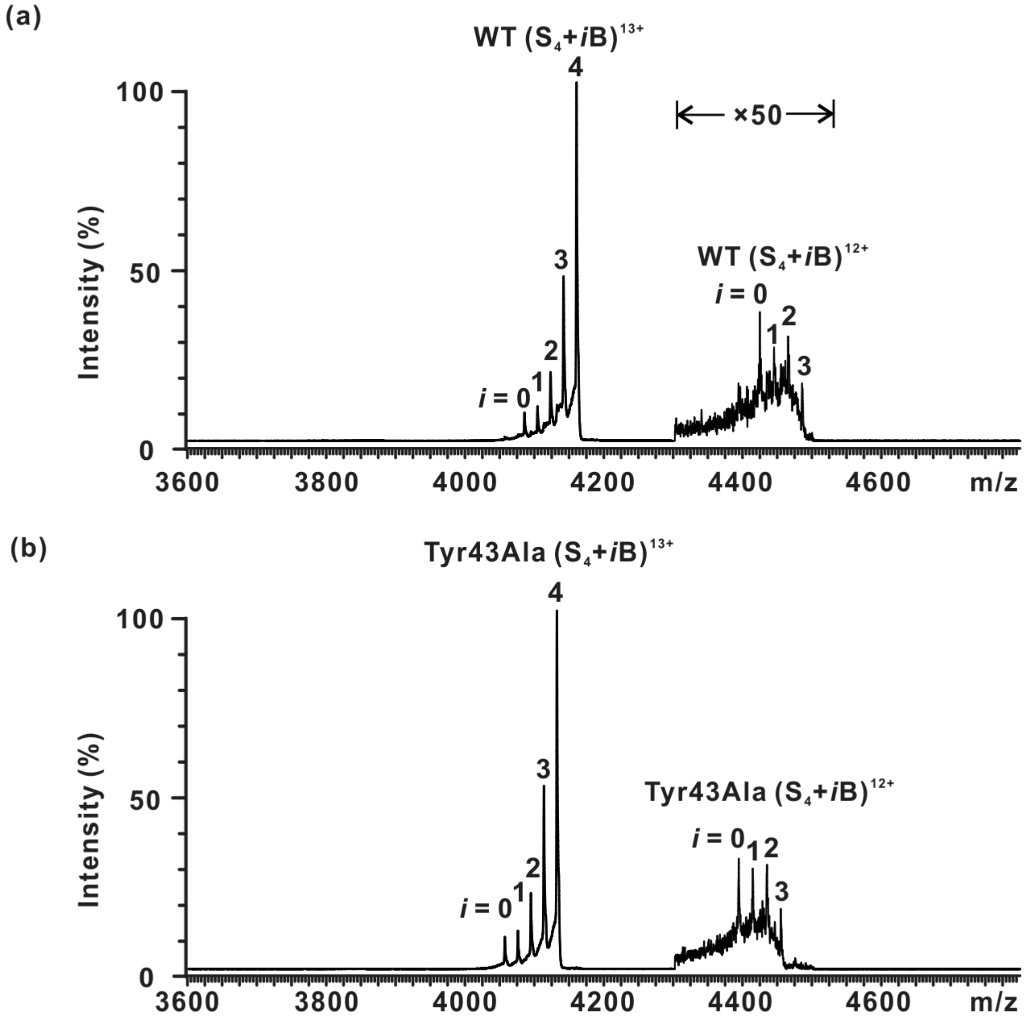


Figure 5.10. CID mass spectrum acquired for protonated $(S_4 + 4B)^{13+}$ ions, composed of (a) WT or (b) Tyr43Ala S_4, with a collision energy of 390 eV. The region from *m/z* 4300-4500 was magnified by a factor of 50.

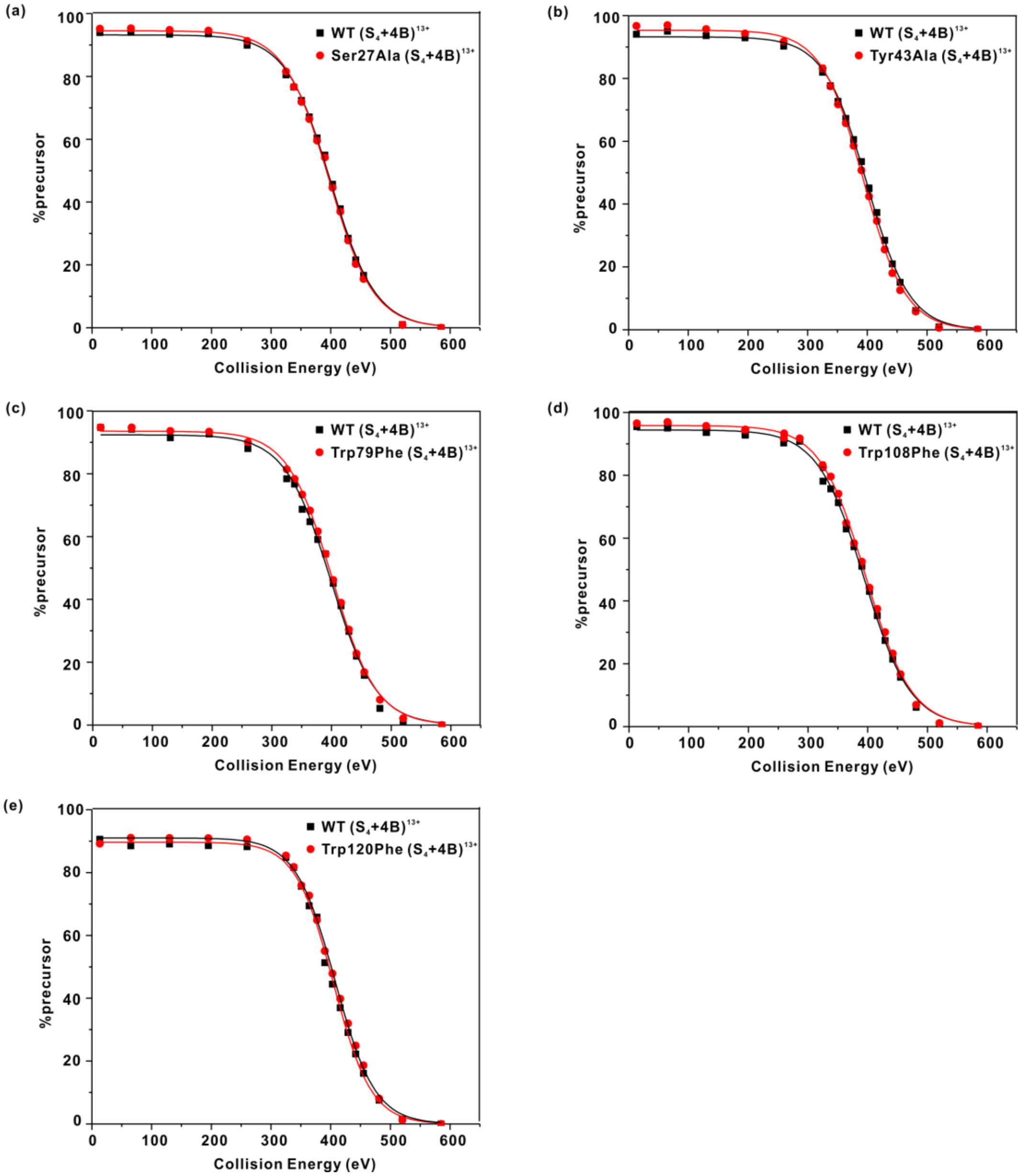


Figure 5.11. CID profiles for pairs of protonated $(S_4 + 4B)^{13+}$ ions involving WT and (a) Ser27A, (b) Tyr43Ala, (c) Trp79Phe, (d) Trp108Phe, and (e) Trp120Phe. The solid curves were determined from sigmoid fitting to the collision energy-dependent %precursor values.

Table 5.4. Comparison of Ec_{50} with Arrhenius parameters (E_a, A) determined for the loss of neutral B from the protonated $(S_4 + 4B)^{13+}$ ions, where S_4 = WT, Ser27Ala, Tyr43Ala Trp79Phe, Trp108Phe, or Trp120Phe.

S_4	Ec_{50} [a] (eV)	ΔEc_{50} [b] (eV)	E_a [c] (kcal mol^{-1})	ΔE_a [d] (kcal mol^{-1})	A [c] (s^{-1})
WT_1	394.2 ± 1.4	-	43.2 ± 1.1	-	$10^{23.2 \pm 0.6}$
Ser27Ala	392.7 ± 1.4	1.6 ± 1.9	38.5 ± 1.0	4.7 ± 1.4	$10^{20.7 \pm 0.6}$
WT_2	393.7 ± 1.3	-		-	
Tyr43Ala	389.6 ± 1.3	4.0 ± 1.9	37.8 ± 0.9	5.4 ± 1.4	$10^{20.4 \pm 0.5}$
WT_3	392.3 ± 2.3	-		-	
Trp79Phe	395.9 ± 1.1	-3.6 ± 2.5	40.7 ± 0.3	2.5 ± 1.1	$10^{21.8 \pm 0.2}$
WT_4	389.3 ± 1.7	-		-	
Trp108Phe	393.2 ± 1.5	-3.9 ± 2.3	44.8 ± 0.5	-1.6 ± 1.2	$10^{24.1 \pm 0.3}$
WT_5	397.2 ± 1.4	-		-	
Trp120Phe	400.9 ± 1.1	-3.7 ± 1.7	40.6 ± 0.4	2.6 ± 1.2	$10^{21.8 \pm 0.2}$

a. Errors correspond to the uncertainties of Ec_{50} values (δEc_{50}) that were calculated using error propagation. b. $\Delta Ec_{50} = Ec_{50}(WT) - Ec_{50}(mutant)$; errors correspond to the uncertainties from error propagation. c. E_a and A values taken from ref 10; errors correspond to one standard deviation. d. $\Delta E_a = E_a(WT) - E_a(mutant)$; errors correspond to the uncertainties from error propagation.

Inspection of the values in Table 5.4 reveals that mutation of Tyr43 results in a small (4.0 eV) decrease in Ec_{50}, suggesting that this residue stabilizes the gaseous $(S_4 + 4B)^{13+}$ ion. This finding is consistent with the result of the BIRD-FGR mapping study. [10] In contrast, mutation of Ser27 produced no significant change in Ec_{50}. The apparent absence of an interaction at Ser27 is inconsistent with the BIRD-FGR results, which revealed a strong (~5 kcal mol^{-1}) interaction.[10] Unexpectedly, mutation of Trp79, Trp108 and Trp120 results in a small increase (3 – 4 eV) in Ec_{50}, suggesting that the mutations enhance the energetic stability of the complex. However, according to the BIRD results, replacement of Trp79 or Trp120 with Phe leads to a decrease in E_a (~2.5 kcal mol^{-1}) [10], while mutation of Trp108 (to Phe) does not change the E_a value. Taken together, these results reveal that the CID-FGR approach, when implemented using binding site mutants, lacks reliability and is prone to both false positives and false negatives.

5.4 Conclusions

The present study represents the first direct comparison of the CID-FGR and BIRD-FGR methods for mapping intermolecular interactions in the gaseous ions of protein-ligand complexes. The results obtained for protonated ions of the (scFv + L1) and $(S_4 + 4B)$ complexes demonstrate conclusively that structural modification that lead to changes in the dissociation E_a do not necessarily lead to measurable changes in the Ec_{50} values. Similarly, changes in Ec_{50} are not always indicative of the loss of intermolecular interactions. Therefore, while the CID-FRG approach can, in some instances, correctly identify residues or functional groups that stabilize the complex, the method is prone to both false positives and false negatives. It was also shown that trends in Ec_{50} values for structurally related complexes do not necessarily correlate the trend in E_a values or thermal rate constants. Moreover, protein mutation and ligand modification do not necessarily produce equivalent effects in terms of Ec_{50} values. The discrepancies between the CID-FGR and BIRD-FGR results can be explained qualitatively by the fact the CID dissociation curves reflect the dissociation rate constants, which are functions of both energy and entropy, convoluted over the internal energy distributions achieved over a range of collision energies. Taken together, the results of this study suggest that the CID-FGR method, as implemented here, does not represent a reliable approach for identifying intermolecular interactions in the gaseous ions of protein-ligand complexes.

5.5 References

1. Kitova, E.; El-Hawiet, A.; Schnier, P.; Klassen, J. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431-441.

2. Kitova, E. N.; Seo, M.; Roy, P.-N.; Klassen, J. S. *J. Am. Chem. Soc.* **2008**, *130*, 1214-1226.

3. Price, W. D.; Schnier, P. D.; Jockusch, R. A.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1996**, *118*, 10640-10644.

4. Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.

5. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *J. Am. Chem. Soc.* **2002**, *124*, 5902-5913.

6. Kitova, E. N.; Bundle, D. R.; Klassen, J. S. *Angew. Chem. Int. Ed.* **2004**, *43*, 4183-4186.

7. Liu, L.; Bagal, D.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2009**, *131*, 15980-15981.

8. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2010**, *132*, 17658-17660.

9. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 3054-3060.

10. Deng, L.; Broom, A.; Kitova, E. N.; Richards, M. R.; Zheng, R. B.; Shoemaker, G. K.; Meiering, E. M.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 16586-16596.

11. McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614.

12. Tešić, M.; Wicki, J.; Poon, D. K. Y.; Withers, S. G.; Douglas, D. J. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 64-73.

13. Hopper, J. T. S.; Oldham, N. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1851-1858.

14. Zdanov, A.; Li, Y.; Bundle, D. R.; Deng, S. J.; MacKenzie, C. R.; Narang, S. A.; Young, N. M.; Cygler, M. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6423-6427.

15. Bundle, D. R.; Baumann, H.; Brisson, J.-R.; Gagne, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183-5192.

16. Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1754-1758.

17. Wang, W.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2005**, *77*, 3060-3071.

18. Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2006**, *79*, 416-425.

19. Hopper, J.; Rawlings, A.; Afonso, J.; Channing, D.; Layfield, R.; Oldham, N. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1757-1767.

20. Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. *Anal. Chem.* **2010**, *82*, 9557-9565.

21. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protocols* **2008**, *3*, 1139-1152.

22. Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *The Journal of Physical Chemistry* **1996**, *100*, 16082-16086.

23. Vékey, K. *J. Mass Spectrom.* **1996**, *31*, 445-463.

24. Mayer, P. M.; Poon, C. *Mass Spectrom. Rev.* **2009**, *28*, 608-639.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This work describes the development and application of ESI-MS methods to study the structure and physicochemical properties of protein complexes in the gas and condensed phases. Research efforts focused on the quantification of the stoichiometry and absolute affinity of protein-metal ion binding in solution (Chapter 2), the quantification of the dissociation kinetics of the protein-ligand interaction in solution (Chapter 3), and investigation of the structure and stability of protein-ligand ions in the gas phase (Chapters 4 and 5).

In Chapter 2, the application of the direct ESI-MS assay for quantifying the stoichiometry and absolute affinity of protein-metal ion binding *in vitro* was described. Control ESI-MS experiments performed on solutions containing calcium chloride or calcium acetate and a pair of proteins that do not bind calcium ions in solution revealed that the nonspecific association of metal ions to proteins during ESI is a random process, independent of protein size and structure. These results establish the reliability of the reference protein method for quantitatively correcting ESI mass spectra for the occurrence of nonspecific metal ion binding to proteins during ESI-MS analysis. To demonstrate the utility of the direct ESI-MS assay, when carried out using the reference protein method, the calcium binding stoichiometry of bovine α-lactalbumin and the calcium ion affinity of bovine β-lactoglobulin were established.

Work in Chapter 3 demonstrated the quantification of k_{off} for the model high affinity interaction between B and S_4 at pH 7 and temperatures ranging from 15 to 45 ℃ using ESI-MS. Two different approaches to data analysis were employed, one based on the initial rate of dissociation of the (S_4 + 4B) complex, the other involving non-linear fitting of the time-dependent relative abundances of the (S_4 + *i*B) species. The two methods were found to yield k_{off} values that are in good agreement, within a factor of two. The Arrhenius parameters for the dissociation of the biotin-streptavidin interaction in solution were established from the k_{off} values determined by ESI-MS and compared to values measured using a radiolabeled biotin assay. Importantly, the dissociation activation energies determined by ESI-MS agree, within 1 kcal mol^{-1}, with the reported value[1]. In addition to providing a quantitative measure of k_{off}, the results of the ESI-MS measurements revealed that the apparent cooperative distribution of (S_4 + *i*B) species observed at short reaction times is of kinetic origin and that sequential binding of B to S_4 occurs in a non-cooperative fashion with the four ligand binding sites being kinetically and thermodynamically equivalent and independent.

Chapters 4 and 5 focus on the study of the structure and kinetic stability of gaseous protein-ligand complexes ions. In chapter 4, the BIRD-FGR approach was employed on the model high-affinity protein-ligand interaction between B and S_4. A comparison of E_a measured for the loss of B from the $(S_4 + 4B)^{13+}$ ions composed of WT streptavidin and five binding site mutants (Trp79Phe, Trp108Phe, Trp120Phe, Ser27Ala and Tyr43Ala) suggests that at least some of the specific intermolecular interactions are preserved in the gas phase. The results of MD simulations performed on WT $(S_4 + 4B)^{12+}$ ions with different charge configurations support this conclusion. The most significant finding here is that the gaseous WT $(S_4 + 4B)^{n+}$ ions at n = 12 – 14, owing to a much larger E_a (by as much as 13 kcal mol^{-1}) for the loss of B, are dramatically more stable kinetically at 25 ℃ than the $(S_4 + 4B)$ complex in aqueous neutral solution[1]. The differences in E_a values measured for the gaseous $(S_4 + 4B)^{n+}$ ions and solvated $(S_4 + 4B)$ complex can be largely accounted for by a late dissociative *TS* and the rehydration of B and the protein binding cavity in solution. In chapter 5, intermolecular interactions in the gaseous ions of two protein-ligand complexes, $(scFv + L1)^{10+}$ and $(S_4 + 4B)^{13+}$, were investigated using a CID-FGR strategy. Intermolecular interactions were identified based on decreases in the collision energy required to dissociate 50% of the reactant ion (i.e., the Ec_{50} value) upon modification of protein residues or ligand functional groups. These results were compared with interaction maps generated using the BIRD-FGR method.[2,3] For the $(scFv + L1)^{10+}$ ion, it was found that deoxygenation of L1 (at the Gal C3 and C6 and Man C4 and C6 positions) or mutation of His101 (to Ala) resulted in decreases in the Ec_{50} values. These results suggest that the four hydroxyl groups, as well as His101, participate in intermolecular H-bonds. These findings are consistent with those obtained using BIRD-FGR method. However, the CID-FGR method failed to reveal the relative strengths of the intermolecular interactions or establish Man

C4 OH and His101 as a H-bond donor/acceptor pair. The CID-FGR method correctly identified Tyr43 as a stabilizing contact in the $(S_4 + 4B)^{+13}$ ion, but failed to identify Trp79, Ser27 and Trp120 as being involved in stabilizing interactions. In fact, mutation of Trp79 and Trp120, led to an increase in the Ec_{50} value. Taken together, the results of this study suggest that the CID-FGR method, as implemented here, does not represent a reliable approach for identifying intermolecular interactions in the gaseous ions of protein-ligand complexes.

6.2 Future work

There are several possible extensions of the current studies. As mentioned in Chapter 2, the results demonstrated that the direct ESI-MS assay, when implemented with the reference protein method, can serve as a reliable technique to establish the stoichiometry of calcium and, presumably, other metals binding to proteins in solution and to quantify protein-metal ion affinities. Metal binding to carbohydrate binding proteins is of particular interest, especially where metal binding is essential for protein-carbohydrate interactions[4].

Another extension to the binding study of protein-ligand interactions by ESI-MS is the library screening of potential binders to a target protein. Recently, our group has demonstrated the application of a catch-and-release (CaR)-ESI-MS approach for screening libraries of carbohydrates against target proteins.[5,6] Briefly, direct ESI-MS measurements were performed on solutions containing a target protein and a library of carbohydrates. Protein bound ligands were identified from the detection of protein-ligand complex iones in a single mass spectrum. Additionally, the absolute affinities were estimated from the abundance of free and ligand bound protein ions with the use of a reference protein method to correct the nonspecific binding during the ESI process. The release of the carbohydrate ligands using CID performed on the deprotonated protein-carbohydrate complexes, followed by IMS performed on the released carbohydrate ions, allowed for the positive identification of isomeric ligands. In particular, the CaR-ESI-MS approach could be used for screening libraries of potential inhibitors against a target protein, P38 mitogen-activated protein (MAP) kinase. The P38 MAP kinase plays a crucial role in regulating the production of proinflammatory cytokines, such as tumor necrosis factor and interleukin.[7-9] Blocking this kinase may offer an effective therapy for treating many inflammatory diseases. This work is carried out in collaboration with Dr. Paul Schnier, Amgen, CA. Unfortunately, preliminary ESI-MS analysis showed no P38-inhibitor complex ions by direct spraying three P38 inhibitors (K_a values ranging 10^7-10^8 M^{-1} measured by SPR, unpublished results) with P38 and a reference protein (scFv) together. Two possible reasons for this are the loss of P38 binding activity to inhibitors in the ESI solution buffer (50 mM AmAc, pH 7) which is different from buffer used in SPR (Tris-Cl, pH 7), and/or the in-source dissociation of the P38-inhibitor complexes. To support this, SPR experiments will be carried out to determine the binding affinity of P38 to the inhibitors in 50 mM AmAc to check whether binding is preserved in these solution conditions. If SPR data shows that the loss of binding occurs in the AmAc solution, ESI-MS binding assay could be modified to facilitate production of the protein and ligand complexes ions from non-volatile buffer (e.g. Tris-Cl). In this case, a desorption electrospray ionization (DESI) source[10] which allows the use of nonvolatile buffer can be applied. In order to eliminate in-source dissociation, a series of small molecules[11], such as imidazole, sorbitol, monosaccharides, disaccharides, and amino acids, will be added to the ESI solutions containing P38 and the inhibitor to test whether they can assist with gas phase stabilization of the P38-inhibitor complexes. After quantification of the binding affinity of P38 and known inhibitors are achieved, we will explore the application of CaR-ESI-MS approach for screening libraries of potential inhibitors against P38. At present, we have a potential inhibitors library composed of 500 organic small molecules. The hits (binders) could be identified by the appearance of protein-inhibitor complexes ions in the mass spectra after correction for the nonspecific binding. The binding affinities of a specific inhibitor to P38 could be quantified from the abundance of free and ligand bound protein ions after correction for the nonspecific binding. Additionally, isomeric ligands could be separated with the use of IMS, performed on the ligands following their release from the complex.

In Chapter 3, we demonstrated the quantification of the dissociation kinetics of the model high-affinity protein-ligand interaction between B and S_4 using the direct ESI-MS analysis. The streptavidin-biotin interaction is unusually kinetically stable and it takes several days to achieve an equilibrium distribution of (S_4 + iB) species at the temperatures investigated. However, it is important to note that this same experimental approach could be applied, in an on-line fashion, to determine k_{off} and, in principle, k_{on} for protein-ligand interactions that require less time to reach an equilibrium distribution. An on-line nanoESI device will be applied in the group to facilitate kinetic studies by ESI-MS on the ms – s scale. A nano-flow pump will be used to deliver protein and ligand solutions to the reaction chamber, where the reaction is initiated. The mixture can be directly delivered to ESI source by a "reaction capillary"[12] or transferred to a second chamber where the reaction is quenched prior to ESI analysis. The reaction time can be controlled by the length of the reaction capillary and solution flow rate.

To date, there are three protein-ligand complexes studied by the BIRD-FGR approach.[2,3,13] All of them have demonstrated the power of this approach to identify a certain intermolecular interaction and quantify its strength as well as provide the insight into solvent effect. In the future, the BIRD-FGR approach could be used to investigate the effect of glycosylation on the protein-ligand binding process. Glycosylation is an important and highly regulated mechanism of secondary protein processing within cells. It plays a critical role in determining protein structure, function and stability.[14] Of particular interest to the development of new therapies is the role that glycosylation plays in affecting protein-ligand interactions. The presence of glycosylation has been shown to modulate intermolecular association between protein ligands and their receptors[15,16], but their function is less well understood. The gas phase stability of complexes composed of avidin/neutravidin and biotin will be tested using the BIRD technique. A comparison of Arrhenius parameters for the dissociation of biotin from the avidin-biotin complexes with those of neutravidin-biotin complexes can provide an insight into the glycosylation effect on the avidin biotin interaction.

6.3 References

1. Klumb, L. A.; Chu, V.; Stayton, P. S. *Biochemistry* **1998**, *37*, 7657-7663.

2. Kitova, E. N.; Seo, M.; Roy, P.-N.; Klassen, J. S. *J. Am. Chem. Soc.* **2008**, *130*, 1214-1226.

3. Deng, L.; Broom, A.; Kitova, E. N.; Richards, M. R.; Zheng, R. B.; Shoemaker, G. K.; Meiering, E. M.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 16586-16596.

4. Balzarini, J. *Nat. Rev. Micro.* **2007** *5*, 583-597.

5. El-Hawiet, A.; Shoemaker, G. K.; Daneshfar, R.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2011**, *84*, 50-58.

6. El-Hawiet, A.; Kitova, E. N.; Klassen, J. S. *Biochemistry* **2012**, *51*, 4244-4253.

7. Pargellis, C.; Tong, L.; Churchill, L.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Grob, P. M.; Hickey, E. R.; Moss, N.; Pav, S.; Regan, J. *Nature Structural Biology* **2002**, *9*, 268-272.

8. Wilson, K. P.; McCaffrey, P. G.; Hsiao, K.; Pazhanisamy, S.; Galullo, V.; Bemis, G. W.; Fitzgibbon, M. J.; Caron, P. R.; Murcko, M. A.; Su, M. S. S. *Chem. Biol.* **1997**, *4*, 423-431.

9. Tong, L.; Pav, S.; White, D. M.; Rogers, S.; Crane, K. M.; Cywin, C. L.; Brown, M. L.; Pargellis, C. A. *Nature Structural Biology* **1997**, *4*, 311-316.

10. Ferguson, C. N.; Benchaar, S. A.; Miao, Z.; Loo, J. A.; Chen, H. *Anal. Chem.* **2011**, *83*, 6468-6473.

11. Sun, J.; Kitova, E. N.; Klassen, J. S. *Anal. Chem.* **2006**, *79*, 416-425.

12. Konermann, L.; Rosell, F. I.; Mauk, A. G.; Douglas, D. J. *Biochemistry* **1997**, *36*, 6448-6454.

13. Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. *J. Am. Chem. Soc.* **2012**, *134*, 3054-3060.

14. Funakoshi, Y.; Suzuki, T. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2009**, *1790*, 81-94.

15. Margraf-Schönfeld, S.; Böhm, C.; Watzl, C. *J. Biol. Chem.* **2011**, *286*, 24142-24149.

16. Lee, K.-H.; Ahn, J.-I.; Yu, D.-H.; Jeong, H.-S.; Lee, S.-H.; Kim, K.-S.; Chung, I.-Y.; Kim, J.-H.; Lee, Y.-S. *Biochem. Biophys. Res. Commun.* **2001**, *286*, 707-713.